國立交通大學

光電工程研究所

碩士論文

複合式奈米孔洞搭配表面微結構之光穿透增益研究

Extraordinary Transmission Enhancement in Corrugated Composite Nano-aperture

研究生：洪偉琮

指導教授：田仲豪　助理教授

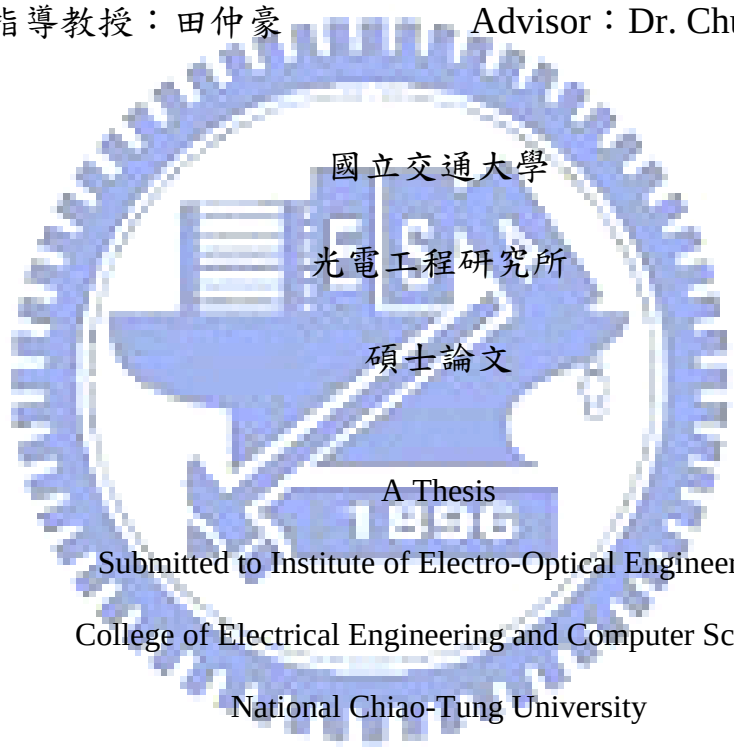
中華民國九十七年七月

複合式奈米孔洞搭配表面微結構之光穿透增益研究

Extraordinary Transmission Enhancement in Corrugated Composite Nano-aperture

研 究 生：洪偉琮　　Student：Wei-Chung Hung

指導教授：田仲豪　　Advisor：Dr. Chung-Hao Tien



國立交通大學

光電工程研究所

碩士論文

A Thesis

Submitted to Institute of Electro-Optical Engineering

College of Electrical Engineering and Computer Science

National Chiao-Tung University

in Partial Fulfillment of the Requirements

for the Degree of Master

in

Electro-Optical Engineering

July 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年七月

複合式奈米孔洞搭配表面微結構之光穿透增益研究

碩士研究生：洪偉琮　　指導教授：田仲豪 助理教授

國立交通大學

光電工程研究所

摘要

隨著光學儲存技術的進步,大量的光學儲存容量需求也與日俱增.這個需求也驅使著研究方向朝著有高儲存密度的微小光點發展.通常是使用開口端微小的奈米探針把光於近場範圍內侷限在超越繞射極限的奈米尺度內.這種方法的最大缺陷就是出光強度太弱不利於應用.Bethe 對於尺度遠小於入射波長的奈米孔洞提供了清楚的物理圖像.穿透孔洞的光強度與孔洞尺寸的四次方成正比,也就是說穿過小尺度孔洞的微小光點通常都是非常微弱的.

然而 Bethe 的理論孔洞建立在不符合實際的數學假設上:厚度可以忽略的完美金屬薄膜.實際的條件是違反這個假設.因此,奈米孔洞的光穿透增益是可能產生的此論文中由以下三種機制來探討奈米孔洞的特性:(1)表面結構決定的攫取光子機制(2)由金屬厚度決定的能量傳輸機制(3)出光面幾何結構所決定的光場分布.

由以上三個面向的討論,我們提出了改良的複合式奈米孔洞結構,同時具備了光穿透增益以及微小光點的特性.此外,我們在孔洞周圍的表面制作微結構來提供水平方向的動量以激發自由電子的集體震盪,由此進一步增強光穿透之增益.我們提出的改良結構與原始的方型奈米孔洞比較,可以提高光穿透增益到達兩百倍左右.

Extraordinary Transmission Enhancement in Corrugated Composite Nano-aperture

Master student：Wei-Chung Hung **Advisor**：Dr. Chung-Hao Tien

Department of Photonics & Institute of Electro-Optical Engineering

National Chiao Tung University

ABSTRACT

As the development of optical data storage, the demand for higher data storage capacity inspires researchers to produce the tiny and intense light spot. The nano-probe is widely used to confine the input light into the small area at the near field. The main issue lies in its low throughput. The Bethe's formula provides a clear physical insight to the nano-aperture. When the aperture size on an infinitely thin and perfectly conducting film is much smaller than the incident wavelength, the transmission of the circular aperture is ~ $(d/\lambda)^4$.

The mechanism of nano-aperture is divided into three parts: (1) photon capture ability determined by the entrance geometry, (2) energy transportation influenced by the film thickness, and (3) the exit radiance distribution.

Based on the prior literature, we designed a composite aperture to keep the intense throughput and the tiny spot at the same time. The metal surface in the vicinity of the aperture is corrugated to supply horizontal momentum and induce the "extended surface plasmonic resonance", which results in the further enhancement of the transmission. Our proposed corrugated composite aperture was able to enhance power throughput by ~200X compared with the comparable square aperture.

致謝

我首先要感謝我的指導教授 田仲豪老師提供良好的研究環境，讓我在兩年的研究生涯學到了很多:除了外語與研究上的訓練之外，還傳授許多為人處事的方式，不僅幫助我順利完成此論文，也獲得許多寶貴的人生經驗。

還必須感謝璧如學姐在各方面對我的指導及協助，讓我在做研究思維上以及技巧上有顯著的進步。實驗室的學長:小陸 進哥 健翔 子翔也對我的生活與學習提供了很多的幫助，同屆的小潘 囂張 柏毅 小余 貢丸陪我共患難渡過了許多研究上的瓶頸，可愛的學弟妹 阿昇 宗瑋 牛奶提供的歡笑豐富了煩悶的研究生活。你們帶給我的點點滴滴都是兩年研究生活中的難忘記憶。

最後我的家人，感謝你們從小的栽培 照顧以及陪伴，讓我無後顧之憂的專心於研究。要謝謝一路上陪我走過的所有人，有你們的支持我才能走到這裡，把完成一個段落工作的喜悅與你們分享。

Table of Contents

Figure Captions

List of Tables

Chapter 1

Introduction

1.1 Motivation

As the coming of the tera-era, the demand of the recording capacity is exploding. The capacity is influenced greatly by the spot size which correlates closely with numerical aperture (NA) and the incident wavelength (λ) of the optical data storage system. Therefore, the optical data storage (ODS) system is developed to achieve a higher density with shorter wavelength laser diodes and higher numerical aperture objective lens, as shown in Fig. 1-1 .

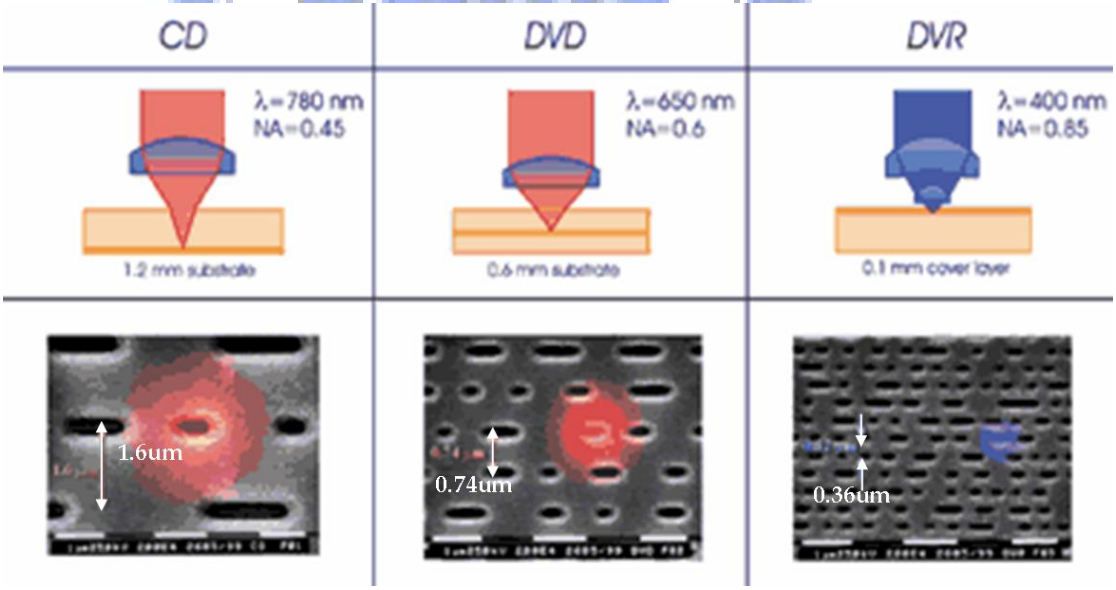


Fig. 1-1 Evolution of the CD, DVD and DVR system and electron microscope photographs of the information pits of the three systems.

Regardless of which mechanism for data reading/writing, breaking the far-field diffraction limit is necessary evidently. However, the method associated with near-field optics comes into exist.

Employing a sub-wavelength aperture in an opaque mask demonstrated by Betzig [1] is to overcome the far-field diffraction limit, as presented in Fig. 1-2. But, the extremely low transmittance of 10^{-5}~10^{-6} is a fatal defect for practical applications. How to enhance the light transmission under the sub-wavelength aperture has gathered great researching interests in recent years.

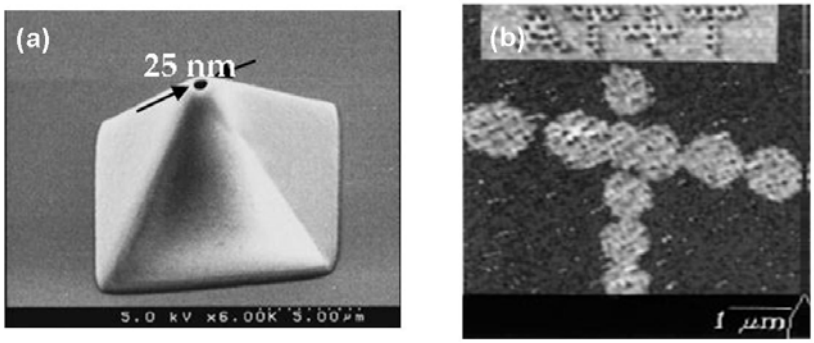


Fig. 1-2 (a) fabrication of sub-wavelength aperture, (b) Betzig's experiments result with resolution ~60nm.[1]

1.2 Objectives

The crucial role for the enhancement is the free electrons in the skin depth of the metal. The collective motion of these electrons is so-called "surface plasmon." How it works for the enhancement will be discussed clearly in this thesis by analyzing the optical properties of light transmission through the metallic sub-wavelength aperture. The objectives are to design the special aperture with suitable corrugation to enhance the light transmission under the small spot size presupposition.

1.3 Organization

The objective of this thesis is to design a novel sub-wavelength aperture that can provide adequate transmittance for practicability.

This thesis is organized to introduce the basic knowledge of the surface plasmon and the simulation tools in Chapter 2. Then, we explain the extraordinary light transmission mechanism in the C- and I-shaped and the slit aperture. From this groundwork, our composite aperture will be given. All the simulation results and the optical properties are discussed in Chapter 3. In Chapter4, the entrance corrugation, which provides the additional momentum to excite "extended surface plasmon resonance", is studied for the composite aperture. Finally, we end the thesis with the conclusions and future works in Chapter5.

Chapter 2

Theory & Simulation Model

In this chapter, we begin with the primary research to the light transmission through a sub-wavelength aperture—Beth' s formula. The fundamental theorems of the surface plasma resonance (SPR) will be introduced with its mechanism (grating coupling, attenuated total internal reflection, and composite diffraction evanescent wave) in the following. In addition, the numerical approximation, finite difference time domain (FDTD) solution, and the Drude model, the description of the material response to the incident electromagnetic wave, are briefly interpreted as well. The optical system and some terminologies we used in this thesis will be given finally.

2.1 Bethe's Formula

As widely known, the geometric optics is a macroscopic description of light as the objectives sizes in the optical system are much larger than the wavelength scales. When the scales are approached to the nano-meters, the wave-like property and the boundary condition can't be ignored any more. Bethe's theory [2] is the most fundamental one to figure out the transmitted power through a sub-wavelength aperture in a thin perfect conducting film. It indicates the transmission , normalized to the hole area, is proportional to $(d/\lambda)^4$, where λ is the wavelength and d is the hole diameter, as shown in Fig. 2-1.

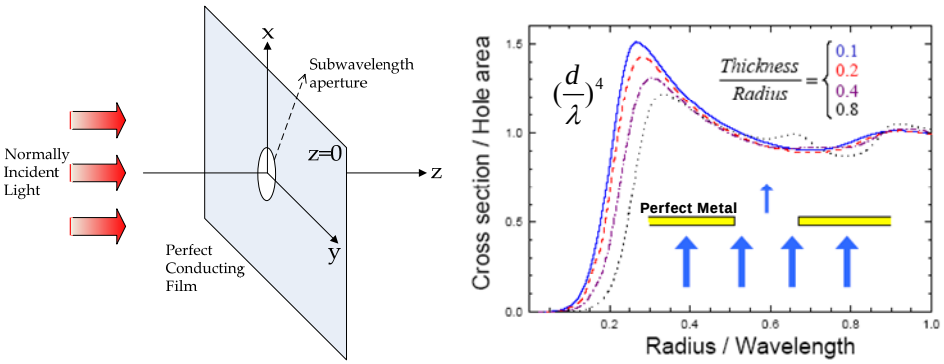


Fig. 2-1 Model of Bethe's theory and transmission cross section of a cylindrical hole drilled in a perfect-metal film as a function of the hole radius for different ratios of the slab thickness to the radius (see labels). The light is coming perpendicular to the film

The result presents that the transmission is extremely low. Moreover, the real metal property and the film thickness are ignored in Bethe's discussions. These parameters will cause the extraordinary enhancement in the reality which we will discuss in the following..

2.2 Surface Plasmon Resonance

Surface plasmon resonance (SPR) effect has become the critical factor to the light transmission enhancement through the metallic sub-wavelength aperture since T.W. Ebbesen et al. published "the extraordinary high transmission on 2D periodic nanohole array" [3]-[12]. We use the fundamental electromagnetism theory to introduce surface plasmon resonance effect in the following. [13]

If the EM wave is incidence from dielectric media (z>0) to metal (z<0) and the polarization is TM (H_z=0) in an infinite large metal layer, as shown in Fig. 2-2.

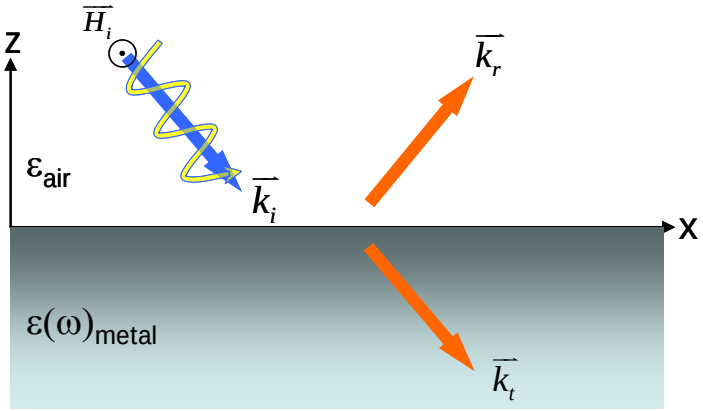


Fig. 2-2 Structures of surface plasmon excitation with TM-polarization incidence from air to metal.

The electric and magnetic fields are shown as: [14]

In air (z>0)，$\vec{H}_1 = \begin{pmatrix} 0 \\ H_{1y} \\ 0 \end{pmatrix} e^{ik_{1z}z} e^{i(k_x x - \omega t)}$ & $\vec{E}_1 = \begin{pmatrix} E_{1x} \\ 0 \\ E_{1z} \end{pmatrix} e^{ik_{1z}z} e^{i(k_x x - \omega t)}$ (2-1)

In metal (z<0)，$\vec{H}_2 = \begin{pmatrix} 0 \\ H_{2y} \\ 0 \end{pmatrix} e^{-ik_{2z}z} e^{i(k_x x - \omega t)}$ & $\vec{E}_2 = \begin{pmatrix} E_{2x} \\ 0 \\ E_{2z} \end{pmatrix} e^{-ik_{2z}z} e^{i(k_x x - \omega t)}$ (2-2)

The parallel wave vector component are the same even in different media here ($k_{1x} = k_{2x}$). Applying them (2.1) & (2.2) into the Maxwell equations :

$$\nabla \bullet \vec{E} = \frac{\rho_f}{\varepsilon} \quad (2\text{-}3)$$

$$\nabla \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t} \quad (2\text{-}4)$$

$$\nabla \bullet \vec{H} = 0 \quad (2\text{-}5)$$

$$\nabla \times \vec{H} = \varepsilon \frac{\partial \vec{E}}{\partial t} + J_f \quad (2\text{-}6)$$

and the Boundary conditions : $E_{1x} = E_{2x}$ & $H_{1x} = H_{2x}$ (2-7)

Where E is electric field, ρ_f is free electric charge density, ε is permittivity, H is magnetic field, J_f is free current density, and μ is permeability. Because the

system is source free and ε are only frequency dependent, the dispersion relation for surface plasmon wave can be described:

$$\frac{k_{1z}}{\varepsilon_{air}} + \frac{k_{2z}}{\varepsilon(\omega)_{metal}} = 0 \tag{2-8}$$

There are two necessary condition to excite the surface plasmon wave. The first one, $\varepsilon_{air} \times \varepsilon(\omega)_{metal} < 0$, is easily to be satisfied due to that the dielectric constant of metal are negative in visible light range. However, the parallel component of the wave vector k_x appears in the form as:

$$k_{sp,x} = \frac{\omega}{c}\left[\frac{\varepsilon_{air}\varepsilon(\omega)_{metal}}{\varepsilon_{air} + \varepsilon(\omega)_{metal}}\right]^{1/2} \tag{2-9}$$

Its absolute value is larger than the parallel component of incident wave vector, as shown in Fig. 2-3. The red curve, the surface plasmon excitation dispersion relation, is larger than the black light line in any k_x values. It is obviously impossible to excite surface plasmon wave for general incidence from air to metal without other mechanism. The two primary coupling methods for extra parallel momentum k_x are going to be introduced briefly as following.

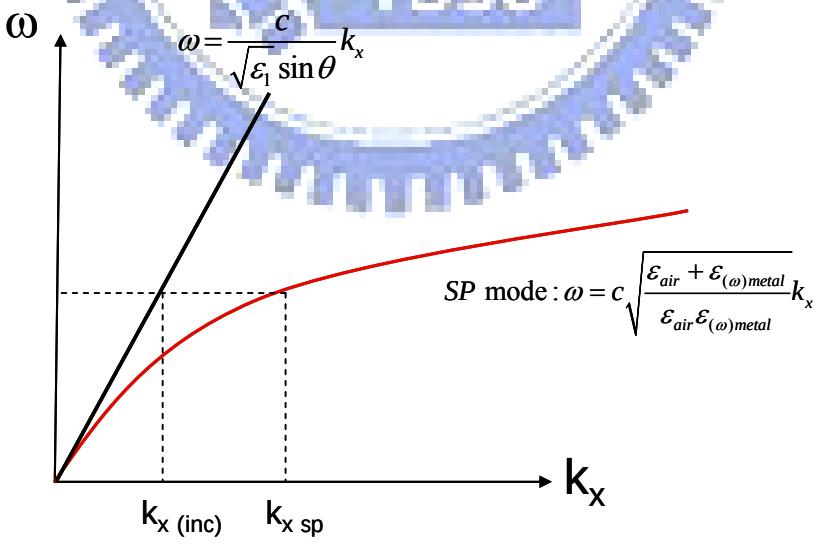


Fig. 2-3 Dispersion curve relation between light incidence in air and the surface plasmon mode.

2.2.1 Grating Coupler Method

Due to the reciprocal lattice vector theory, the incident electromagnetic field would gain extra parallel momentum k_x in the periodical grating structure, as shown in Fig. 2-4.

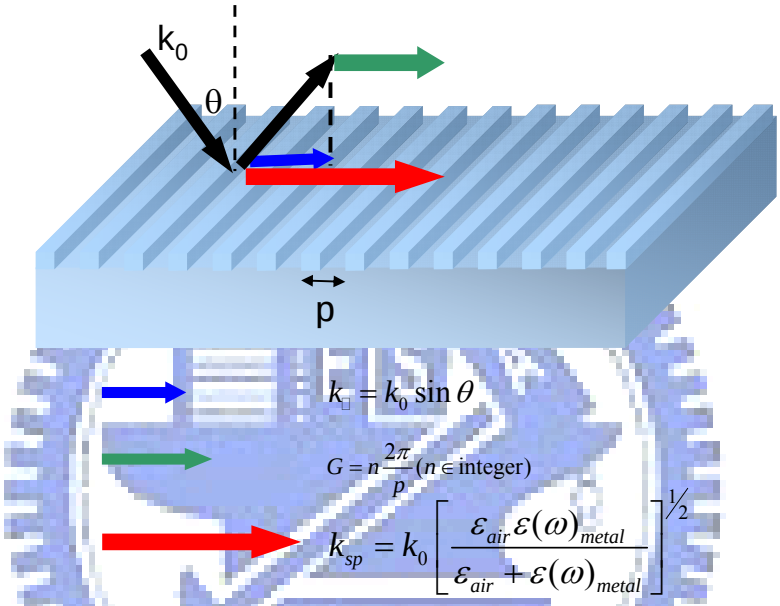


Fig. 2-4 Surface plasmon wave is excited by grating coupler.

The titled incident light could gain the extra horizontal momentum (G) which is inversely proportional to lattice constant supported by the periodical grating as the mathematical expression:

$$k_{sp} = k_0 \left[\frac{\varepsilon_{air}\varepsilon(\omega)_{metal}}{\varepsilon_{air} + \varepsilon(\omega)_{metal}} \right]^{1/2} = k_0 \sin\theta + G \qquad (2\text{-}10)$$

In $\omega - k$ plot, the gratings shift the light line (black solid) for compensation on the momentum difference to the modified curve (green dash) to match the surface plasmon mode (red). We show the dispersion curve to illustrate this phenomenon in Fig. 2-5.

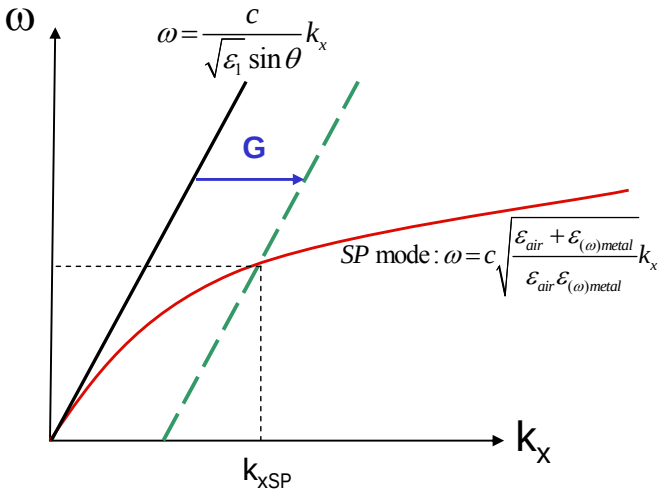


Fig. 2-5 Dispersion curve relation between light incidence in air (black-solid) and the surface plasmon mode (red), and the modified curve by grating coupler (green-dash).

2.2.2 Attenuated Total Internal Reflection Method

Putting a prism with higher dielectric constant above the metal layer is so-call Otto or Kretschmann configuration, as shown in Fig. 2-6. The light would process higher momentum as traveling in the prism and then have the total internal reflection (TIR) at the dielectric-metal interface. The evanescent wave appears to travel along the boundary and decay in the z direction with imaginary wave vector (ik_z).

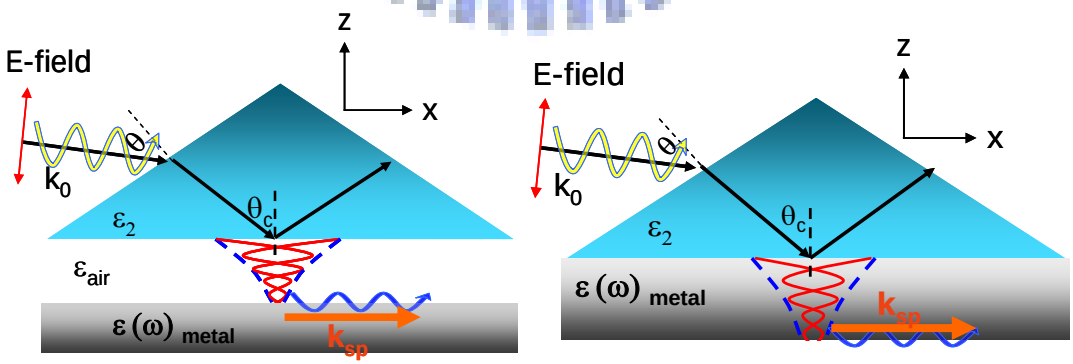


Fig. 2-6 Scheme of excitation mechanism of surface plasmon wave by attenuated total internal reflection (ATIR) method, Otto configuration (left) and Kretschmann configuration (right).

According to the momentum conservation at the interface:

$$k_{o(air)}{}^2 = k_{0(\varepsilon_2)}{}^2 = k_{x(\varepsilon_2)}{}^2 + k_{z(\varepsilon_1)}{}^2 \tag{2-11}$$

After the TIR, $\vec{k}_z = ik_z$:

$$k_{o(air)}{}^2 = k_{x(air)}{}^2 + (ik_{z(air)})^2 \tag{2-12}$$

Obviously, the imaginary z component enlarges the x component:

$$k_{x(air)} > k_{x,incidence} = k_o \sin\theta \tag{2-13}$$

the extra x momentum could excite surface plasmon wave. In the $\omega - k$ plot, the TIR change the light line slope from black solid to green dash, the intersection with surface plasmon mode means $k_{x,sp}$, as presented in Fig. 2-7.

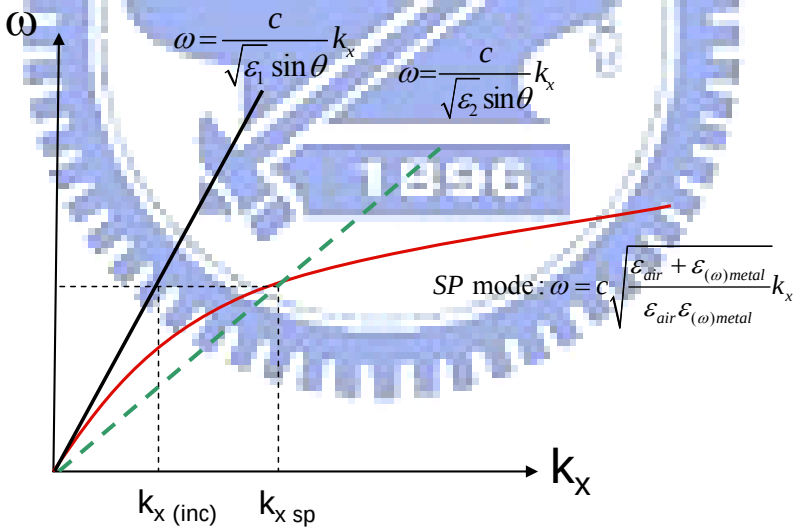


Fig. 2-7 Dispersion curve relation between light incidence in air (black-solid) and the surface plasmon mode (red), and the modified curve by ATIR coupler (green-dash).

2.3 An Alternative Theory for Transmission Enhancement: Diffracted Evanescent Wave

Besides the two coupling mechanism we mentioned above, H.J. Lezec and T. Thio [15] suggested that the light which is incident to the sub-wavelength aperture with a width of d would be diffracted into two parts, shown in Fig. 2-8:

1. For $k_x < k_o$, the light propagating into free space is so-called a radiative ("homogeneous") mode.
2. For $k_x > k_o$ and $k_z = i(k_x^2 - k_o^2)^{1/2}$, the light propagating along the surface is named as an evanescent ("inhomogeneous") mode.

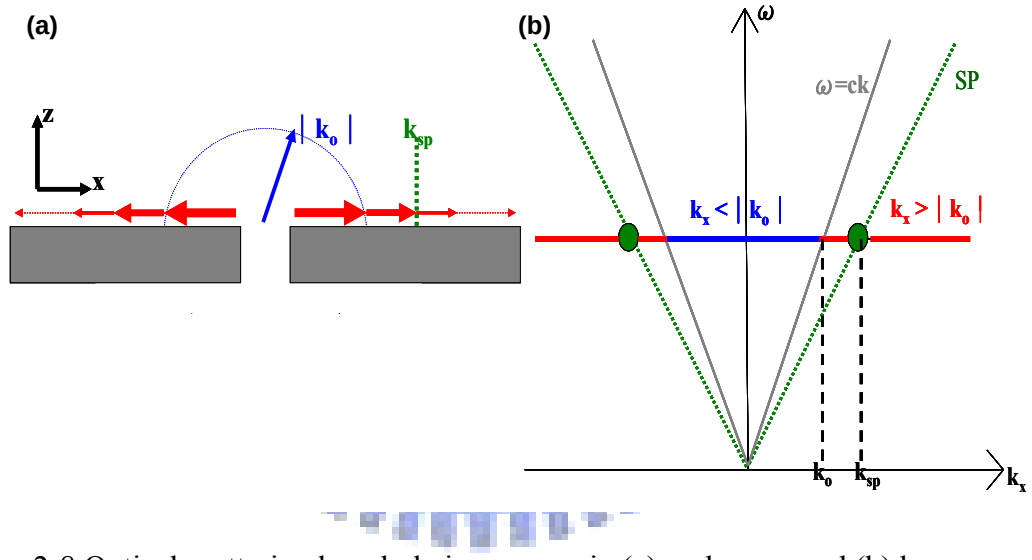


Fig. 2-8 Optical scattering by a hole in a screen in (a) real space and (b) k-space. $k_x < k_o$ are radiative modes; $k_x > k_o$ are evanescent modes

As d/λ_o is reduced smaller, the fraction of the total power emerging from the aperture which is diffracted into evanescent modes grows larger. Of this evanescent modes distribution, only one single mode (the green line in Fig. 2-8(a) and the green dot in Fig. 2-8(b)) matches the SP wave vector k_{sp}, where $k_{sp} = n_{sp}(2\pi/\lambda_o)$ and

$$n_{sp} = \frac{k_{sp}}{k_o} = \left(\frac{\varepsilon_m \varepsilon_d}{\varepsilon_m + \varepsilon_d}\right)^{1/2} \tag{2-14}$$

The SP is the TM guided mode supported by metal/dielectric interface only when $\varepsilon_m < \varepsilon_d$ which is the real part of the permittivity for metal and dielectric, respectively.

Diffraction scattering from a sub-wavelength aperture does provide the necessary momentum to couple to an SP (see in Fig. 2-8(b)). However, the fraction of the total evanescent power coupled into the SP is very low. In other words the SP does not constitute a preferred channel for diffraction compared to the other diffracted lossy modes which are not long-range guided modes of the surface.

To understand the fundamental characteristics of the total evanescent field diffracted by a sub-wavelength, we consider the light diffracted a single slit of width d as in Fig.2.3-1(a). The electric field amplitude of the total diffracted field at the surface is given by:

$$E(x, z=0) = -\frac{E_o}{\pi}\left[Si\left(k_o\left(x+\frac{d}{2}\right)\right) - Si\left(k_o\left(x-\frac{d}{2}\right)\right)\right] \tag{2-15}$$

Where x is the distance from the center of the slit, E_o is the source field amplitude, and $Si(\alpha) = \int_0^{\alpha} (\sin(t)/t)dt$. For $d > \lambda/2$, the field outside the slit can be accurately approximated by:

$$E(x) = \frac{E_o d}{\pi x}\cos\left(k_o x + \frac{\pi}{2}\right) \tag{2-16}$$

The total diffracted evanescent field can be described as a running wave propagating with a well-defined wave vector lying along the plane of the surface. This wave is designated as a “composite diffracted evanescent wave” (CDEW), where “composite” refers to the summation of all diffracted inhomogeneous modes performed. It plays a crucial role to transmission enhancement for an aperture on a smooth metal film.

2.4 Simulation Model

The finite difference time domain method, which is a rigorous solution to Maxwell's equation, is convenient for dealing with complicated geometry structure with a finite size. And the Drude model makes some assumptions for the electrons in the metal to simplify the analysis for the optical properties of metals.

2.4.1 Finite Difference Time Domain Method

Since the exact analytical solution to the practical electromagnetic problem are often formidably difficult, it is necessary to apply numerical calculations. Finite difference time domain (FDTD) method is one of the most powerful approaches due to its rigorous calculation without any physical assumptions or premises. Its main thought is to discretize Maxwell equations by using central difference operators to replace differential operators in both space and time variables. [16]

When E and H fields are represented by the discrete values on the spatial grids and are advanced in time steps of Δt, the field vectors are staggered so that every components of the E field is surrounded by four circulating H components and vice versa, as described in Fig. 2-9. The contour integrals of E and H along the cell's edge in Faraday's and Ampere's law circulate around the corresponding components at the cell face's center. The algorithm solves for both E and H fields in a systematic approach as follow: all the E computations in the modeled space are completed and stored in computer memory for a particular time-point using previously stored H data. Then all H computations in the space are completed using the E data just computed. Inside a loop, this process continues until time-stepping is concluded. The computer simulation runtime increases with the size of the model and total number of time-steps. It must be emphasized that the FDTD method directly solves the electromagnetic

model in both the space and time domains meaning that the system under excitation will exhibit a transient response before going to the steady state.

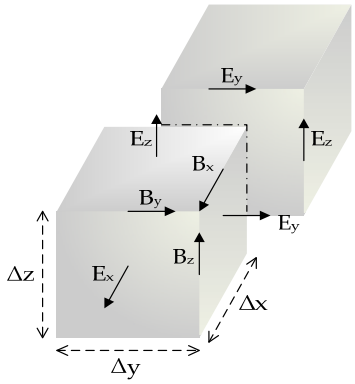


Fig. 2-9 Unit cell of FDTD mesh, components of E and H field are shifted by a half-pixel in x, y, z directions so that each E field component normal to the cell face is surrounded by the circulation of those H field components defined on the cell edges.

In 3D simulations, at least six field components must be stored and updated at each grid, which leads to considerable memory consumption and CPU requirement. Fortunately, the time update of any field component involves nearby fields one or two cells away from the grid. This translation from physical space into computer memory access allows for efficient implementation of FDTD on shared and distributed memory parallel platforms.

A general setup of the FDTD space is divided into tree regions, as shown in Fig. 2-10.

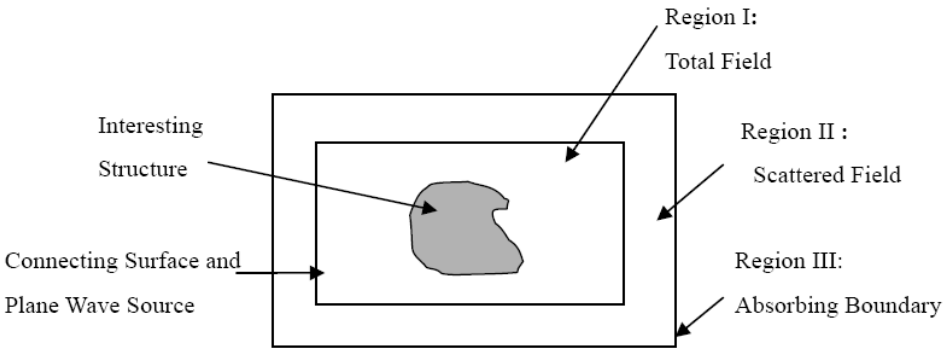


Fig. 2-10 FDTD simulation space setup

In the Region A, fields contain both incident waves and waves scattered by the interesting object. The scattered wave propagates to the Region B and is finally absorbed in the Region C. The incident wave is set around the boundary between the Region A and Region B. The region C, absorbing boundary, is purposed to model open region problems and to avoid non-physical noise which would corrupt the results in the main grid. It is used to truncate the lattice grid since the tangential components of the E and H fields along the outer boundary of the problem space cannot be updated using the normal time-stepping relations.

In most cases, low-reflection boundary conditions, such as the "Mur" and perfectly matching layer (PML) boundary condition [17], are popularly applied for those problems requiring open boundaries to simulate propagation into infinity. Therefore, combining all these three regions will allow accurate evaluation of the incident, reflected, and the transmitted waves throughout the computational domains

Additionally, several used material-dispersion models such as Debye, Drude, and Loretz can be readily incorporated with the time-dependent formulation [17] via either a linear recursive relation [18] or auxiliary differential equation methods [19]. Drude model what we adapted in this thesis will be explained briefly at the next section.

Although the numerical discretization in FDTD method does not introduce additional dissipation into the physical problem, it contributes to the numerical dispersion error in mathematics. In the commonly used implementation of FDTD, this error is proportional to the cell size h square, as a second order error function $O(h^2)$ .In this matter, a grid with about 30 points per wavelength is anyhow desired practically in order to keep the numerical dispersion errors under control.

2.4.2 Drude Model

The Drude model [20] for the dielectric function assumes that a metal can be approximated by a gas of free electron. It means that the valence electrons of atoms can move freely and independently in a solid. The complex relative permittivity ε_r^* in conductors can be written in the form as:

$$\varepsilon_r^*(\omega) = \varepsilon_r(\omega) + i\frac{\sigma(\omega)}{\varepsilon_o \omega} \tag{2-17}$$

, where σ is the conductivity of metal. In principle, both ε_r and σ are complex: the imaginary part of ε_r and the real part of σ contribute to Im[ε_r^*] which means dissipation, absorption, and dispersion; Re[ε_r^*] representing the index of refraction and the phase velocity comes from the real part of ε_r plus the imaginary part of σ.

At sufficiently high frequencies $\omega\tau >> 1$ corresponding to the visible light range and ultra-violet, where τ is the electron scattering time, $\varepsilon_r \sim 1$ (the bound electrons are too slow to respond) and $\sigma(\omega) = i\sigma_o / \omega\tau$, where σ_o is the Drude conductivity at $\omega = 0$,

$$\varepsilon_r^*(\omega) = 1 - \frac{\sigma_o}{\varepsilon_o \omega^2 \tau} = 1 - (\frac{\omega_p}{\omega})^2, \text{ for } \omega_p = (\frac{4\pi ne^2}{m})^{1/2} \tag{2-18}$$

, where ω_p is the plasma frequency, n is the electron density, m is the effective mass of the conduction electrons. In the range $0 < \omega < \omega_p$, the dielectric constant is negative and no modes are allowed in the bulk of the metal. This accounts for the fact that metals are extremely opaque at optical frequencies. However, it becomes effectively transparent for $\omega > \omega_p$ which is in contradiction with the observed reddish color for copper and yellowish color for gold. The reason is that the metal band structure can no longer be neglected and the interband absorption effect arises for high frequency.

2.5 Optical Model

The optical system we use to discuss the sub-wavelength aperture is depicted in Fig. 2-11. A light source ,emitted from –z to z, with two orthogonal linear polarized excitations (E_x & E_y) can be considered a plane wave normal into the aperture since the spot size for general incident wave is much larger than the aperture area.

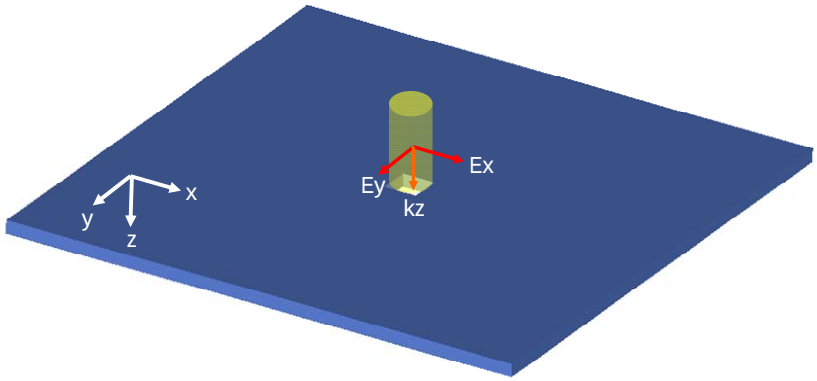


Fig. 2-11 Schematic diagram of the relation between the polarization of incidence and the subwavelength aperture.

In addition, the dispersion relation of the silver film with n (refraction index) and κ (absorption index) [21] versus incident wavelength is shown in Fig. 2-12, which implies the skin depth around 30nm at incident wavelength $\lambda = 633nm$. We should note that the film thickness must be larger than it to prevent the light from leaking accordingly.

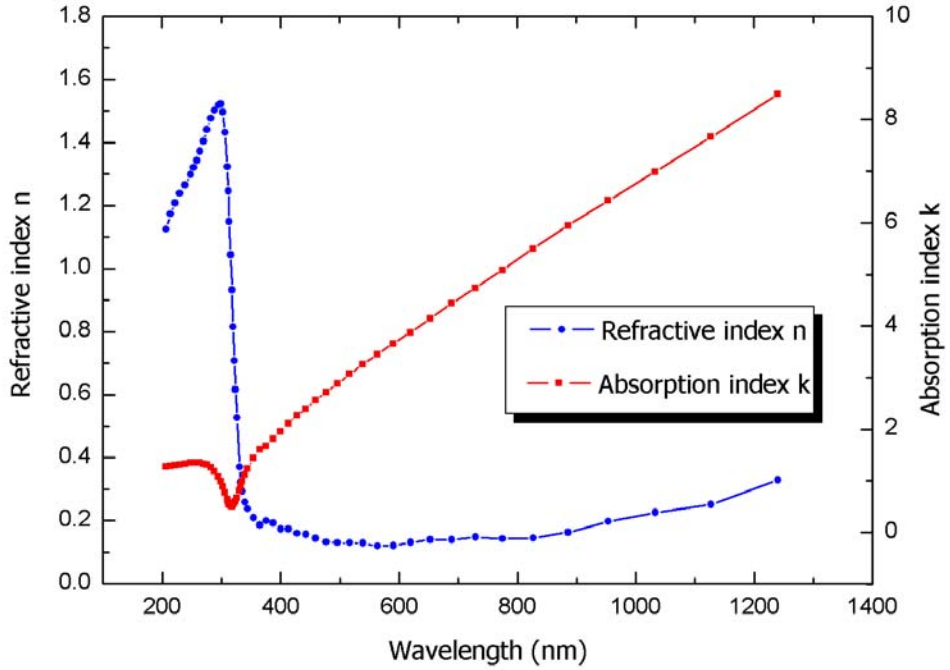


Fig. 2-12 The relation between n (refractive index), κ (absorption index) and incident wavelength .

Moreover, the curve above exhibit the refractive index n almost keeps the same around 0.14.And the absorption index κ which is linear trend in visible range (400nm-800nm) **means** the skin depth is reduced with the larger wavelength. The evaluation factors applied in the following discussion are going to be introduced.

(1) Power throughput (PT): It is dimensionless and defined as the ratio the total transmission power to the product of incident power density and the aperture area.

$$PT=\frac{\text{Total transmitted power}}{\text{Incident power density}\times\text{Aperture area}}$$

Its physical meaning is to quantify the photon capture ability of the aperture. PT less than unity indicates the photon even impinging on the aperture cannot tunnel the aperture whereas excess over unity signifies those incident photons beyond the aperture area can still be captured, as schematically plotted in Fig. 2-13.

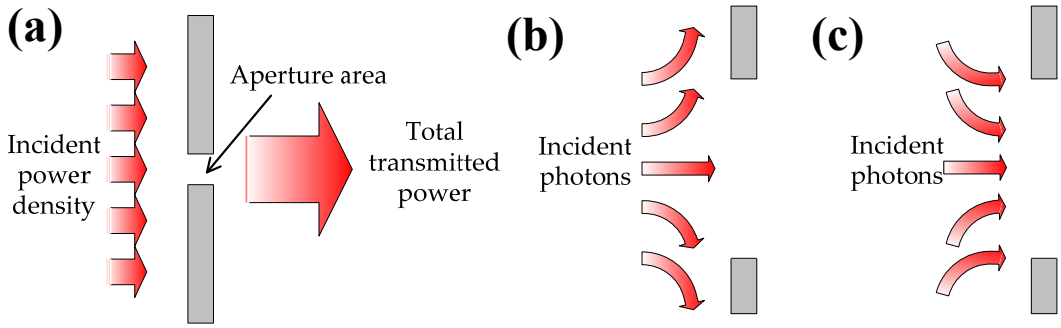


Fig. 2-13 Schemes of (a) power throughput, (b) photons incapable of tunneling and (c) photons captured by the aperture

(2) Power throughput density (PTD): It is defined as the ratio of the power through put to the spot size.

$$PTD = \frac{\text{Power throughput}}{\text{Spot size } S_x \times \text{Spot size } S_y}$$

Spot size which is the area that the tunneled photons spread out can be considered as the photon delivering ability of the aperture accordingly. From the perspective, PTD can be interpreted as

$$PTD = \frac{\text{Photon capturing ability}}{\text{Photon escaping ability}}$$

Therefore, PTD can represent the optical performance of the aperture thoroughly because it takes both incidence and emission into consideration simultaneously.

Chapter 3

Composite Aperture

The researchers had brought up the idea that localized surface plasmon (LSP) excited along the slit under perpendicular incident polarization causes the extraordinary transmission through the sub-wavelength aperture. Many improved ridge-based designs are extended from this relationship between the incident polarization and the aperture entrance geometry, such as C-shaped aperture provided by Prof. Hesselink in Stanford University and Japanese Prof. Tanaka's I-shaped aperture. In this chapter, we discuss the detailed mechanism for all these apertures from different perspectives: photon capture ability, energy transportation, and spot performance. Afterward, we design a composite aperture to combine the advantages of the slit and the ridge-based aperture.

3.1 A Sub-wavelength Aperture in Real Metal

From the previous work done by other researchers, there are some important characteristics (a) spectrum analysis of different shape sub-wavelength aperture, (b) polarization dependence, and (c) the resonance with the film thickness which support great information for our aperture reforming design.

3.1.1 Polarization Dependence on the Geometric Shape of the Aperture

The French researcher Prof. T.W. Ebbesen et al [12] have analyzed the circular

(perfectly symmetric) and rectangular (non-symmetric) aperture to verify the existence of the localized surface plasmon (LSP). At the beginning, the relationship between the transmission of the cylindrical holes with a given diameter d=270nm and different incident wavelength with various film thickness *h* are measured as Fig. 3-1. When $\lambda > 600nm$, almost twice the diameter of the hole, the transmission is sensitive to the depth of the hole. No propagating mode could be sustained at that situation.

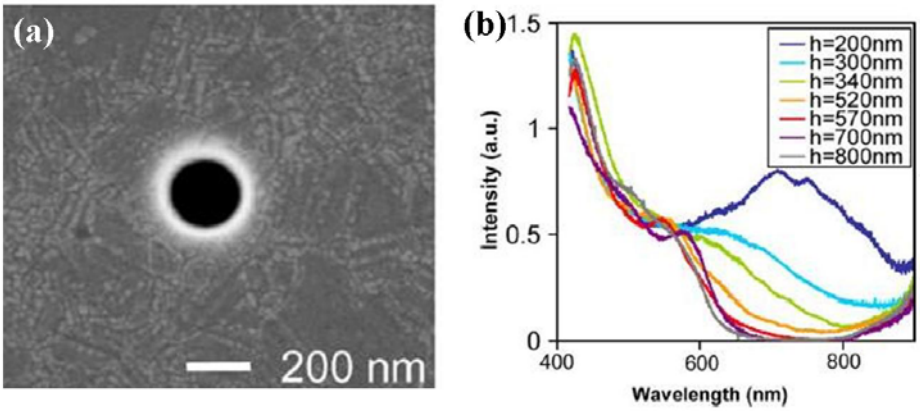


Fig. 3-1 (a) The SEM micrograph image of an isolated sub-wavelength aperture in a suspended Ag film. (b) Transmission spectra at normal incidence for cylindrical holes of diameter d=270 nm, for a range of hole depths h. Each curve is an average of the spectra of several isolated holes of the same dimensions.

Furthermore, the LSP can be clarified by examining rectangular sub-wavelength holes in Fig. 3-2. The result exhibits the transmission spectrum of a rectangular aperture in a 700-nm thick silver film for various incident linear polarizations. Two distinct peaks can be observed by simply changing the angle θ between the electric-field and the longitudinal direction. As θ varies from 0° to 90°, there are two obviously different modes at wavelength 450nm and 700nm. The intensity of the two peaks as a function of incident orientation θ obey the classical Malus law for polarization, which means that the LSP modes are launched by the components of the electric field orthogonal to the straight edge.

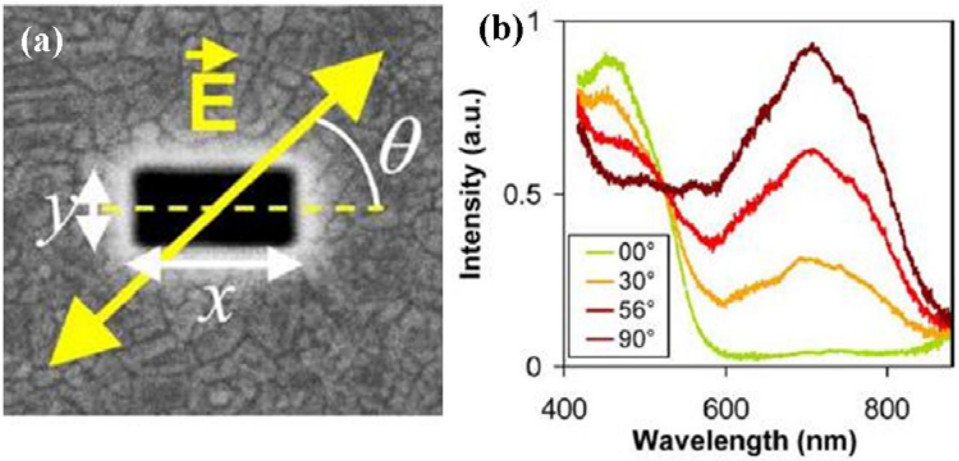


Fig. 3-2 (a) SEM micrograph image of a rectangular aperture in a suspended Ag film. Also shown are the notations adopted in the text: θ is the angle between the electric field and the longitudinal direction, x(y) is the longitudinal (transverse) dimension of the rectangle. (b) Transmission spectra of an isolated rectangular aperture for various linear polarizations (h=700 nm, x=310 nm, y=210 nm).

The total different transmission spectrum for a cylindrical hole and a rectangular aperture indicate the peak resonance wavelength is high dependence on the geometric structure of the aperture and the polarization of incidence.

3.1.2 Phenomena of Light Transmission through the 2-D Slit

At the time, the pioneer experiments results are published by the Ebbesen' s group, Prof. M. Mansuripur in University of Arizona does fruitful contributions on the theoretical analysis [22]-[24] . We are going to briefly summarize his works [25] following.

The incident light in this system is declared first such as a Gaussian beam, having full-width at half-maximum intensity FWHM = 1.5 μm at the waist (located at z = 1200 nm). And it propagates along the negative z-direction with vacuum wavelength of the light, λ_o = 1.0 μm which is in the near-infrared range. The film is

silver with $\varepsilon = 0.23 + 6.99i$ at $\lambda_o = 1.0$ μm of thickness *t* and has a slit aperture of width *W* at center z=0. Since illumination is uniform in the x-direction and the setup is independent of x, the electromagnetic fields are decoupled into $E_{\parallel}$ and $E_{\perp}$ modes in the 2D analysis, as indicated in Fig. 3-3: The $E_{\parallel}$ mode consisting of E_x , H_y and H_z field components means its electric field is along the length of the slit; the $E_{\perp}$ mode consists of H_x , E_y and E_z, with E_y being perpendicular to the slit length.

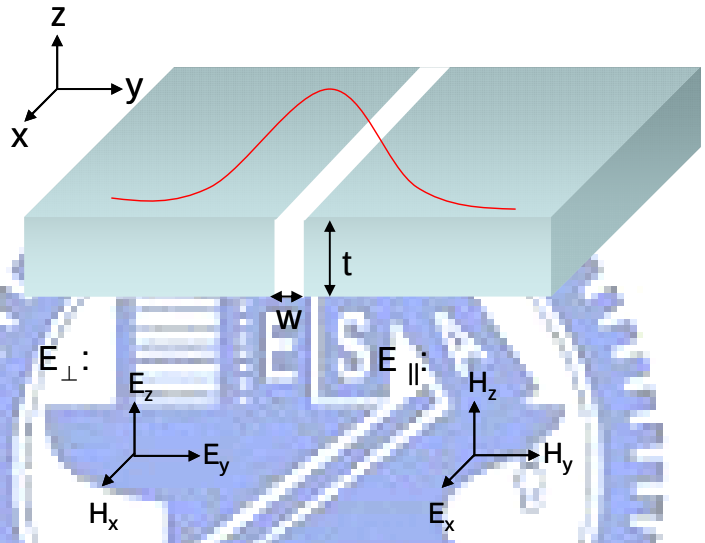


Fig. 3-3 A slit aperture is *W* wide in a metallic film of thickness t. The material of the film (silver) has refractive index $n + i\kappa = 0.23 + 6.99i$ and $\varepsilon = (n + i\kappa)^2 = -48.8 + 3.16i$ at $\lambda_o = 1.0$ μm. The incident beam is uniform along x, and has a broad Gaussian profile along the y-axis; its linear polarization state, denoted by $E_{\parallel}$ or $E_{\perp}$, indicates the incident E-field direction relative to the slit's long axis. The relevant E- and H-field components for the two polarization states are shown below the slit.

The slit is called "sub-wavelength" aperture because of its width smaller than a half of wavelength ($W < 1/2\,\lambda_o$). According to the waveguide theory, the fundamental mode could not be sustained. Anomalously, the stronger localized surface plasmon (LSP) induced by the specific incident polarization will be a crucial role to make the light pass through the sub-wavelength slit in near-field range. The discussions aim at two orthogonal incident polarizations was given as following.

(1) $E_{\parallel}$ excitation:

$E_{\parallel}$ illumination resulting in that no surface charges accumulated at the edges of the slit causes the E-field to decay rapidly in the slit. On the other hand, the incident polarization also drives the surface currents to run along the slit length. Then the surface currents support the H-field everywhere in the vicinity of the slit edge and bend the H-field lines into the aperture, then hovering around in the incident space. They are all obviously observed in Fig. 3-4. The simulation case is under the cut off condition W=0.4 λ_o, the incident optical energy, aside from a small fraction that was absorbed within the metal's skin depth, was reflected back toward the source. Only very little light could get through the slit, which matches the waveguide theory as we may expect.

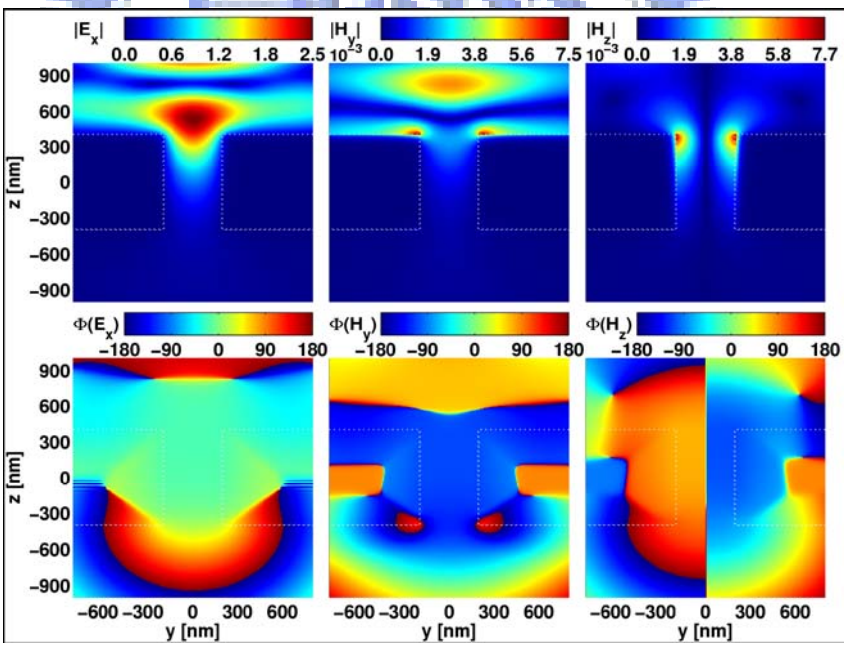


Fig. 3-4 Computed plots of Ex , Hy and Hz for the case of $E_{\parallel}$ illumination: (top) magnitude, (bottom) phase. Film thickness t = 800nm; slit-width W = 400nm < 1/2 λ。

(2) $E_{\perp}$ excitation:

In contrast, $E_{\perp}$ mode has totally different field distributions from previous

description. The E-field plot shows the surface charges accumulated at the edge of the slit because of the discontinuity of the incident E_y on the perpendicular aperture structure. Additionally, the accumulated charges pushing by the incident photons flow along the walls of the gap. The surface current deposits at the exit plane finally. The motion of the electrons produces the field distributions as shown in Fig. 3-5. The periodic time-oscillations of the charges make the E_y and H_z transportation in the form of standing wave.

Since the sub-wavelength slit (W=400nm<1/2 λ_o) is applied, the transmission could not be precisely explained by the waveguide theory. The light transmission for $E_\perp$ mode is larger than that for $E_\parallel$ excitation. It exhibits that the absence of a cutoff for $E_\perp$ illumination, even for an aperture as narrow as W =100 nm (<< 1/2 λ_o).

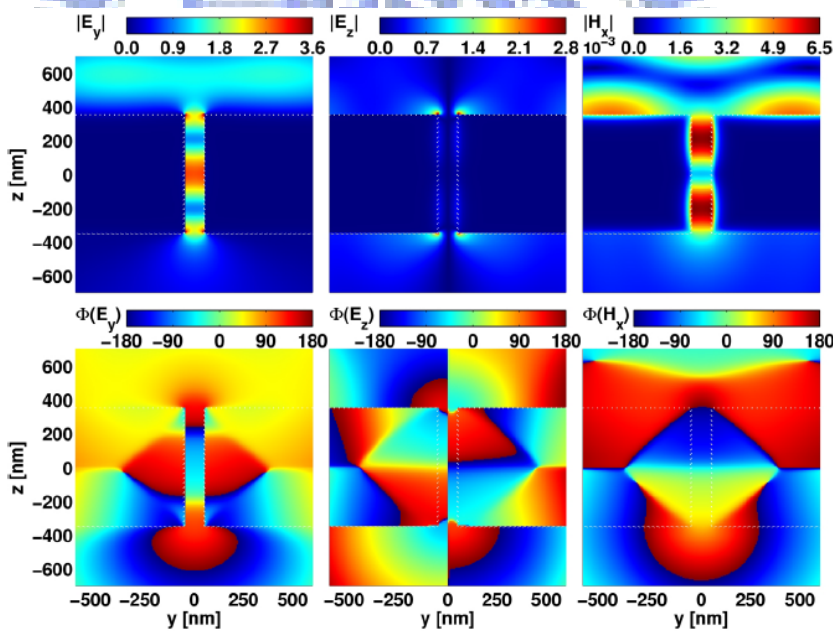


Fig. 3-5 Computed plots of Ey, Ez, Hx for $E_\perp$ illumination: (top) magnitude, (bottom) phase.t = 700 nm; W = λo/10 = 100nm

In Fig. 3-6, the influence of the film thickness on the transmission for $E_\parallel$ and $E_\perp$ modes are discussed. Generally speaking, S_z decays rapidly with the growing

thickness no matter what kind of excitation it is when the slit width is under cutoff length (W<1/2 λ_0). However, the $E_\perp$ illumination with specific thickness (t=300nm) will reach the resonance condition that the sharp peak of S_z appears suddenly, in Fig. 3-6(a) & (b). Beyond the cutoff length (W=600nm>1/2 λ_0), S_z keeps stationary with various thickness. The film thickness dependence of S_z is rather insignificant, in Fig. 3-6(c).

Fig. 3-6 (d) summarizes the film thickness influence. When W>1/2 λ_0, the film thickness almost does nothing to S_z. On the contrary, W<1/2 λ_0, S_z drops fast with the increasing film thickness under $E_\parallel$ illumination because of the absorption of the wall of the slit. Only for incident $E_\perp$ polarization, S_z oscillates periodically with the film thickness.

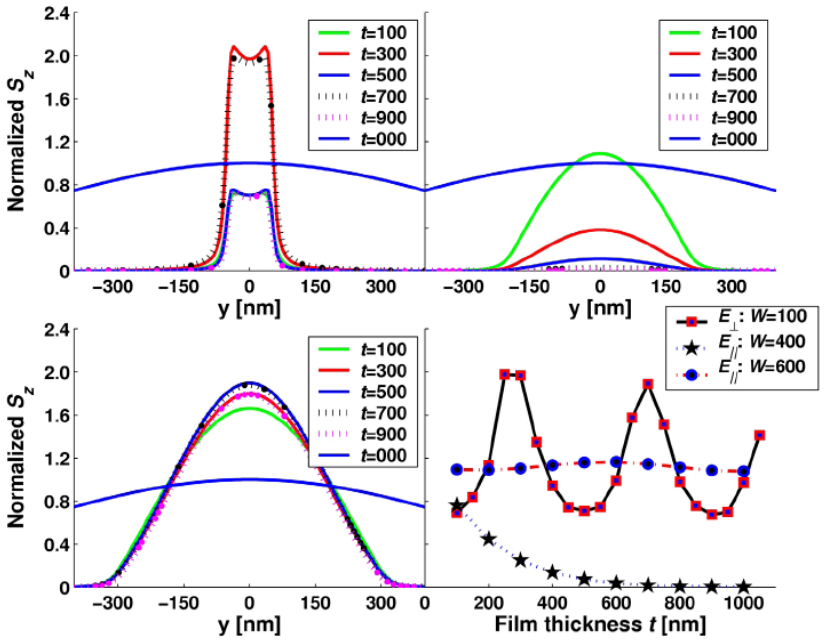


Fig. 3-6 Computed plots of the energy flux density S_z at the output aperture (z = −½ t) for different film thicknesses ranging from t = 100 nm to 900 nm. The broad blue line in (a)−(c) represents the incident beam's S_z at z = 0 (in free-space). (a) $E_\perp$ illumination, W = 100nm; (b) $E_\parallel$ illumination, W = 400nm; (c) $E_\parallel$ illumination, W = 600nm; (d) total transmitted S_z versus t for the slits depicted in (a)−(c).

3.2 Aperture Morphing

3.2.1 Slit Aperture

Although previous work demonstrated that the incident light having the perpendicular polarization to the slit long axis could produce powerful transmission, the theoretical analysis is restricted in two dimensions. We discuss the three dimensional real case by starting with the square aperture. Then the boundary effect for the charge-accumulated phenomena will be analyzed by changing the aspect ratio of length to width.

In Fig. 3-7, the square aperture (60×60 nm^2) transmission spectrum decreases with the incident wavelength. The quartic fit reveals the tendency matching Bethe's formula even that the film thickness contradicts the Bethe assumption.

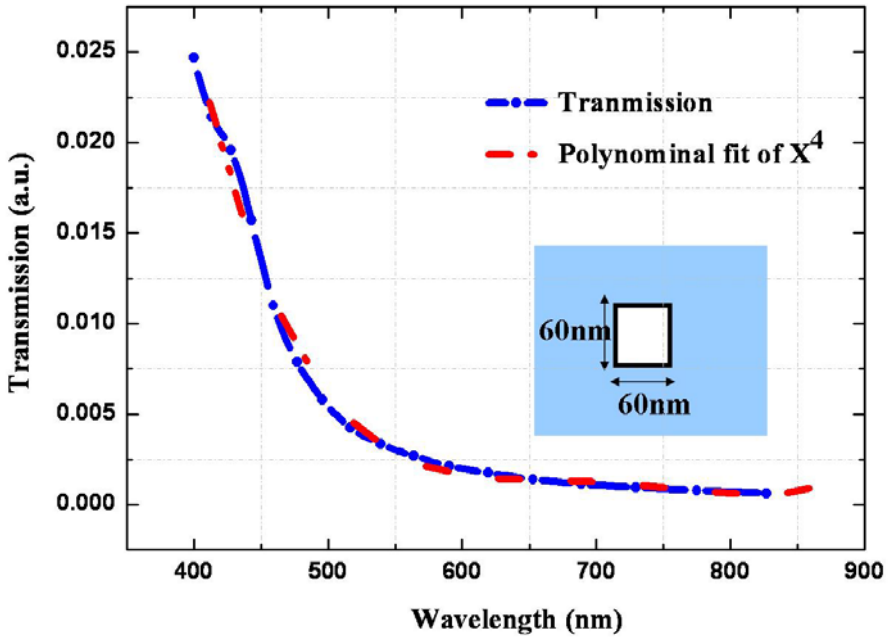


Fig. 3-7 The transmission spectrum of the square aperture (60×60 nm^2) with 200nm film thickness under normal incidence.

Besides the film thickness influence, the shape of the aperture also affects the

transmission spectrum. As previous mention, the $E_{\perp}$ excitation produces stronger transmission than $E_{\parallel}$ illumination. However, the enhancement will not keep growing with increasing length of the slit. The E_x field distributions for various rectangular apertures with different aspect ratio are shown in Fig. 3-8.

Because surface plasmon will change the refractive index of the metal $\sigma_y(\omega) = \sqrt{\varepsilon_2/(1+\varepsilon_2)}$ and the permittivity from the Drude model is a function of ω_p : $\varepsilon_2(\omega) = 1 - \omega_p^2/(\omega^2 + i\gamma\omega)$ .Therefore, the characteristic function of surface plasmon can be expressed in the form: $\sigma_y(\omega) = \sqrt{[(\omega^2 - \omega_p^2) + i\gamma\omega]/[(2\omega^2 - \omega_p^2) + 2i\gamma\omega]}$. Surface plasmon will change the response of the metal to the external field. As a result, the surface plasmon phenomenon isn't illustrated in the oscillation of the valence electron and is presented in the electric field distribution indirectly.

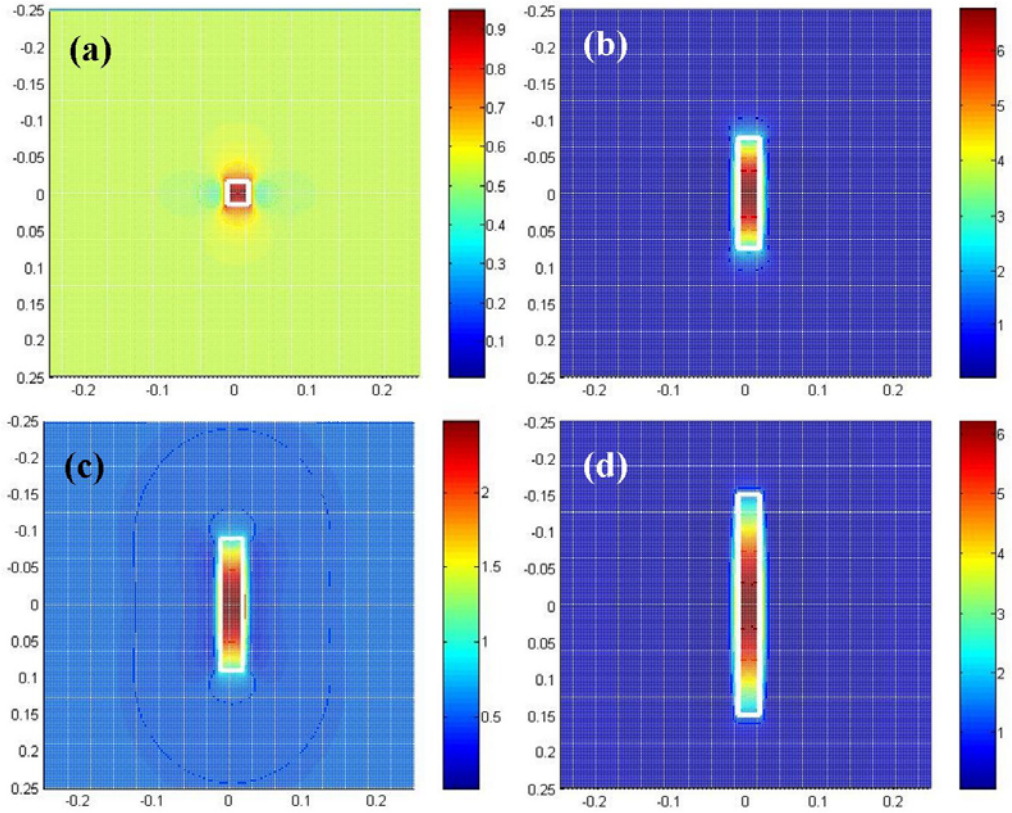


Fig. 3-8 The E_x distributions at entrance plane for different apertures in 200nm thick Ag film with fixed width 30nm and (a)L=30nm(b)L=150nm(c)L=180nm(d)L=300nm, respectively

In the square aperture (30×30 nm^2) with 90° rotational symmetry, the responses to orthogonal polarizations are the same. The few accumulated charges along the short edge traveling along the wall of the aperture results in weak transmission in Fig. 3-8(a). The accumulated charges form localized surface plasmon (LSP), the dominant effect to high transmission, as the aperture reformed from square to slit. However, the energy will leak out without the specific length as (c).

Besides powerful transmission, the spot size should be taken into accounts. The slit with longer length, as shown in Fig. 3-8(d), which sacrifices the smaller spot size just induce the same field intensity as slit (b) with exact length. Therefore, the longest slit does not match our needs. The relationship between PTD and aspect ratio in Fig. 3-9 supports the reference for our following aperture design.

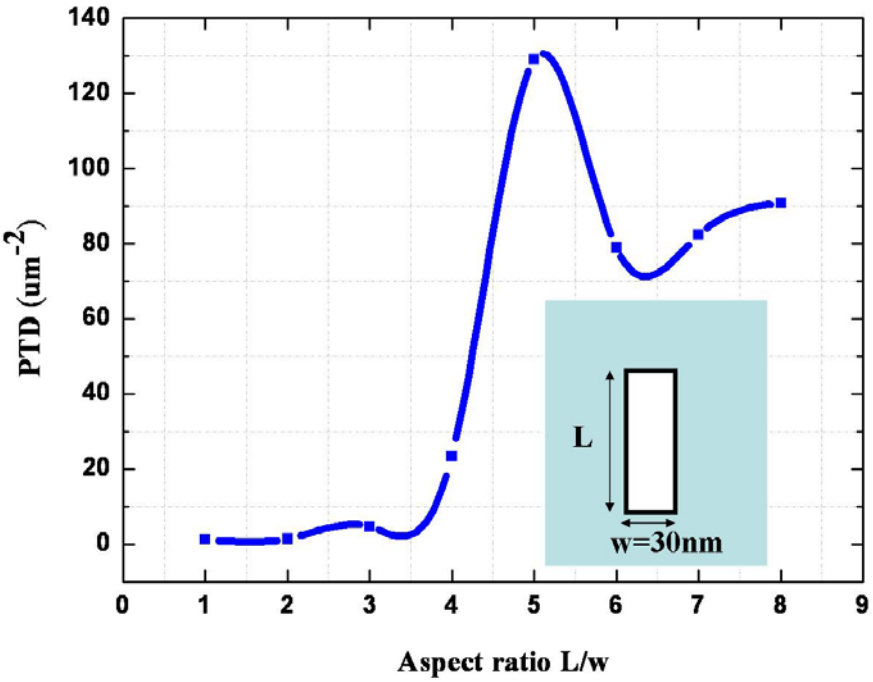


Fig. 3-9 The relationship between PTD and aspect ratio for a 30nm wide slit aperture

3.2.2 Ridge-based Aperture

Various geometric apertures are applied to maintain the strong transmission and small spot size at the same time. I and C shaped apertures are reformed from the slit, as presented in Fig. 3-10. The corresponding cutoff wavelengths of the fundamental propagation mode are 420nm, 577nm, and 600nm for slit, C-, and I- shaped aperture, respectively [26]. No propagation mode could be sustained in our discussion having 633nm incident wavelength.

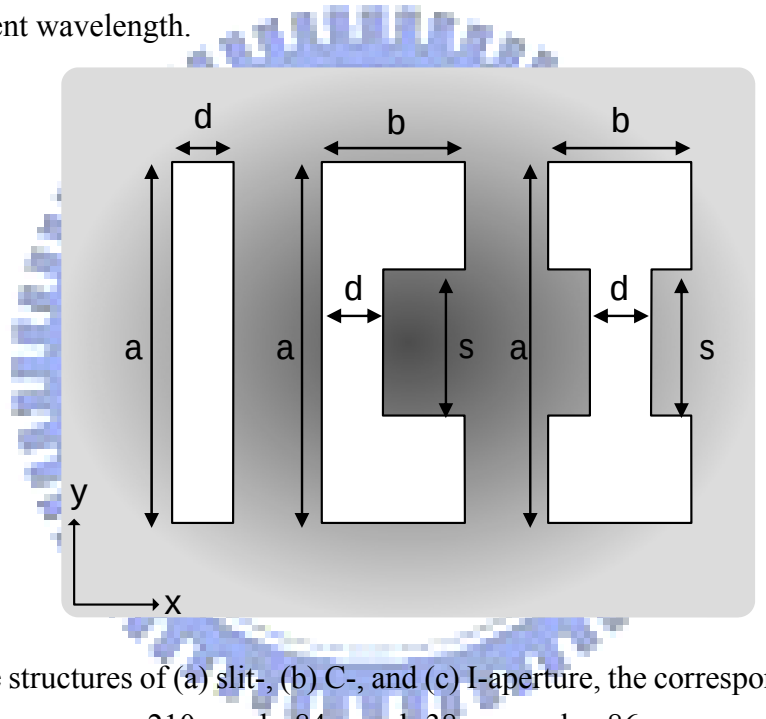


Fig. 3-10 The structures of (a) slit-, (b) C-, and (c) I-aperture, the corresponding lengths are a=210 nm, b=84 nm, d=38 nm, and s=86 nm.

When the E_x polarization illuminates the slit aperture in 200nm thick silver film, the entrance E_x-field distribution is depicted as Fig. 3-11(a). The accumulated charges along the long axis because of the discontinuity of the incident polarization result in the narrow E_x filed profile at horizontal cross section at y=0 and the other broad one at x=0, as shown in Fig. 3-11(b).

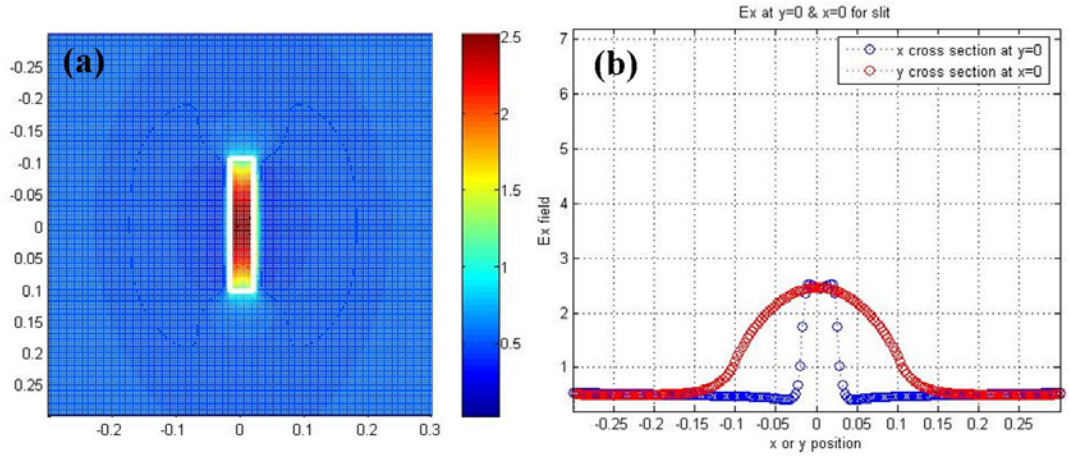


Fig. 3-11 (a)E_x distribution and (b) E_x cross section profile at the entrance of the slit aperture (38nm×210nm) in 200nm thick Ag film.

Other improved design, such as C- and I- shaped aperture, maintain the ridge part and add two extra arms to confine the charges distribution, such as electrical dipole concentrated around the central part of the aperture .The cross section profiles in Fig. 3-12 show the centralized diploes perform stronger filed intensities almost two times than uniform diploes in the slit.

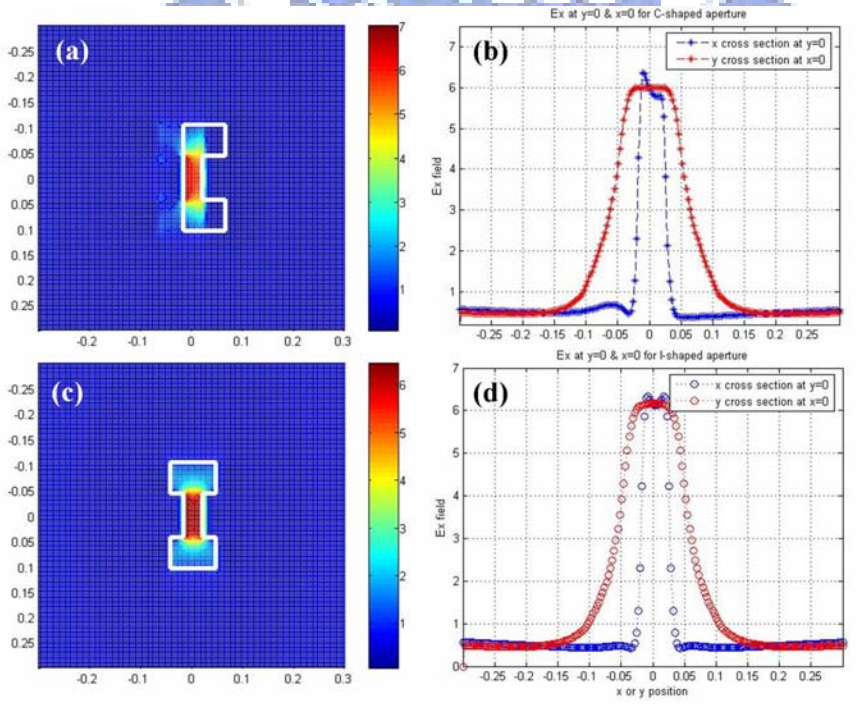


Fig. 3-12 The E_x-filed distribution (a)(c)and E_x cross section profile (b)(d) at the entrance of the C- and I-shaped aperture in 200nm thick Ag film.

Even though the ridge-based aperture excites more intense LSP field intensity at specific cross section, the geometry restricts the LSP distributions over all the open aperture area. The cross section profile in Fig. 3-13 expresses that the two extra arms of the ridge-based aperture reduce the full width of half maximum (FWHM) narrower than the slit in the y direction. It explains that the slit excites the LSP most efficiently along its edges. The slit owns better photon capture ability than the ridge-based aperture.

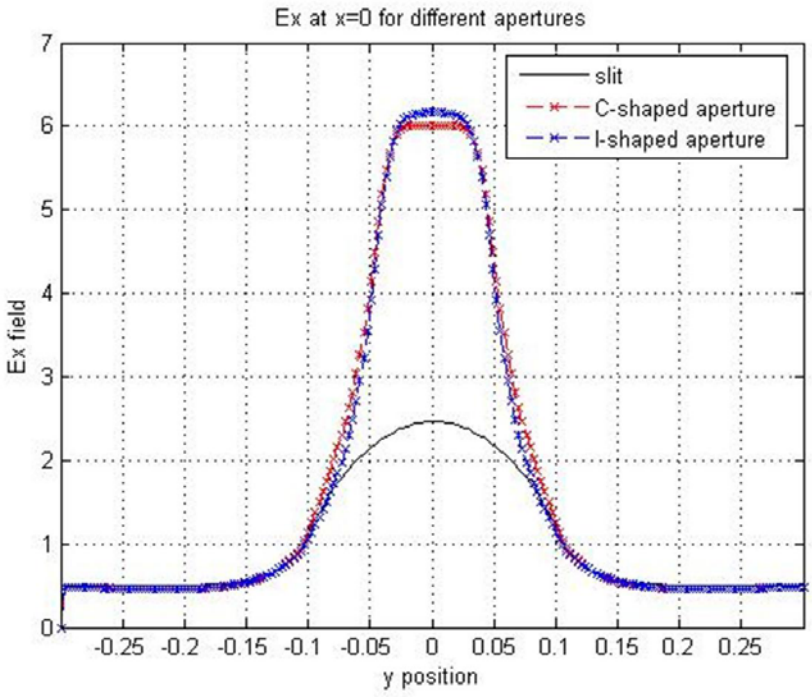


Fig. 3-13 The E_x-field cross section profile at the x=0 for the slit, C- and I- shaped aperture

After the discussion on the LSP excitation at the entrance plane, the energy transportation issue will be presented next. However, the Bethe's rule is limited in the

infinite thin film. After passing through the thick film, the beam profile at the output is not the same as the entrance filed distribution. Furthermore, the fundamental propagating mode disappears at the sub-wavelength scale on the waveguide theorem.

Fortunately, Prof. M. Mansuripur [27] supports the "radiation pressure" concept to solve this energy transportation problem. When the E_x polarized plane-wave illuminates the narrow gap of the width $\delta << \lambda_o$ between two dielectric hosts with refractive index n_o as Fig. 3-14. The field amplitudes inside the medium are $(E_x,H_y) = (E_o,H_o) = (E_o,n_oE_o/Z_o)$.And the electromagnetic field inside the gap is the superposition of two evanescent plane-waves, each of which must satisfy the constrain $k \bullet k=k_0^2$, $k \bullet E=0$, and $(k/k_o)E=Z_oH$ imposed by Maxwell's equations.

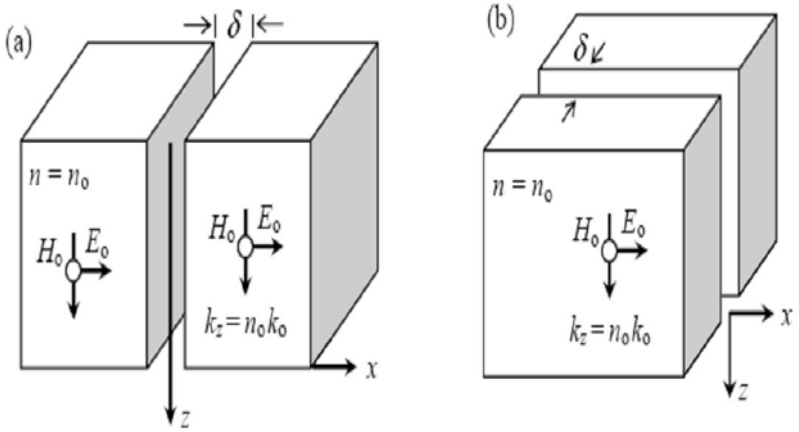


Fig. 3-14 A linearly polarized plane-wave propagates along the z-axis in a dielectric host of refractive index no. A narrow gap of width $\delta << \lambda_o$ is assumed to exist in this medium; the plane of the gap is yz in (a) and xz in (b).

The fields of these evanescent waves must satisfy the boundary conditions of the both walls of the gap. In case (a), the continuity is required of the perpendicular D-field, $D_x = \varepsilon_o n_o^2 E_o$, and the tangential H-field, $H_y = H_o$.The k-vector and field amplitudes are in the following forms.:

$$k_{\pm}/k_o = \pm i\sqrt{n_o^2 - 1}\hat{x} + n_o\hat{z} \tag{3-1}$$

$$E_{\pm} = 1/2(n_o^2\hat{x} \mp in_o\sqrt{n_o^2 - 1}\hat{z})E_o \tag{3-2}$$

$$H_{\pm} = 1/2(n_o E_o / Z_o)\hat{y} = 1/2H_o\hat{y} \tag{3-3}$$

The momentum density p in the gap is derived from the Poynting vector component S_z along the propagation direction, namely.

$$p = < S_z > /c^2 = 1/2\,\mathrm{Re}(2E_x \times 2H_y^*)/c^2 = 1/2n_o^2 E_o H_o / c^2 \tag{3-4}$$

Otherwise, in case (b), it is E_x and $B_y = \mu_o H_y$ that must be continuous. The similar analysis yields

$$k_{\pm}/k_o = \pm i\sqrt{n_o^2 - 1}\hat{x} + n_o\hat{z} \tag{3-5}$$

$$E_{\pm} = 1/2E_o\hat{x} \tag{3-6}$$

$$H_{\pm} = 1/2(n_o\hat{y} \mp i\sqrt{n_o^2 - 1}\hat{z})E_o / Z_o = 1/2(\hat{y} \mp i\sqrt{1 - n_o^2}\hat{z})H_o \tag{3-7}$$

$$p = < S_z > /c^2 = 1/2\,\mathrm{Re}(2E_x \times 2H_y^*)/c^2 = 1/2E_o H_o / c^2 \tag{3-8}$$

The general expressions for dielectric material are not the same as our metal structure. The metal will screen the fields' penetration and no fields will appear inside the metal. However, we adapt the Drude model to take the surface charges within skin depth into accounts. The accumulated charges resulting from the incident polarization which is perpendicular to the slit can still produce stronger radiation pressure at the entrance of the thick metal film, as shown in Fig. 3-15. The phenomenon matches the theoretical analysis on the dielectric gap.

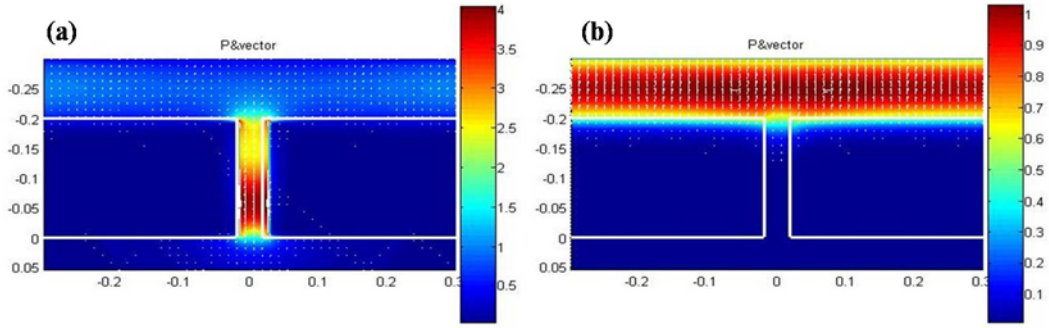


Fig. 3-15 The energy transportation inside the slit for (a) TM and (b) TE incidence

When the accumulated charges radiate the energy through the aperture, the aperture depth influences the energy transportation. The power throughput oscillated with the aperture depth as Fig. 3-16 (a). For the slit with x-polarized incidence, the oscillating period is almost half the incident wavelength. And the C- and I- shaped aperture perform shorter oscillating period physically. The enhancement inside the sub-wavelength aperture results from the Fabry Perot-like resonance effect, as depicted in Fig. 3-16 (c).

Furthermore, PT decays sharper in the slit than C- and I-aperture under the cut-off condition with y-polarized incidence, such as in Fig. 3-16 (b). Because the two extra arms of the C- and I- shaped apertures excites LSP under y-polarized illumination, which wouldn't occur for the slit aperture. However, the y-polarization which induced weak LSP at entrance and faint driving radiation pressure inside the aperture causes PT for these structures all less unity.

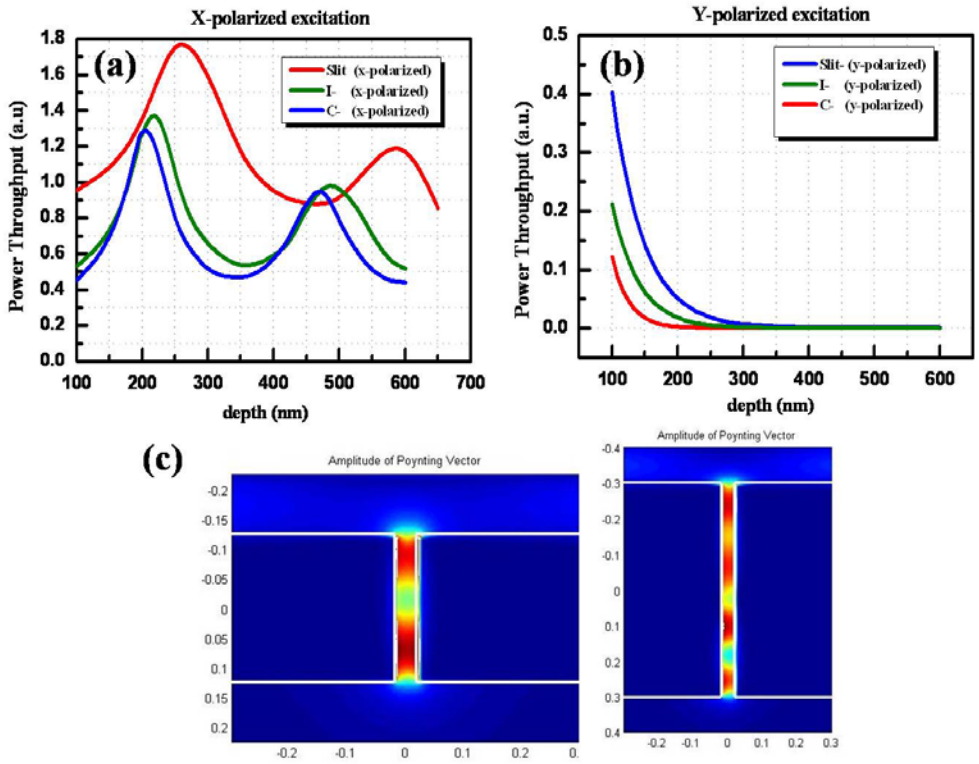


Fig. 3-16 Power throughput comparison of slit-, C- and I- aperture versus the depth of aperture under (a) x- and (b)y-polarized illumination.(c) illustrates the Fabry Perot-like resonance for X-polarized excitation corresponding to 250nm and 600nm film thickness.

We discuss the spot size by measuring the FWHM of Poynting vector at 50nm away from the aperture. For the C-shaped aperture, it is not right-left symmetrical at the line x=0 as the slit and I-shaped aperture. The two extra arms shift the energy distribution of the C-shaped aperture away from the center at horizontal cross section, in Fig. 3-17 (a). It is also the anti-symmetry avoids LSP to locate at the ends of the long axis and keeps the concentrated output performance in the vertical cross section in Fig. 3-17 (b).

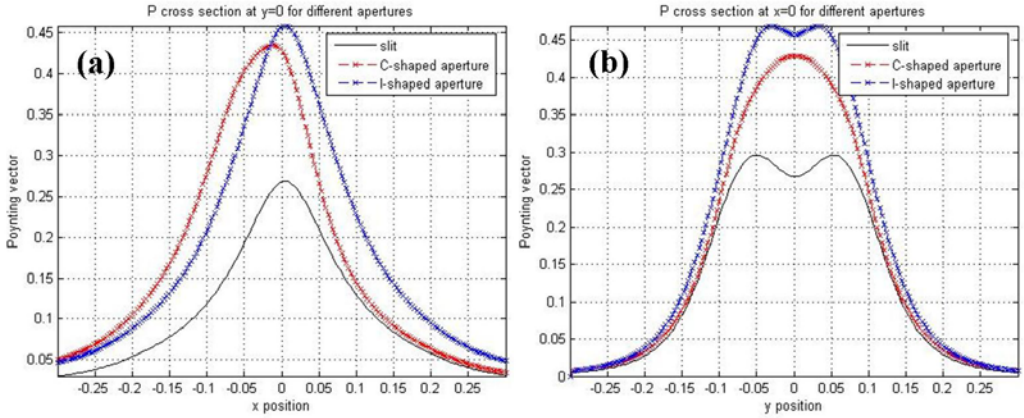


Fig. 3-17 The spot size x-(at y=0) and y-(at x=0) cross section Poynting vector profile for the slit (black solid), C-(red dash), and I-shaped (blue dash) aperture at 50nm away from the exit plane.

Generally speaking, the simple geometry of the slit captures the photon most efficiently. And the wide-distributed LSP for the slit raises the enhancement by sacrificing the energy confinement (spot size). The improved ridge-based aperture re-distributes the charges accumulation. It confines the stronger LSP excitation into the smaller area. The reformed apertures compress the light into smaller spot size by sacrificing the photon capture ability by the two extra arms at the entrance surface. Although the C- and I-shaped apertures both own the better performance in PTD than the slit in Tab. 3-1, the asymmetric (along x axis) C-shaped aperture produces more concentrated spot size.

Aperture	Area (λ^2)	PT (a.u.)	Spot size (λ^2)	PTD (μm^{-2})
Slit	**0.020**	**1.30**	**0.049**	**66.1**
C	**0.032**	**1.27**	**0.038**	**89.5**
I	**0.032**	**1.27**	**0.038**	**89.5**

Tab. 3-1 Output performance comparison between the slit, C-and I-shaped aperture. The “area” denotes the size of opening at entrance plane.

3.3 Composite aperture

According to the previous discussion, the slit produces power enhancement because of large LSP distribution coming from the preferable geometry of aperture. However, the C-shaped aperture could condense the energy distribution more concentrated. Taking both advantages of the slit and C-shaped aperture, we proposed a new composite structure shown in Fig. 3-18. At the incidence plane on a metallic thick film, it forms as the slit, and as C shape at the exit side. Alternatively, the structure can be viewed as a slit but with two arm-grooves at the exit side.

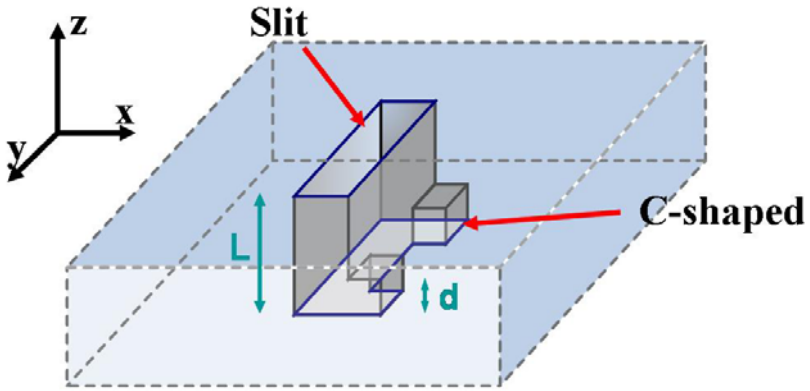


Fig. 3-18 Schematic diagram of the composite structure with parameters L and d.

We should note arm-grooves not only influence the Fabry Perot-like resonance but also affect the LSP distribution, as presented in Fig. 3-19. Even though the composite aperture and the slit open the same geometrical structure at the entrance,

the asymmetric composite aperture concentrates twice the Ex field intensity at the entrance twice than the slit.

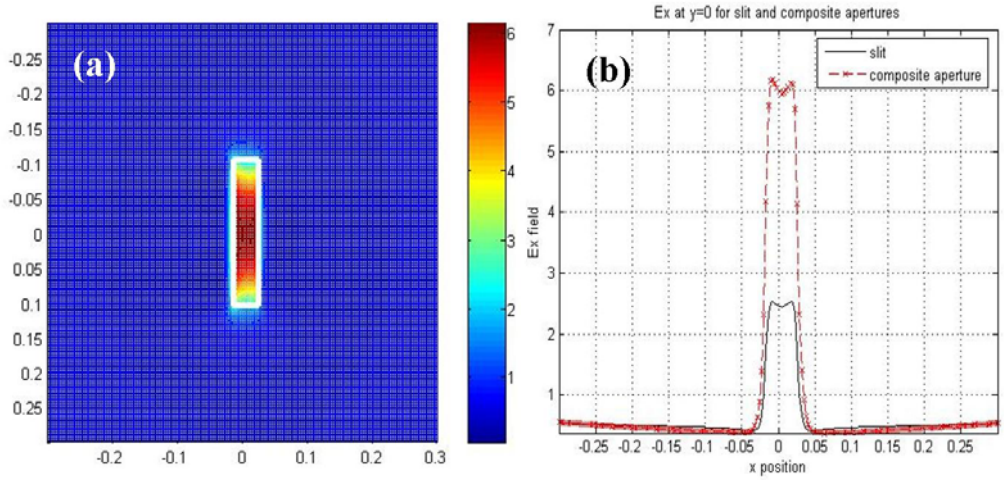
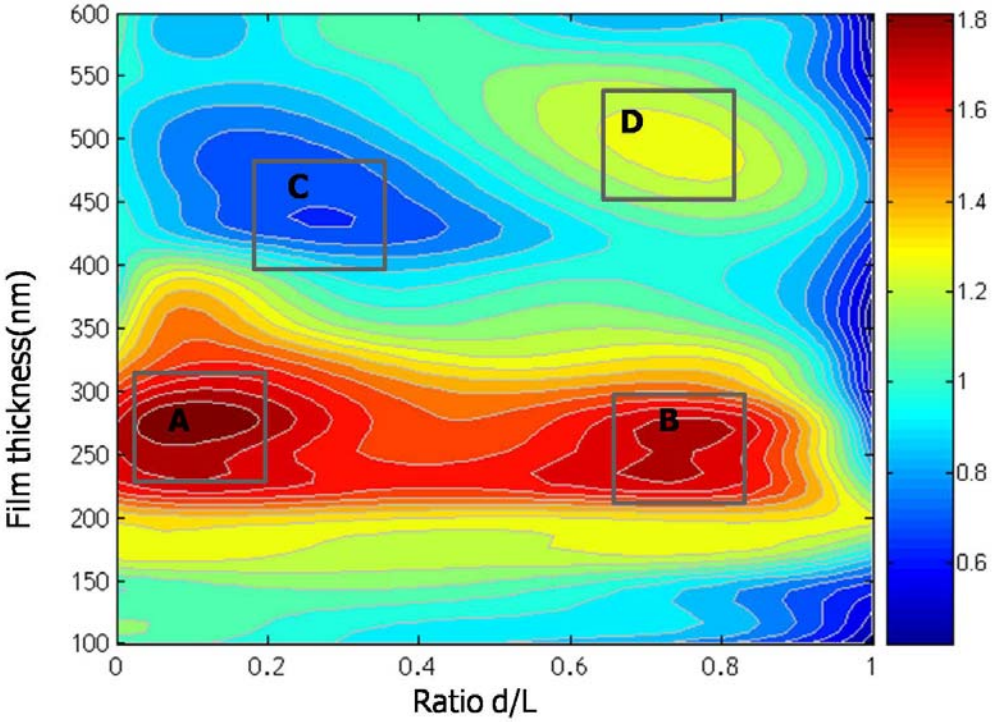


Fig. 3-19 (a) E_x field distribution at the entrance plane for composite aperture. (b) The comparison of E_x intensity at cross section y=0 at the entrance comparison between composite and slit aperture.

The composite aperture can be viewed as an asymmetric energy transportation channel. The Fabry Perot-like behavior along this channel is hardly expressed in the analytical form. Therefore, we vary the arm-grooves depth d and the film thickness L to observe the relationship between the PT value and the geometry parameters. Where d/L=0 (or d=0) represents that the aperture is a slit. And the C-shaped aperture has d/L=1 (or d=L). Otherwise, it is a composite aperture we proposed.

Fig. 3-20 is the PT for the different film thicknesses L with various arm groove depth d. Each point corresponds to a specific geometrical structure. The colors correspond to different PT performance. The four local extreme, A~D, are conditions of four different energy transportation modes, as shown below.


Film thickness(nm)
Ratio d/L
A
B
C
D

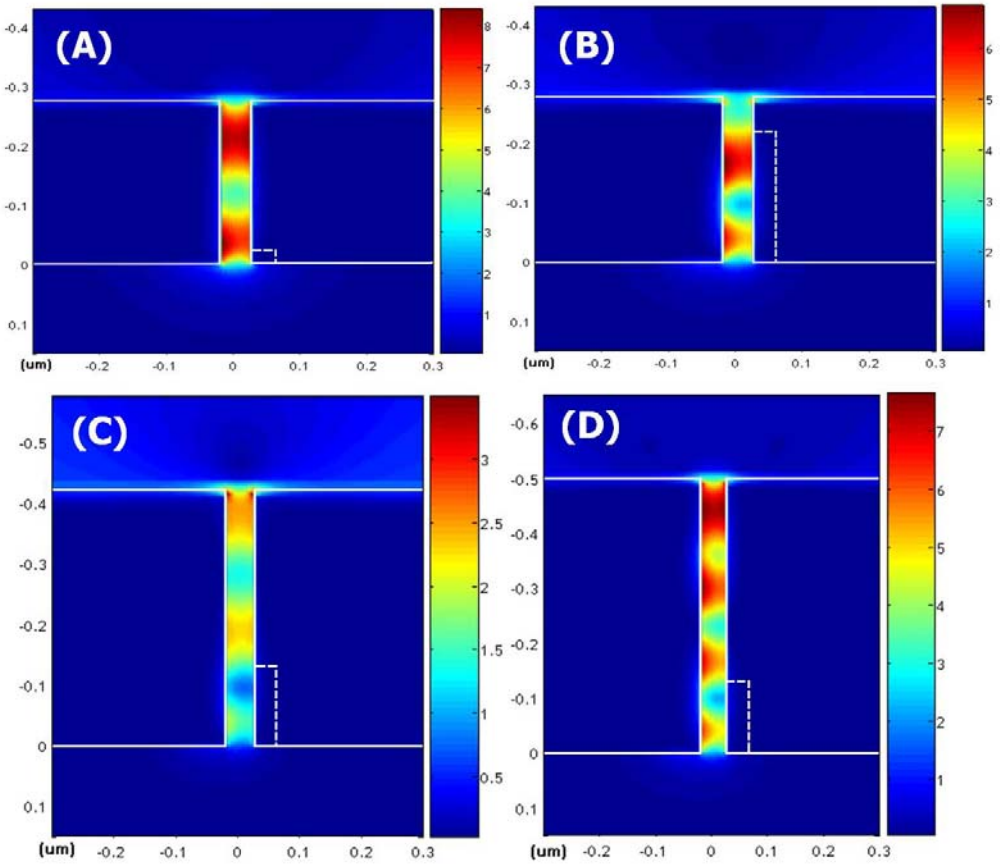

(A)
(B)
(C)
(D)
(um)

Fig. 3-20 The illustration of PT for apertures with different film thicknesses and various groove depths. The four local peaks correspond to four different transportation mode: L=275nm (A) ratio=0.1 and (B) ratio=0.75, (C) L=425nm & ratio=0.3,and (D) L=500nm & ratio=0.75. ABD are local maximums; C is local minimum

The results show that PT oscillates not only with the film thicknesses but also with the ratios. When the thickness is so thin that the localized surface plasmon on both sides will interact with each other and the light could directly penetrates the metal screen. When film is 100~200nm thick, the lower ratio structures have higher PT. It is beyond 200nm, there are two oscillation peaks along the channel. Then scope of this thesis, the PT would gradually decay with the increasing film thickness. In other words at the interval 200~300nm, two transmission enhancement peaks corresponding to the maximum value of PT appear.

We suppose that the various boundaries at the entrance plane might affect the transmission but could not reform the spot size (or field distribution). On the contrary, the different exit boundary will reform the spot performance directly. In FDTD simulation, we choose the parameter L=275nm, which the PT is significant enough to prove our assumption. For (a) pure slit- aperture, (b) pure C-aperture, and (c) slit – C composite aperture. All the parameters including width, length, material, and incident condition were consistent with our previous discussion, L=275 nm and d=206.25 nm, as shown in Fig. 3-21. And the results were shown in Tab. 3-2.

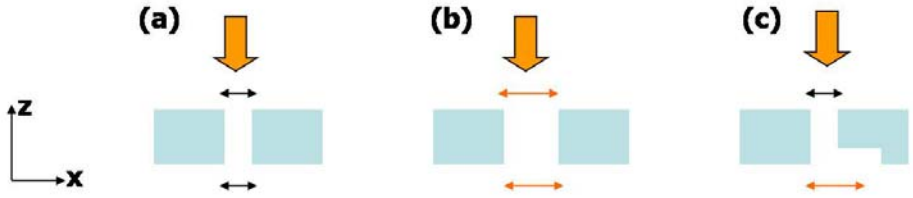


Fig. 3-21 Schematic diagrams of four different apertures, (a) slit, (b) C-aperture and (c) C-at entrance side and slit at exit

Aperture	PT (a.u.)	Spot size (λ^2)	PTD (μm^{-2})
Slit	1	0.049	1
C	0.4	0.038	0.52
Slit-C	1.02	0.04	1.27

Tab. 3-2 Comparisons of the aperture in Fig. 3-21

PT and PTD are normalized to a bare slit

3.4 Brief Summary

The mechanism of extraordinary transmission of a nano-aperture has been suggested by localized surface plasmon effect at the entrance plane and Febry Perot-like resonance inside the aperture. The results in 3D FDTD simulation showed that geometry of the opening at the exit plane plays a crucial role for spot size performance. Fig. 3-22 illustrates the enhancement mechanism of a single nano-aperture clearly.

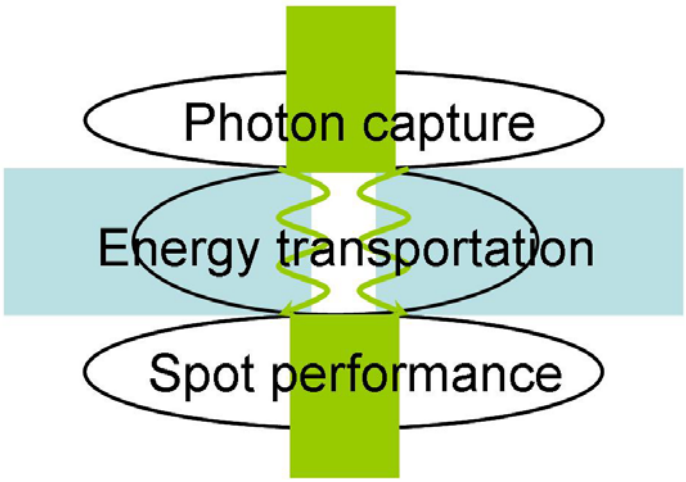


Fig. 3-22 Conceptual scheme of LSP-assisted transmission mechanism which triggered the design of composite structure.

Therefore, a new composite structure of aperture with slit-form at the incident plane and C-shape at the exit plane has been demonstrated. It maintains the efficient photon capture ability as the slit and converges the beam similar to the C-shaped aperture.

Chapter 4

Corrugated Composite Aperture

The enhancement mechanism of the single aperture has been discussed detailed previously. The aperture perpendicular to the incident polarization excites localized surface plasmon most efficiently. However, the incident light does not interact with the "extended surface energy" on a smooth film because the energy and momentum conservation requirements are not obeyed simultaneously. If a periodic surface structure is present, which provides the grating couple, the transmission through the aperture is greatly enhanced compared to that through an aperture in a smooth metal film. In this section, the enhanced transmission of adjacent slits pair resulting from communication between two individual apertures will be presented at the beginning. Then, the 2D analysis on "extended surface plasmon" for the periodic corrugations surrounding the aperture follows. In the reality, the corrugation shape for the 3D aperture will be discussed in the end.

4.1 Extended Surface Plasmon

Since extraordinary transmission of sub-wavelength hole arrays was published by Ebbessen *et al* [3]-[12], the relation between single-aperture transmission has been a topic of debate. [25] Prof. M. Mansuripur in the University of Arizona confines the attention to the interaction between a pair of slits in a fairly thick metal film, and present preliminary simulation results that clarify the nature of cooperation between

adjacent apertures: a cooperation that results in enhanced transmission.

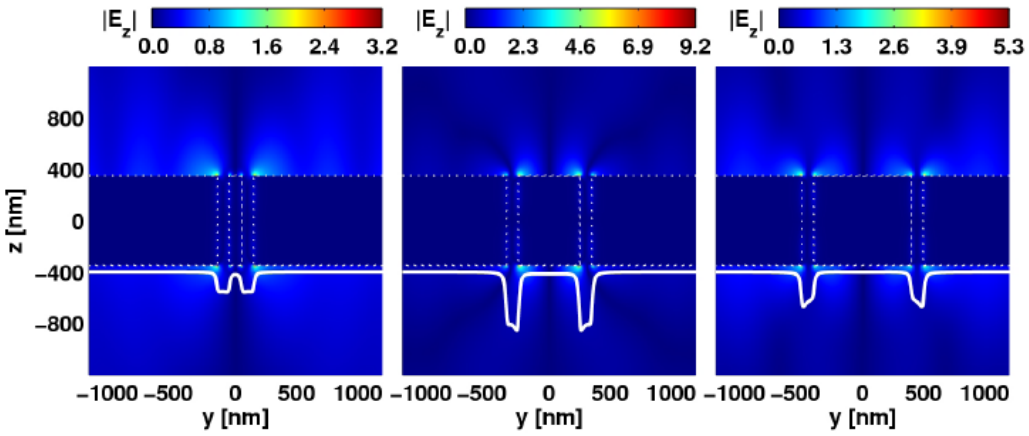


Fig. 4-1 Computed plots of E_z in the case $E_{\perp}$ illumination with λ_o=1000nm, showing the interaction between a pair of adjacent slits (W=100nm) in a 700nm-thick silver film. Left to right: center-to-center spacing of the slits is d=200nm, 500nm, 900nm. The white curves beneath each slit show S_z at the exit facet.

As Fig. 4-1 shown, when d=200nm, the separation is so small that the positive and negative charges around the edges of the two slits tend to cancel each other out , thus weakening the throughput of both slits. As the distance between the slits widens to ~500nm, the interaction between the surface charges and currents near the edges of the slits intensifies and the transmission reaches its maximum, which is ~50% greater than that for each slit alone. Beyond d=500nm, the interaction weakens and, by the time d reaches ~900nm, the two slits have begun to act more or less independently.

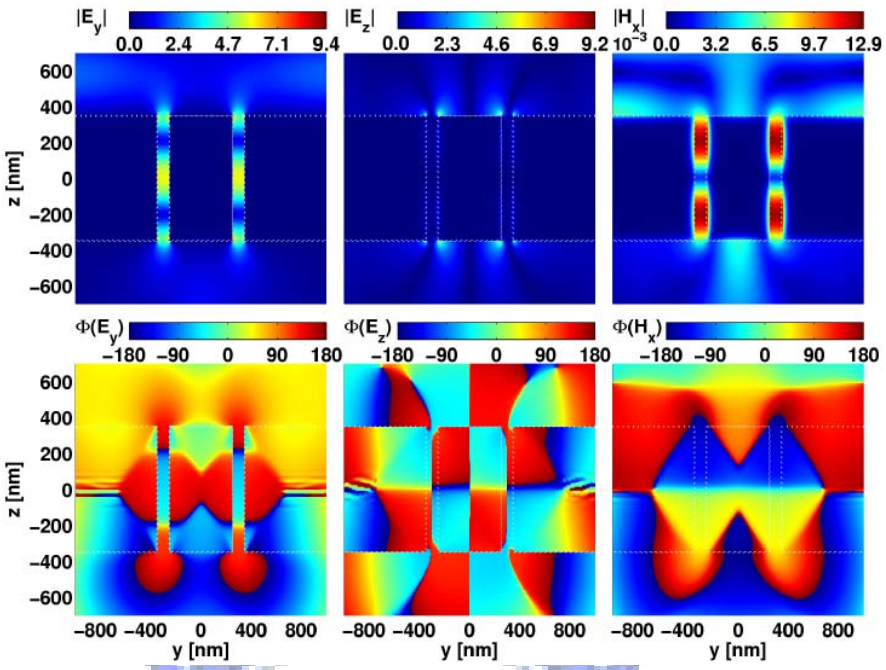


Fig. 4-2 Left to right: plots of Ey,Ex,Hx for a pair of adjacent slits (W=100nm) in a 700nm-thick silver film under E illumination: (top) magnitude, (bottom) phase. The center-to-center spacing of the slits is *d*=600nm.

The center-to-center spacing of the slit is adjusted to *d*=600nm. In addition to the asymmetry distribution charge near the slit's edges (visible in the $|E_z|$ plot), the presence of a strong surface current (in the $|H_x|$ plot) in between the two slits on the film's bottom facet. Inside each slit the fields have left-right symmetry, which indicates the absence of excited odd modes. We believe the efficiency improvement is due to an increase in the effective coupling coefficient between beam and the guide mode of each slit.

French Prof. T.W. Ebbesen expands the previous interaction between two separated apertures by corrugating a metal surface in the vicinity of a slit [10], as depicted in Fig. 4-3.

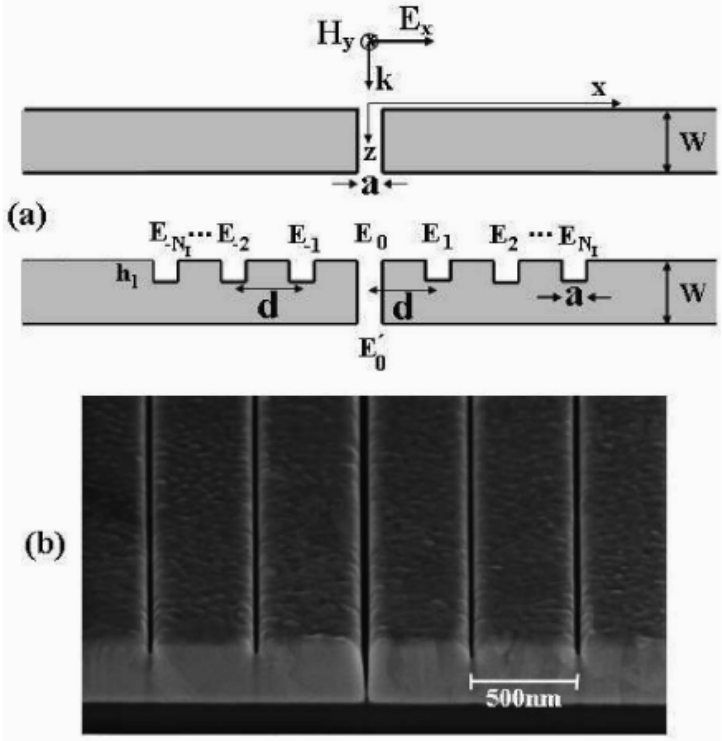


Fig. 4-3. (a) Schematic pictures of a single slit of width a in a metallic film of thickness W symmetrically surrounded in the input surface by $2N_1$ grooves of depth h_1. The separation between adjacent indentations is d and all groove widths are also a. A sketch of the p-polarized normal incident radiation is also shown. (b) Electron micrograph image of the devices with a=40nm, W=350nm, h_1=100nm, and d=500nm.

He identifies three main mechanisms for this structure that can enhance optical transmission: groove cavity mode excitation (controlled by the depth of the grooves), in-phase groove reemission (controlled by the period of the groove array), and slit waveguide mode (controlled by the thickness of the metal film). His simulation results are shown as Fig. 4-4

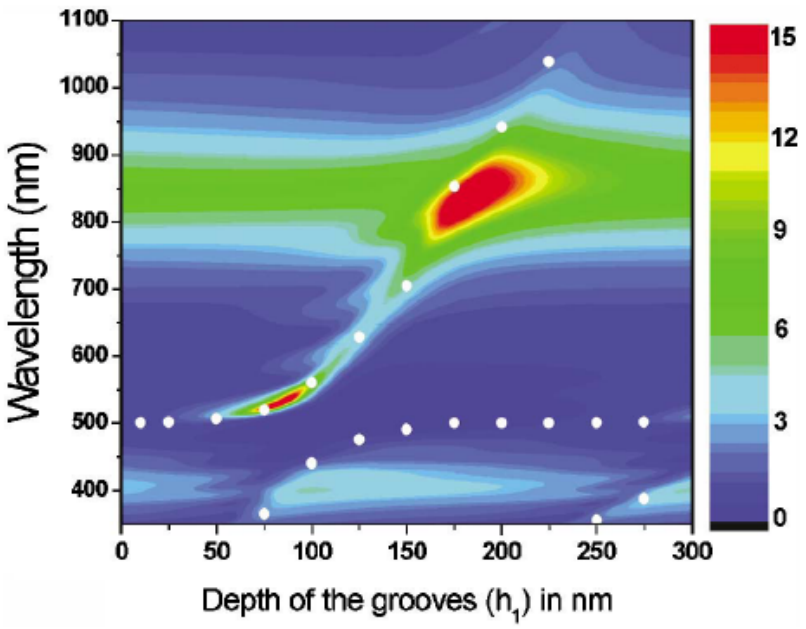


Fig. 4-4 Contour plot of calculated transmittance v.s. both groove depth and wavelength for 10 grooves, a=40nm, d=500nm, and W=350nm. White dots correspond to the locations of surface EM modes of an infinite 1D array of rectangular grooves with different depths (h_1).

The plot illustrates how the different mechanisms influence the transmittance. There are several regions where transmission is enhanced: the slit waveguide modes at $\lambda \approx 400$nm and 850nm, and the line that stars at $\lambda = d$=500nm for small h_1 going on as $\lambda \approx 4h_1$ for large h_1. It can be shown that the latter line corresponds to the excitation of a surface EM resonance of the corrugated input metal surface, originated by the interplay between the groove cavity mode and the in-phase groove reemission mechanisms. Additional insight into this surface resonance can be gained by considering what occurs in reflection gratings. Their absorption anomalies are also due to the excitation of surface resonances. We have calculated the spectral positions of surface EM modes of a reflection grating with the same parameters as our finite structure and represented them in Fig. 4-4 (white dots). The plot clearly shows that surface modes in reflection gratings originate from the anti-crossing of two primary

modes, each one of them associated with the two mechanisms described above. It also shows that the transmission is further enhanced when two mechanisms coincide. This is the case for $\lambda \approx 850nm$, $h_1 \approx 175nm$ (coincidence of slit and grooves cavity resonance),and $\lambda \approx 525nm$, $h_1 \approx 75nm$ (when the mixing of groove resonance and in-phase reemission is important). A huge boost in transmission will be predicted when all three mechanism coincide. In this case, the transmission is boosted by an additional factor of 10 with the optimal parameters.

4.2 Corrugations for Composite Aperture

The mechanism for the sub-wavelength aperture surrounded with periodic corrugation can be decomposed into two parts. Fig. 4-5 illustrates it clearly: the incident polarization perpendicular to the slit excites the localized surface plasmon (LSP) to be accumulated around the edges as the discussion in the Chapter 3; the periodic structure scatters incident light by supplying extra horizontal momentum. The valence electrons are induced to vibrate collectively as "extended surface plasmon resonance". However, localized surface plasmon is confined at the aperture nearby; extended surface plasmon resonance could propagate through the gratings to a certain range.

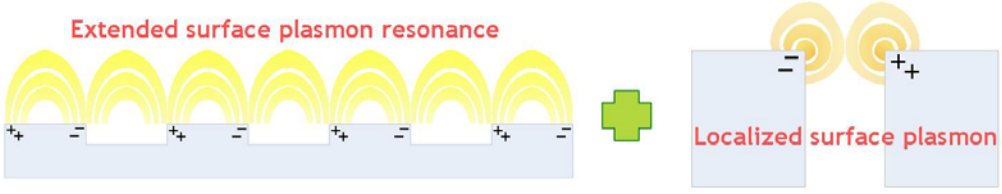


Fig. 4-5 The illustration of the "Extended surface plasmon resonance" and "Localized surface plasmon"

We broaden the corrugation width wider than the composite aperture. The accumulated localized surface plasmon also appears at the edge of the corrugation. And the extended surface plasmon resonance has been still excited under specific corrugation periodicity. The extended surface plasmon resonance further improve the concentration of localized surface plasmon around the aperture, as shown in Fig. 4-6. The central LSP intensity is stronger than that of a single aperture with the help of the extended surface plsmon. Generally speaking, the entrance corrugation improves the photon capture ability and enlarges the power throughput enhancement of the single aperture.

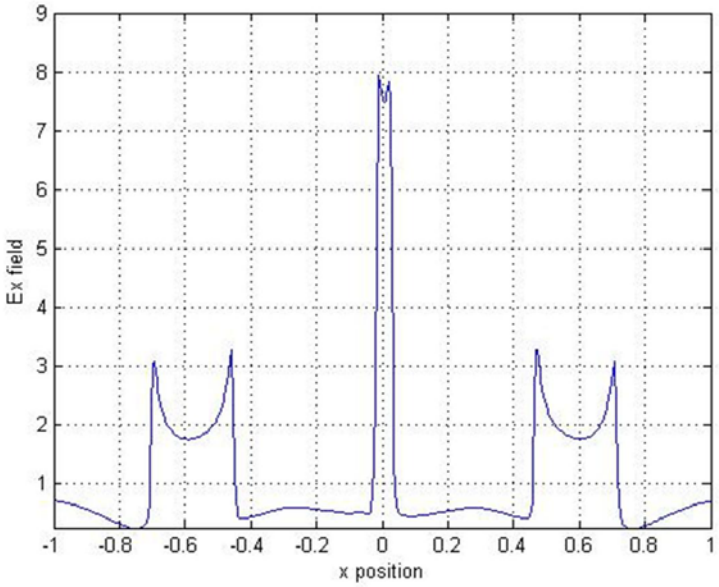


Fig. 4-6. Ex field distribution at cross section at y=0 for the composite aperture surrounded by the corrugation.

With a suitable corrugation periodicity for the extended surface plasmon, the incident p wave couples to the LSP most efficiently. However, the previous discussion is restricted in x cross section. The next energy distribution is expanded to entrance x-y plane. Such as Fig. 4-7 (a), Slit corrugation is disable to manipulate the y-component of incidence. Meanwhile, the circular corrugation around the aperture excites the extended surface plasmon in radial direction, as presented in Fig. 4-7 (b).

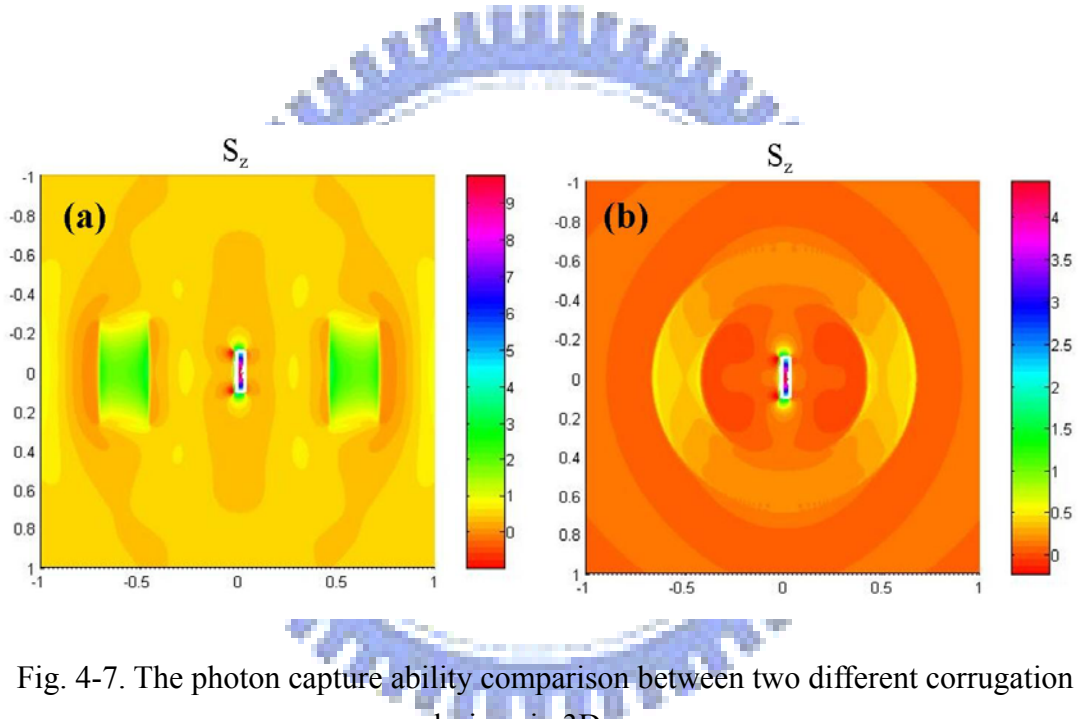


Fig. 4-7. The photon capture ability comparison between two different corrugation designs in 3D case

As a result, both cases sustain the small spot size as the C-shaped aperture. Nevertheless, the aperture in case (b) with the intact circular extended surface plsmon wave in the entrance plane performs the better power throughput enhancement than case (a), as the comparison in Tab. 4-1.

Aperture	PT (a.u.)	Spot size (λ^2)	PTD (μm^{-2})
Corrugation a	2.22	0.038	144.63
Corrugation b	3.15	0.038	199.03

Tab. 4-1 Output performance comparison between two different 3D corrugations

4.3 Summary

For a single aperture, only the localized surface plasmon is accumulated at the aperture edge. Incident light hasn't enough momentum to induce collectively electronic motion in a resonant way. So the fraction of the total power coupled into surface plasmon is unsatisfactory. Therefore, the entrance periodic grating, which scatters incident photons by supplying extra horizontal momentum extra horizontal momentum, is added to the single aperture.

However, our composite aperture with finite size is different with the previous works on 2D analysis on the infinite length aperture. The localized surface plasmon around the aperture edges behaves as a electric dipole source. The circular corrugation coincides with the concentric field distribution of a dipole and ultimately enhances the coupling efficiency of incident light.

.

Chapter 5

Conclusions and Future Works

5.1 Conclusions

In the beginning, the Bethe's formula claimed the nano-aperture can produce the tiny spot beyond the diffraction limit. However, its throughput is too weak to be used. The practical nano-aperture breaks the Bethe's theoretical assumptions: infinitely thin and perfect conducting film. It achieves the extraordinary transmission enhancement in reality. The mechanism of nano-aperture is divided into three parts: photon capture ability determined by the entrance geometry, energy transportation influenced by the film thickness, and the exit spot performance.

Expanded from the square aperture, the ridge-based aperture enhances PT powerfully because of the "localized surface plasmon" which is accumulated along the ridge. Furthermore, the improved composite aperture from the combination between the slit and the C-shaped aperture transforms the "Fabry Perot-like resonance" condition to maintain the intense transmission and the tiny spot simultaneously. Finally, the entrance corrugation scatters incident photons by supplying the additional horizontal momentum to induce the "extended surface plasmon" which increases the photon capture ability. The mechanism is organized as Tab. 5-1

Aperture	Square	Ridge-based	Composite	Corrugated composite
Mechanism	Diffraction limit	Localized surface plasmon	Fabry Perot-like resonance	Extended surface plasmon resonance

Tab. 5-1 Mechanism for different apertures

The comparison shows that the proposed corrugated composite aperture produces almost 200x times larger PT than the comparable square aperture. It greatly improves the original feeble throughput.

5.2 Future Work

Our proposed corrugated composite aperture should be milled structures on both sides by FIB (focused ion beam). Consequently the free standing silver film is required by RIE (reactive-ion etching), as shown in Tab. 5-1.

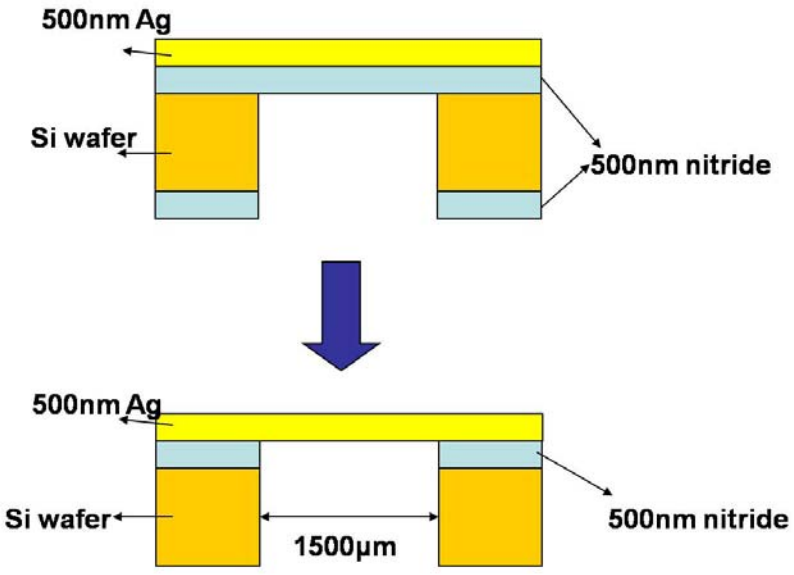


Fig. 5-1 Manufacture process RIE (reactive-ion etching) for free standing silver membrane

However, the horizontal accuracy for FIB is only a quarter of the milling depth. The main issue, ion-charging, which will decrease the manufacture accuracy appears when the metal film is too thick.

The dilemma results from the thin membrane being easily to be broken under the RIE process. How to make balance between these two fabrications is currently a great challenge in existing equipment..

Reference

[1] E. Betzig, J.K. Trautman, R.Wolfe, E.M. Gyorgy, P.L. Finn, M.H. Kryder, and C.-H. Chang, "Near-field magneto-optics and high density data storage," *Appl. Phys. Lett.*,**61**, p.142 (1992)

[2] H.A. Bethe, "Theory of diffraction by small holes", *Phys. Rev.*,**66**, p.163 (1944).

[3] T.W. Ebbesen, H.J. Lezec, H.F. Ghaemi, T. Thio & P.A. Wolff, "Extraordinary optical transmission through sub-wavelength hole arrays", *Nature*,**391**, p.667 (1998).

[4] H.F. Ghaemi, T. Thio, D.E. Grupp, T.W. Ebbesen, and H. Lezec, "Surface plasmons enhance optical transmission through subwavelength holes", *Phys. Rev. B* ,**58**, p.6779 (1998).

[5] T.Thio, J.J. Lezec, and T.W. Ebbesen, "Strongly enhanced optical transmission through sub-wavelength holes in metal films", *Phycia B* ,**279**, p. 90 (2000)

[6] T.Thio, K.M. Pellerin, R.A. Linke, H.J. Lezec, and T.W. Ebbesen, "Enhanced optical transmission through a single sub-wavelength aperture," *Opt. Lett.*,**26**, p.1972 (2001)

[7] T. Thio, J.J. Lezec, T.W. Ebbesen, K.M. Pellerin, G.D. Lewen, A. Nahata, and R. A. Linke,"Giant optical transmission of sub-wavelength apertures: Physics and applications", *Nanotechnology* ,**13**, p.429 (2002)

[8] H.J. Lezec, A. Degiron, E. Devaux, R.A. Linke, L. Martìn-Moreno, F.J. Garcìa-Vidal, and T.W. Ebbesen, "Beaming light from a sub-wavelength aperture", *Science* ,**297**, p.820 (2002)

[9] W.L. Barnes A. Dereux & T.W. Ebbesen,"Surface plasmon sub-wavelength optics", *Nature* ,**424**, p.824 (2003).

[10] F.J. Garcìa-Vidal, H.J. Lezec, T.W. Ebbesen, and L. Martìn-Moreno, "Multiple paths to enhance optical transmission through a single sub-wavelength slit", *Phys. Rev. Lett.* ,**90**(21), p.213901 (2003)

[11] A. Degiron and T.W. Ebbesen, "Analysis of the transmssion process through single aperture surrounded by periodic corrugations", *Opt. Exp*. **12**(16), p.3694 (2004)

[12] A. Degiron, H.J. Lezec, N. Yamamoto, T.W. Ebbesen, "Optical transmission properties of a single sub-wavelength aperture in a real metal", *Opt. Commun.*,**239**,p.61 (2004).

[13] H. Raether, "*Surface Plasmons on Smooth and Rough Surfaces and on*

Gratings",(Springer-Verlag,Berlin,1988).

[14] J. D. Jackson, "*Classical Electrodynamics*", 3rd edition, Wiley, New York, 1999.

[15] H.J. Lezec, and T. Thio, " Diffracted evanescent wave model for enhanced and suppressed optical transmission through sub-wavelength hole arrays", *Opt. Express*, **12**(16), p.3629 (2004)

[16] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Ant. Prop.* **14**, p.302 (1966).

[17] A. Taflove, and S.C. Hagness, "*Computational electrodynamics: The Finite-Difference Time-Domain method"*, 2nd Ed Artech House Publishers, Boston, (2000).

[18] K.S. Kunz and R.J. Luebbers, *"The Finite Difference Time Domain Method for Electromagnetics"*, Boca Raton, CRC Press (1993)

[19] A. Taflove, *"Advances in Computational Electrodynamics"*, Boston, Artech House (2001)

[20] P. Drude, "Research of the Transmission Ehancement through a Nano-Scaled Metallic Slit", *Ann. Phys., Lpz.* **1** ,p.566 (1900)

[21] E.D. Palik, "*Handbook of optical constants of solids*", New York: Academic Press (1985)

[22] T. Vallius, J. Turunen ,M. Mansuripur, and S. Honkanen," Transmission through single sub-wavelength aperture in thin metal films and effects of surface plasmons", . *J. Opt. Soc. Am. A*, **21,** p.456 (2004)

[23] Y. Xie, A.R. Zakharian, J.V. Moloney, and M. Mansuripur, "Transmission of light through periodic arrays of sub-wavelength slits in metallic hosts", *Opt. Express* **14**, p.6400 (2006).

[24] A.R. Zakharian, J.V. Moloney, and M. Mansuripur, "Surface plasmon polaritons on metallic surfaces",*Opt. Express.* **15**, p.183 (2007)

[25] Y. Xie, A.R. Zakharian, J.V. Moloney, and M. Mansuripur, "Transmission of light through slit apertures in metallic films," *Opt. Express* **12**, p.6106 (2004),

[26] J. Helszajn: Ridge waveguides and passive microwave components (IEE, London, 2000) p. 26

[27] M. Mansuripur, " Radiation Pressure on Submerged Mirrors: Implications for the Momentum of Light in Dielectric Media", *Opt. Express*, **15**, p.2677 (2007)