# Chapter 1

Introduction

1.1 History

The first optical recording technology was invented about 5000 years ago, when the Egyptians started using hieroglyphic and alphabetic writing systems. The precursors to the modern data storage system are the tape drive, hard drive, and the compact disk. Today optical data storage refers to storage systems that use light for recording and retrieval of information. Light can be easily used for digital information recording, especially since the invention of the lasers.
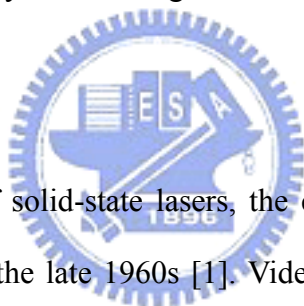
Later with the advent of solid-state lasers, the compact disk (CD) was invented by Klass Compaan at Philips in the late 1960s [1]. Video compact disks (VCDs) and digital video disks (DVDs) were introduced in the 1980s and 90s, respectively.

1.2 overview of data storage

Computer memory can be broadly classified into three groups: primary, secondary, and tertiary [2]. The memory whose data can be obtained by the computer system on a time comparable to the system clock is called the primary memory. Secondary memory comprises memory such as the hard drive and the optical drives. The data from this type of memory is obtained on a longer time scale compared to the primary memory.

Data storage systems from one of the most sophisticated systems involving concepts

from different disciplines. The fundamental physics of recording the servo systems that maintain the position of the recording heads, the signal processing that goes into reducing the bit error rate for a noisy signal, and the lubrication and aerodynamics of the head-disk interface, validate this claim.

Based on the physical principle of storage, conventional data storage systems can be classified into magnetic and optical data storage systems [3]. In magnetic data storage systems, the information stored in the groups with different magnetized direction is written by a magnetic field and is read out by a magnetic field sensitive transducer. In optical data storage systems, the information is stored in more than one way [4-7]. It could be stored in terms of physical marks on the disk as in the case of Compact Disk Read only Memory (CD-ROMs) and Digital Versatile Disks (DVDs), in the form of the magnetic state of a group of atoms as in the case of Magneto-Optical (MO) disks, or in the form of the crystalline order of a group of atoms as in the case of Phase Change (PC) disks. The most popular types include read-only memory (ROM) such as CD-ROMs, write-once read-many (WORM) such as CD-R (CD-recordable) and rewritable such as CD-RWs and MO disks. Optical drives of all kinds operate on the same principle of detecting variations in the optical properties of the media surface. CD and DVD drives detect changes in the light intensity, MO drives - changes in the light polarization. All optical storage systems work with reflected light.

Optical recording systems potentially have much greater reliability than magnetic recording systems since there is a much larger spacing between the read/write element and the moving media. Therefore, there is no wear associated with repeated use of the optical systems. However, there are other possible souses of trouble: the life and stability of the laser, mechanical damage to the relatively soft and exposed-to-the-environment media, mechanical damage due to shock and vibration, etc. Another advantage of the optical

recording systems over the best performing magnetic recording systems - hard drives - is their removability.

Data storage systems can also be classified as removable or non-removable. This classification is based on whether the storage medium can be separated physically from the writing or the reading mechanisms (write or read head). In magnetic hard drives, the head-disk interface is delicate and has to be sealed to prevent contamination by dust. Hence the head cannot be separated from the disk. Other removable magnetic systems like tape and ZIP disks [8] exist. Most optical data storage systems are removable. The systems are not very sensitive to contamination by dust. However, in the future, optical systems that depend on near-field head-disk coupling might be non-removable.

The main disadvantages of optical storage when compared to magnetic are slower random data access. This partially comes from the design of the relatively large (and heavy) optical heads. Moving 100 grams over the disks at the high acceleration and speed needed to match 5 to 10 ms average access time for magnetic hard disk drives is a real challenge, since the effective weight of the moving parts of a hard drive head-stack assembly (actuator arm with suspensions and sliders) is on the order of only a few grams. Plus, unlike the hard disk, an optical disk is usually removable, which limits rotational velocities and, correspondingly, limits access time. Increasing rpm causes the relatively loosely fixed CD-ROM disk to vibrate significantly compared to a stiff, fixed, and balanced hard disk.

1.3 Hybrid recording

Storage densities of hard disk drives using conventional longitudinal magnetic recording have been increasing at rate between 60% and 100% per year. The most recent

laboratory systems have demonstrated storage densities in excess of 100 Gbit/in^2 [9]. However, this trend will likely reach physical diffraction limits at densities below 1 Tbit/in^2. It is the so-called superparamagnetic limit described as follows.

The need to maintain a given signal-to noise ratio (SNR) requires that the size of the grains within a bit must be scaled with the area of the bit. However as the size of the grains is reduced, thermal instability happens. For the recorded magnetization to be thermally stable for a period of ten years the following relation must hold [10]:

$$\frac{K_u V}{k_B T} = 40$$

Here K_u is the media anisotropy energy density, V is the grain volume, k_B is Boltzman's constant and T is the temperature of the medium. Thus, as the volume of the grains is reduced the anisotropy of the material must be increased. However, the coercive field, H_c, is proportional to the anisotropy and any increase in K_u results in an increase in H_c, ultimately making the media unwriteable since it is not possible to generate adequately large write fields with conventional magnetic heads.

These considerations have motivated considerable work in the field of hybrid recording [11]-[15], which uses a heat source such as a focused optical beam from a laser to reduce the coercivity of the medium during the writing process. Thermally assisted magnetic recording that combines the benefits of magnetic and magneto-optical recording was proposed to achieve very high density storage. Perpendicular magnetic recording has attracted much attention for its less thermal fluctuation at very high recording density than in the longitudinal recording media. Such a system allows for the use of high anisotropy

media which are quite stable at room temperature and which could not otherwise be written with conventional thin film heads. Read back can be accomplished with a magnetoresistive sensor. For this type of recording system to become truly practical, however, an integrated read/write head must be developed that brings the optical field, the magnetic field, and the read sensor to the medium in a manner consistent with densities of 1Tbit/in^2 and beyond. The thesis is primarily concerned with the optical data storage systems. However, recent changes in the field of magnetic data storage systems, makes this thesis relevant to heat assisted magnetic recording systems.

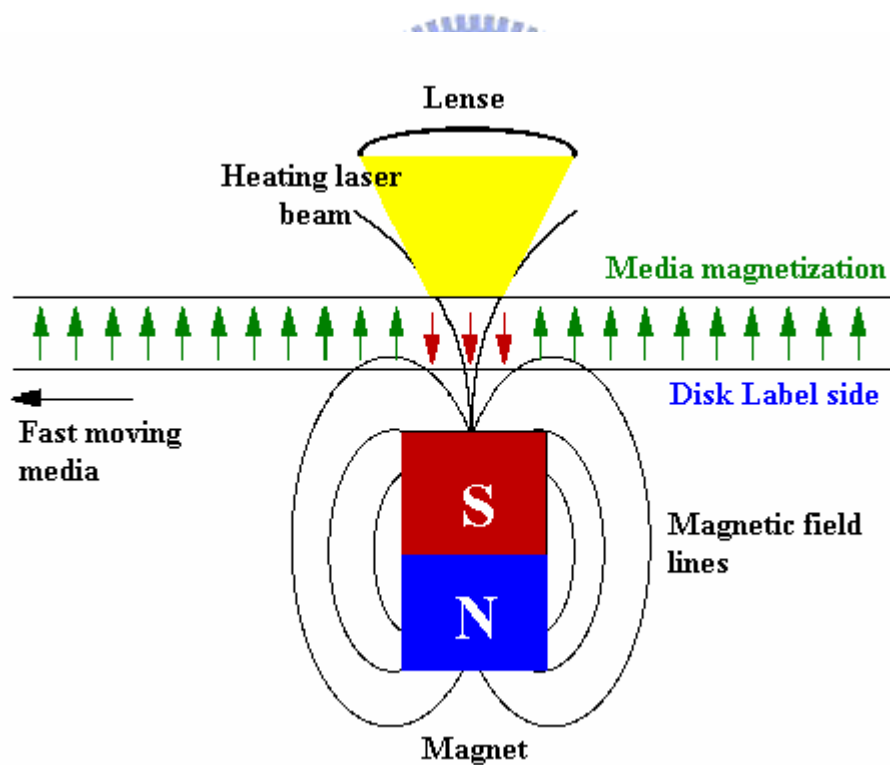### 1.3.1 Magneto-optical writing process

Fig. 1-1 Magneto-optical writing process

In magneto-optical recording systems, signal recording is accomplished by thermo-magnetically writing reversed magnetic domains on a pre-saturated medium.

The recording medium usually has a magnetic easy axis along the film normal such that the magnetization lies perpendicular to the film plane. The coercivity, which is on the order of several thousand oersteds at room temperature, decreases with increasing temperature. Therefore, recording is achieved by heating a micron or sub-micron area on the media with a focused laser spot with a moderately high power (5-10mW) under the influence of a small bias magnetic field of several hundred oersteds. The bias magnetic field is lower than the coercivity of the media at room temperature but higher than the coercivity of the region heated by the laser beam. Therefore, only the magnetic moments in the locally heated spots will be reversed by the bias magnetic field.

1.3.2 Magneto-optical reading process

Unlike traditional magnetic recording systems, which use currents induced in the magnetic heads by the changing magnetic fluxes on the disk surface to read the data, MO systems use polarized light to read the data from the disk.The polarization of reflect light is changed from the incident light because of the direction of magnetization of domain. If a beam of polarized light is impinged on the surface, the light polarization of the reflected beam will change slightly as shown in Fig. 1-2. (typically of less than 0.5^{o}) if it is reflected from a magnetized surface. If the magnetization is reversed, the change in polarization (the Kerr angle) is reversed, too. The magnetized areas - pits - can not be detected in regular light, but only in polarized light. The change is direction of magnetization could be associated with numbers 0 or 1, making this technique useful for binary data storage.
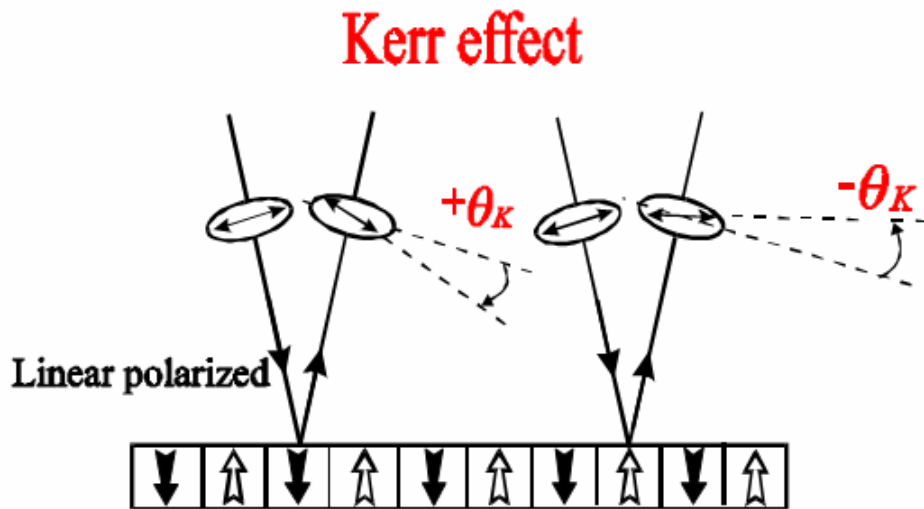
Fig. 1-2 Schematic of magneto-optical Kerr effect

1.4 Miniaturization of pickup heads

Key issues in future pickup improvement include miniaturization, integration, and parallel processing (multiple beams). These technologies are expected to reduce the cost, size, and weight of the opto-mechanical structure, making the device more competitive in terms of price and performance.

Several integrated pickup by MOEMS technology have been proposed to make the head smaller and lighter for fast access, direct overwriting and high data transfer rate. For practical applications, the issues that need to be addressed include:

(1). improvement of fabrication accuracy,

(2). increase of optical throughput power, and

(3). appropriate actuator mechanism for the focusing/tracking servo.

There are many methods will be shown below to achieve miniaturization of pickup heads.

1.4.1 Planar pickup based on glass substrate light guide

Diffractive optical elements (DOEs) can be used to replace individual refractive optics. T. Shiono *et al.* [16] integrated various planar diffractive structures on a glass substrate. The glass substrate was used as a light guide in whom the beam followed a zigzag optical path, as shown in Fig. 1-3. In this approach, the beam hits those diffractive elements in order. The focusing/tracking error detections were based on the foucault/push-pull methods, respectively, which were achieved by the off-axis blazed structure called reflection-twin-focusing beam splitter.
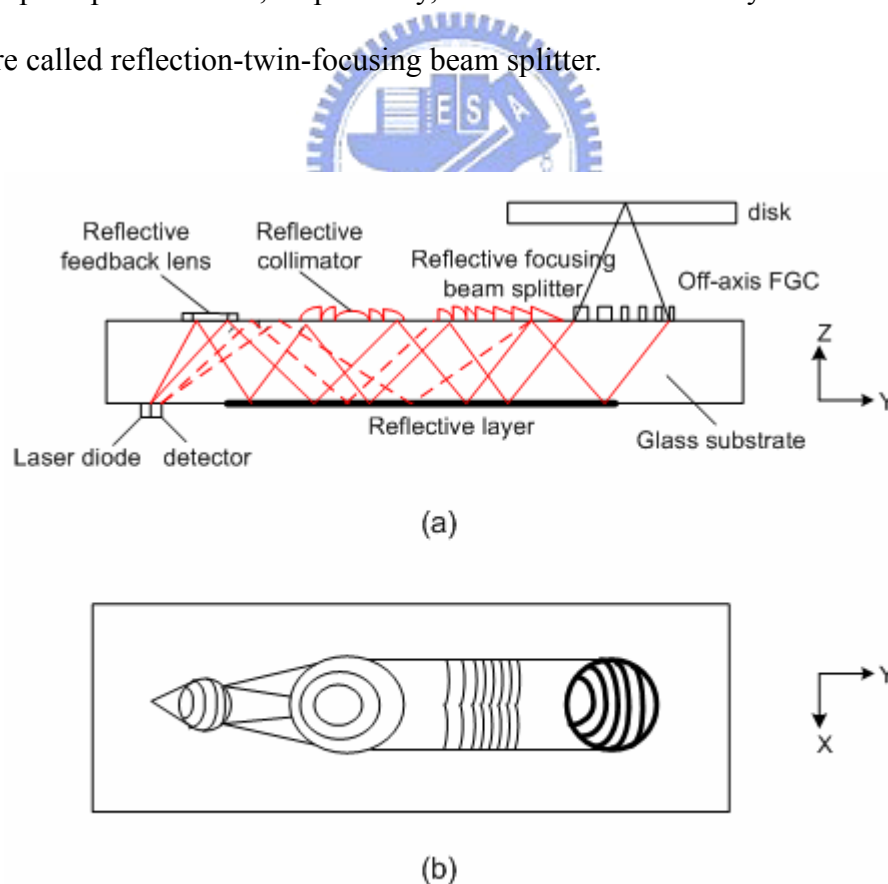
Fig. 1-3 Schematic of the planar optical disk pickup based on glass-substrate based:
(a) side and (b) top view.

The focused spot in FWHM was about 1.0 μm at $\lambda = 0.63\,\mu m$. This value was close to the diffraction limit (0.98 μm). Theoretical efficiency of the system was about 46 %, with reflection loss neglected at each interface. This design required precision alignment on both sides of the substrate and was very sensitive to optical crosstalk due to the high-order diffraction beams.

1.4.2 Free space micro-optical bench

As proposed by M. C. Wu *et al*[17]., a Si substrate was used as a micro-optical bench on which optical elements, micro-positioners, and actuators were monolithically fabricated by silicon surface micromaching. The optical configuration in this approach was the same as the conventional pickup but had an extremely tiny size. The schematic drawing of the pickup is shown in Fig. 1-4. The NA of the focusing micro Fresnel lens was measured to about 0.17. In this structure, fine optical alignment (0.1 μm) and dynamic tracking can be achieved by on-chip micro actuators with electronic driving.
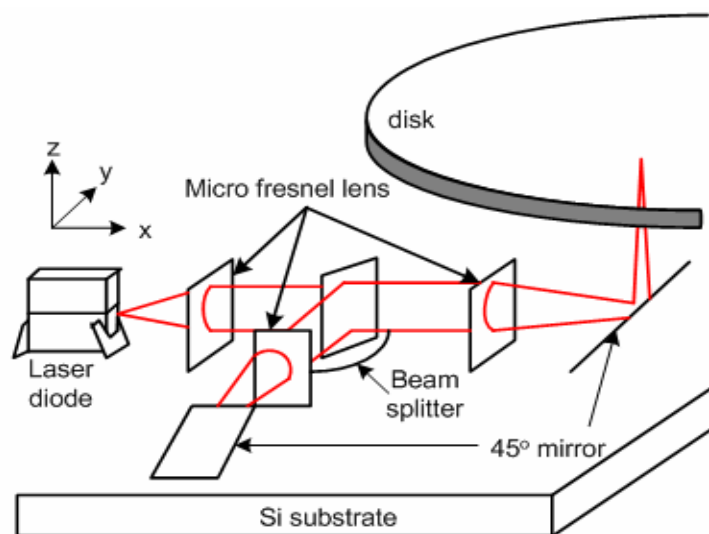

Fig. 1-4 Optical bench module structure

1.4.3 Integrated-optic implementation based on waveguides and holographic components

S. Ura *et al*. [18] employed a focusing grating coupler (FGC) in the pickup. A FGC has a chirped and curved grating pattern to couple a guided wave from the laser diode to a focusing beam in the free space, as shown in Fig. 1-6. Focusing and tracking error detections were based on the foucault/push-pull methods, respectively. The optical system was modified for implementation with diffractive components in a waveguide. The focus spot width was measured as approximately $2\,\mu m$, while the theoretical diffraction-limited value was $1.4\,\mu m$. Major aberrations induced by the fabrication tolerance are astigmatism and coma, which can easily cause the deterioration of the focused spot. Since holographic optical elements (HOE) are dispersive elements, the wavelength deviation is also an issue in the optical design. The fabrication accuracy required to obtain diffraction-limited focused spots becomes move stringent with increasing NA of the focusing grating. For example, the maximum tolerance in the FGC dimension and wavelength deviation are approximately ±0.6% and ±0.9%, respectively, for NA = 0.24. The tolerance is reduced to ±0.09% and ±0.1%, respectively, for NA = 0.45.
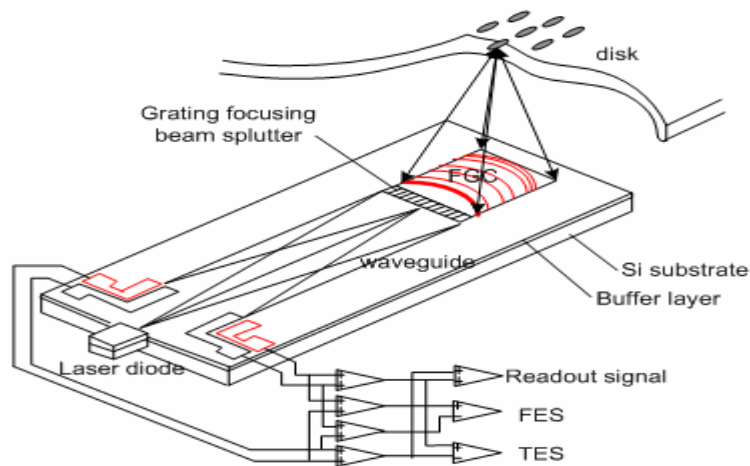
Fig. 1-5 Schematic of the integrated optic disk pickup (IODPU). The curved and chirped grating incorporates the input/output coupling and the focusing by the wavefront conversion based on holography.

1.4.4 Fiber-based integrated optical pickup

The integrated optical pickup can be modified by separating the waveguide chip and the laser diode (LD), and linking them by an optical fiber. In the fiber-based pickup, only the waveguide part is actuated for servo operation. The large reduction in weight of the moving part allows higher speed accessing. Another advantage of the fiber-based pickup is that an optical isolator can be easily inserted between the LD and the fiber to stabilize the lasing. In this thesis, we proposed a novel design to achieve an integrated optic disk pickup (IODPU). In our approach, a microlens formed on the end of single-mode fiber (SMF) is used as a pickup in optical data storage. This fiberlens design can overcome the tight tolerances found in high NA optical path and can reduce the number of components, therefore reducing the cost and weight. Thus, fiber-based system is relatively a good choice for high-speed access in optical flying head (OFH) system.

11

Near-field optics, one of the super resolution techniques, can be used for high recording density. Solid immersion lenses (SIL) and apertures are two near-field optical techniques to generate spot size smaller than the diffraction limit. The optical fiber is combined with a well-controlled mechanical structure including a subwavelength aperture as shown in Fig. 1-6. The laser beam is guided by the fiber, collimated by the front-end fiberlens, and reflected by the 45°mirror. The spot size is further reduced by the near-filed optics including a subwavelength aperture. Smaller spot can read signals with higher contrast and lead to an increase in signal-to-noise ratio (SNR) and recording density. To achieve high recording density and fast data rate, higher tracking precision is also required. The optical head should be miniaturized for fast access motion.

1.5 Research objective

Optical resolution beyond the diffraction limit can be achieved by use of a nano metallic aperture in a near-field optical system. However, the problem encountered is the drastic decrease of detected signal in the region of the subwavelength size because of the use of a much smaller sized aperture to sustain spatial resolution. To solve the problem of the trade-off between signal output and spatial resolution, some methods were proposed such like utilizing a different shape aperture proposed by Xiaolei Shi, an aperture surrounding period corrugation proposed by Tineke Thio [16] or taking use of a SIL proposed by L. Milster [20].

In the thesis, the attempt is made to calculate near-field distribution in vincity of novel subwavelength aperture by Finite-Difference Time-Domain method (FDTD) and characterize the optical properties. Then the results of different shapes of

apertures and SIL-aperture combination and comparisons between them are demonstrated. Furthermore, properties of metals are considered to find the optimum structure to increase power throughput with small spot size. The optimum aperture is fabricated by Focused Ion Beam (FIB), and then the measurement can be carried out by Near-field Scanning Optical Microscope (NSOM) to confirm the simulation results.
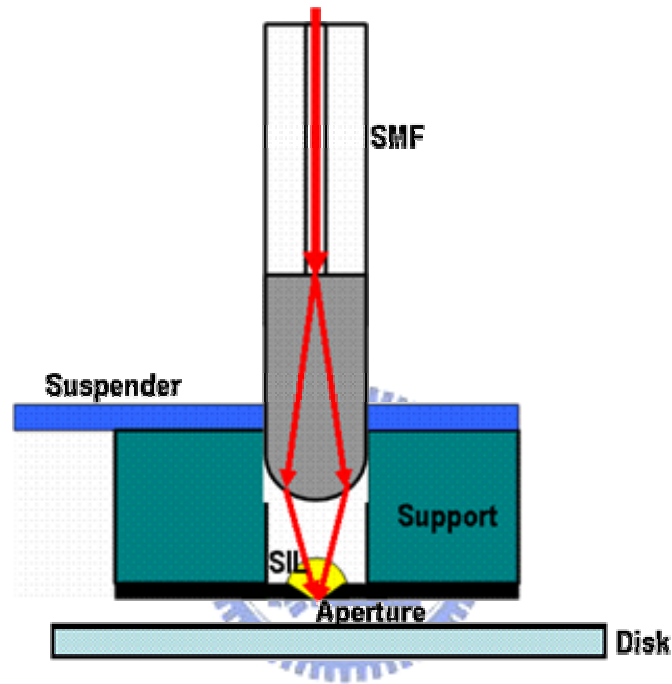

Fig. 1-6 Schematic of the integrated optic disk pickup (IODPU), which integrates a fiberlens, SIL and sub-micrometer aperture.
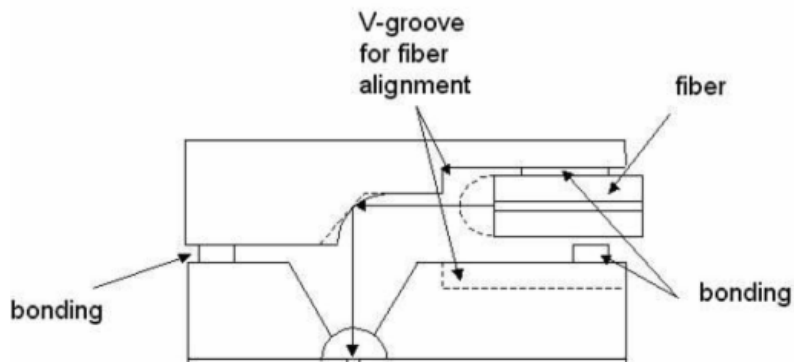

Fig. 1-7 Proposed near-field optical pickup

Chapter 2

Principle

2.1 Introduction

The spot size is determined by the subwavelength aperture in our system [1]-[3]. The electromagnetic interaction and complex boundary condition taking place on the subwavelength scale are relevant to the field distribution. Fresnel and Fraunhofer theories are not appropriate for the study of near-field behavior because they do not adequately account for the behavior of evanescent field. Therefore, the rigorous electromagnetic theories must be used to consider the effect of the presence evanescent fields and the appropriate boundary conditions. Most of the models presented here are based on such rigorous solutions of Maxwell's equations.
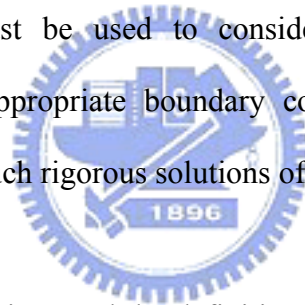
The electromagnetic relations and the definition of power throughput have to be clarified for the subsequent discussions and to normalize the results to compare each other. The results of two kinds of shape: square aperture and c-shape aperture will be introduced. The properties of metals will be taken into considered. Next, the introduction of finite-difference time domain method will be given.

2.2 Electromagnetic Relations

2.2.1 Maxwell's equations and constitutive relations

The Maxwell's equations in stationary materials are given by [4],

$$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \qquad\qquad\qquad \text{(Eq. 2-1)}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad\qquad\qquad \text{(Eq. 2-2)}$$

$$\nabla \cdot \vec{B} = 0 \qquad\qquad\qquad \text{(Eq. 2-3)}$$

$$\nabla \times \vec{B} = \mu_0 (\vec{J} + \frac{\partial \vec{P}}{\partial t} + \nabla \times \vec{M} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t}) \qquad\qquad \text{(Eq. 2-4)}$$

and the constitutive relations are given by

$$\vec{E} = \frac{(\vec{D} - \vec{P})}{\varepsilon_0} \qquad\qquad\qquad \text{(Eq. 2-5)}$$

$$\vec{B} = \mu_0 (\vec{H} + \vec{M}) \qquad\qquad\qquad \text{(Eq. 2-6)}$$

where $\vec{E}$ is the electric field, $\vec{B}$ is the magnetic induction field, $\vec{J}$ is the true current density in the system associated with the physical transport of true charge, $\vec{P}$ is the polarization induced in the system that is consistent with the electric field, $\vec{M}$ is the magnetization, and ρ is the charge in the system(the sum of true charge resulting from the polarization in the system). ε_0 is the permittivity of free space. $\vec{D}$ is the electric displacement vector, and $\vec{H}$ is the magnetic field.

From the equations 2-5 and 2-6, $\vec{E}$ and $\vec{B}$ can be induced by two sources; respectively the true applied charges, current, the polarization charges and magnetization currents.

In our simulation, the materials are assumed non-magnetic and isotropic at a fixed temporal frequency and there are no external currents and charges applied, so $\vec{M} =0$ and $\vec{B} = \mu_0 \vec{H}$. If an external field is applied onto the structure, polarization and current will be induced to influence the fields. The current and the electric displacement vector are given by

$$\vec{J} = \sigma(\omega)\vec{E} \qquad \text{(Eq. 2-7)}$$

$$\vec{D} = \varepsilon(\omega)\vec{E} \qquad \text{(Eq. 2-8)}$$

Because there is no external charge, divergence of $\vec{D}$ is 0

$$\nabla \cdot \vec{D} = 0 \qquad \text{(Eq. 2-9)}$$

Since the current is a result of charge variation, the conservation of charge implies a constraint given by

$$\nabla \cdot \vec{J} + \frac{\partial \rho}{\partial t} = 0 \qquad \text{(Eq. 2-10)}$$

In the absence of external charge or current, $\vec{J}$ and ρ refer to induced current and charge density, respectively.

2.2.2 Energy

When a field is time varying, energy can be transported. The energy transported by the field per unit time across a unit area normal to the direction of transport at any point in space is given by the Poynting vector

$$\vec{S} = \text{Re}(\vec{E}) \times \text{Re}(\vec{H}) \qquad \text{(Eq. 2-11)}$$

The transmitted optical energy is obtained by integrating the normal Poynting vector component, considering it as a far field photo-detector sensed signal when the head travels across the medium. The electric energy distribution $(=|E|^2)$ is calculated for a measure of the near field energy. In order to compare the results under different conditions, Power throughput (P.T.) is defined as the ratio of total incident power over

aperture area to transmitted power [5],

$$P.T. \equiv \frac{\text{(the total transmitted power through the aperture)}}{\text{(incident power density * aperture area)}} \quad \text{(Eq. 2-12)}$$
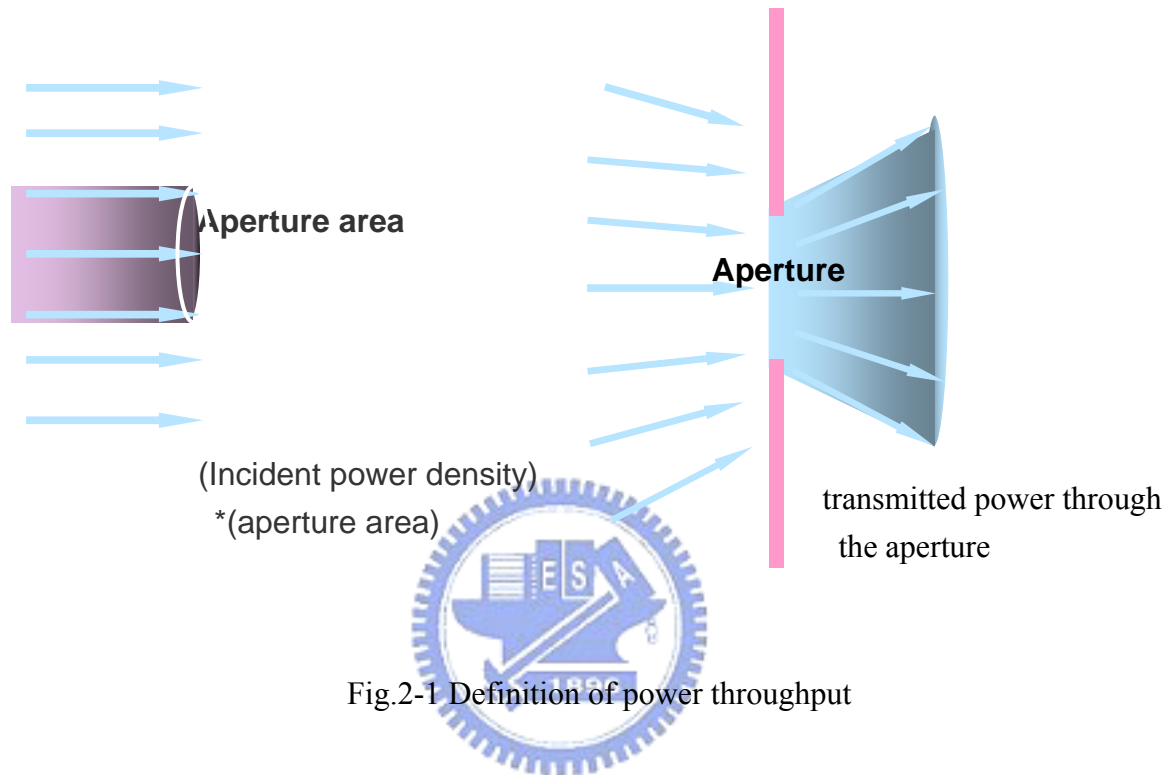

Fig.2-1 Definition of power throughput

2.3 Propagation of light through a subwavelength aperture

The metallic aperture is considered as a metallic waveguide [6]. If the wave frequency is higher than a specific threshold, determined by the mode polarization and by the geometry of the guide normal section, the electromagnetic energy can be transported in the form of a one-dimensional propagating or "radiative" wave. If the frequency is lower than this threshold, there is no possibility of propagation and the field keeps localized at the entrance of the waveguide, in an "evanescent" wave: the field amplitude then decreases rapidly as it penetrates the hollow conductor.

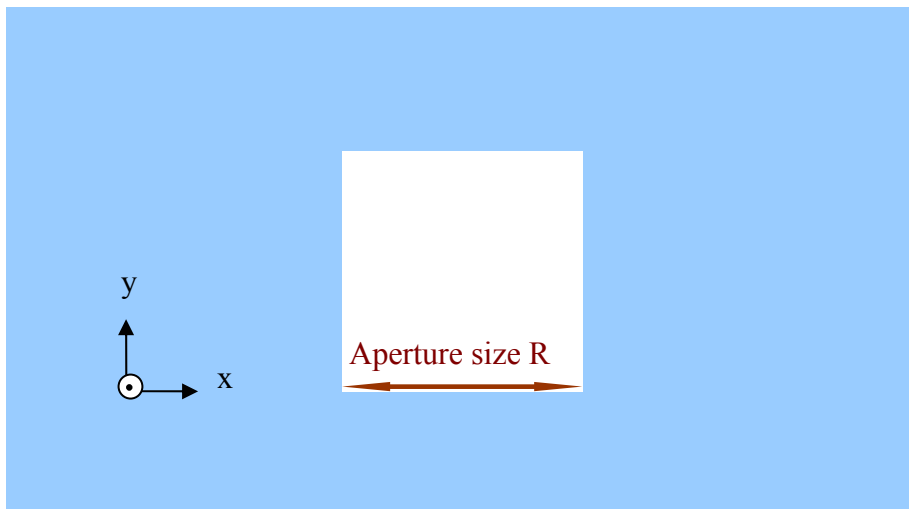**2.3.1 Bethe's small-hole theory for square apeture**

Fig.2-2 Cross-section of a square aperture

The propagation characteristics are mainly determined by the width of the aperture under the x-polarized incident light. At 633 nm wavelength incident, an evanescent wave exists within the range of aperture size interested in because of the high cutoff frequency. The rigorous electromagnetic theories must be used to consider evanescent fields and the appropriate boundary conditions.

One of the most important considerations in near-field recording is the total amount of power that can be coupled through an extremely small aperture. Although rigorous electromagnetic models are used extensively here to explore the near-zone fields of various diffracting structures, an approximation developed by Bethe [7] is used in two examples to calculate the total power transmitted through square aperture having dimensions much less than the wavelength of the imaging radiation. This approximation has been used in circular and square aperture calculations relevant to our discussion.

Bethe's model is based on the assumption that screen containing the aperture is infinitesimally thin perfectly conductor and the fields in the vicinity of a small aperture are approximately those that would exist in the absence of the aperture. Bethe's model shows when the radiation with incident angle θ incident on an aperture of size R. Bethe's model used to calculate the power transmitted through a subwavelength aperture; the total transmitted power is given by

$$P = \frac{64\pi^2}{27} \cdot (\frac{R}{\lambda})^4 \cdot (1 + \frac{1}{4}\sin^2\theta) \qquad \text{(Eq. 2-13)}$$

where θ is the angle between the normal to the screen and the direction of propagation. θ is equal to 0 for our case where the plane wave normally impinged. The power throughput is proportional to the fourth power of the diameter-to-wavelength ratio, a ratio which commonly found for light interacting with small features and is related to an intrinsic dipole-radiation source.

Although the recording medium is required to be placed in the near-field zone of the aperture to prevent diffraction from significantly enlarging the spot size, but, unavoidably, power throughput (P.T.) decays as the fourth power of the aperture size. The square aperture for our system is not ideal because of its low power throughput.

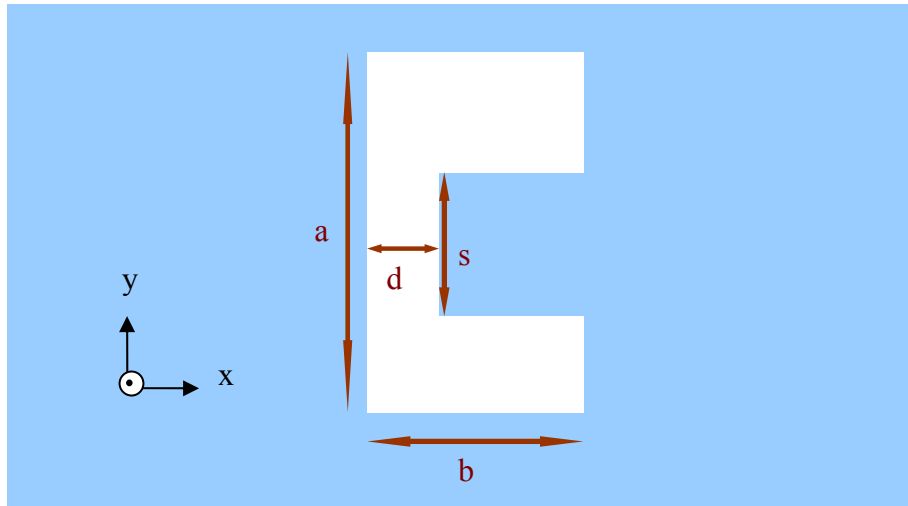**2.3.2 Propagation of light through a C-shape aperture**

Fig.2-3 Cross section of a C-shape aperture. The shaded region is metal

Shi, Thornton, and Hesselink [5] recently described enhanced transmission through a novel C-shaped aperture. The phenomenon they found are that the x-direction polarization was better than the other and this transmission peaked at a particular wavelength. These results can be best understood in slot resonators and ridge waveguides. Slot resonators are well known in the context of the microwave antennas and planar circuits. Neglecting fringing effects, the resonance occurs when the slot length equals one-half wavelength. The C-shape aperture described by Shi et al. can be viewed as a ridge waveguide and also enhances the field intensity at the center of the inner (shorter) side. If the thickness of the metal is negligible compared to wavelength, then the aperture is more accurately modeled as a ridge waveguide. The slot resonance corresponds to the cut-off frequency of the lowest-order TE_{10} mode of the waveguide.

Their simulation results show that c-shape aperture can provide 3 orders of magnitude higher near-field intensity and PT ~1000 times higher than those of a square aperture with a similar near-field spot size. These results strongly suggest that

the high transmission of the C-shape aperture is related to a propagation mode in the aperture.

2.3.3 Ridged waveguide

The square aperture acts like a cut-off rectangular waveguide for the consideration in the wavelength and spot size. The cutoff wavelength of waveguide could have been obtained by using the equivalent transverse circuit [8]-[9] as proposed as follows. If a different geometry from rectangular is needed, one must solve a transcendental equation.

The cross-section and equivalent transverse circuits of rectangular and ridged waveguide are shown in fig. 2-4.
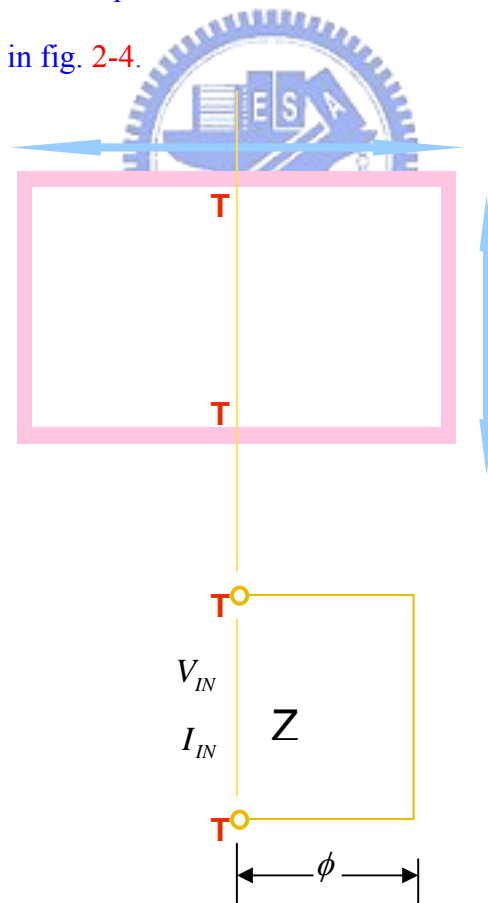

Fig. 2-4 The equivalent transverse circuit of rectangular waveguide

For rectangular waveguide, the input voltage V_{IN} and input current I_{IN} at terminal TT is given by the matrix equation:

$$\begin{bmatrix} V_{IN} \\ I_{IN} \end{bmatrix} = \begin{bmatrix} \cos\phi & jZ\sin\phi \\ \dfrac{j\sin\phi}{Z} & \cos\phi \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{(Eq. 2-14)}$$

The input impedance Z_{IN} is

$$\begin{bmatrix} \dfrac{V_{IN}}{I_{IN}} \end{bmatrix} = Z\tan\phi \qquad \text{(Eq. 2-15)}$$

At cutoff conditions Z_{IN} is infinite, a condition which is satisfied by
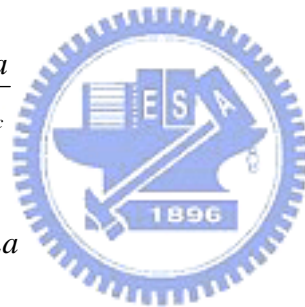
$$\varphi = \frac{\pi}{2} \qquad \text{(Eq. 2-16)}$$

and since

$$\varphi = \frac{\pi a}{\lambda_c} \qquad \text{(Eq. 2-17)}$$

by

$$\lambda_c = 2a \qquad \text{(Eq. 2-18)}$$

But for ridged waveguide, the equivalent transverse circuit of the cross-section at cut-off condition of the TE wave is shown in fig. 2.5 where C is a discontinuity capacity dependent on the height of the ridge and represents the effects of modes of higher order than the fundamental TEM mode. The higher mode is introduced by the presence of the ridge. Z_1 and Z_2 are the characteristic impedances of the parallel lines of heights (b-d) and b, respectively. φ_1 and φ_2 are the electrical lengths associated with these lines.
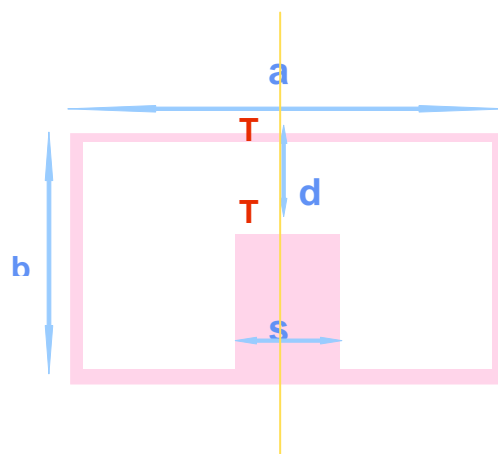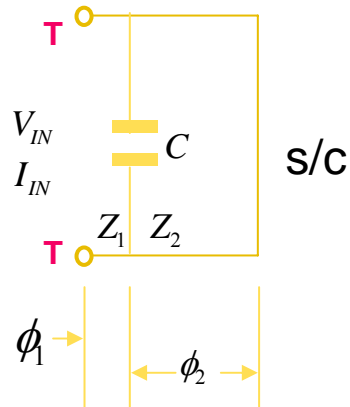
Fig. 2-5 The equivalent transverse circuit of ridged waveguide

The input voltage V_{IN} and input current I_{IN} at terminal TT is given by the matrix equation:

$$\begin{bmatrix} V_{IN} \\ I_{IN} \end{bmatrix} = \begin{bmatrix} \cos\phi_1 & jZ_1\sin\phi_1 \\ \dfrac{j\sin\phi_1}{Z_1} & \cos\phi_1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ jB & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\phi_2 & jZ_2\sin\phi_2 \\ \dfrac{j\sin\phi_2}{Z_2} & \cos\phi_2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{(Eq. 2-19)}$$

The cutoff conditions occurs when Z_{IN} is infinite or when

$$1 - \frac{Z_2}{Z_1}\tan\varphi_1\tan\varphi_2 - BZ_2\tan\varphi_2 = 0 \qquad \text{(Eq. 2-20)}$$

The above equation is a transcendental equation in λ_c. It may be solved by suitably programming a computer.

The normalized discontinuity susceptance BZ_2 for a step ratio $\alpha = d/b$ between parallel plate lines has been obtained by Marcuvitz and is accurate to one percent in the range $b < \lambda_c$. Marcuvitz's expression contains two terms: dc term and ac term. Hopfer utilizes the equivalent circuit methods as shown in Figs.2-4 and 2-5 to give design curves for the cutoff wavelength λ_c of the TE_{10} mode in single and

23

double ridged rectangular waveguide of the usual aspect ration of 0.45. But when the ridge is very narrow or the ridge is much closed to the side wall of waveguide, there will be errors due to fringing effect or approximate effects. Owing to that, there should be a correction factor taken into considered.

In microwave theory, the ridged waveguide is applied for broad-bandwidth requirement. The cut-off frequency of the lowest order TE_{10} mode is considerably lower than that of the second lowest order mode. Moreover, the electric field is confined under the ridge in the lowest order.

J. R. PYLE [8] get computed values of λ_c / a for a single rectangular ridge with b/a=0.45 are shown in Table 1. The computed value of λ_c / a for a single rectangular ridge of any aspect ratio may be expressed in terms of the value with b/a equal=0.45 by the expression
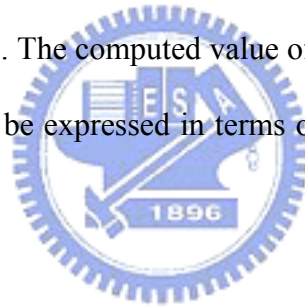
$$\left[\frac{\lambda_c}{a}\right]_{b/a} = \left[\frac{\lambda_c}{a}\right]_{0.45} + \left[\frac{b}{a} - 0.45\right] F \qquad \text{(Eq.2-21)}$$
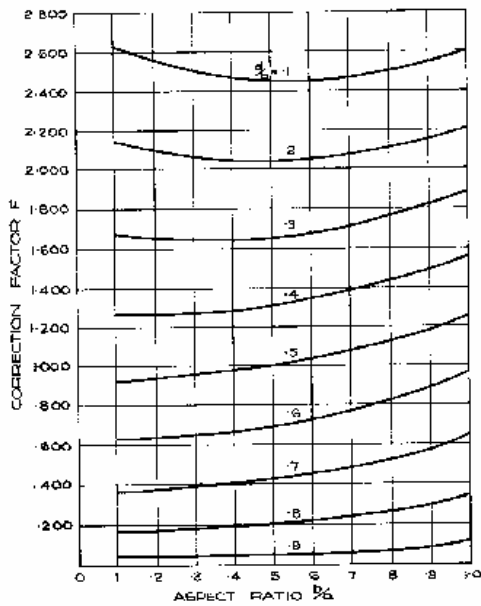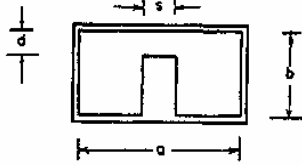
We use this formula above to design the C-shape aperture comparable to ridged waveguide, where F is the correction factor and is a function of b/a. The computed values of F against b/a are plotted in Figs.2-6 to 13 for the following ridge parameters:

s/a=0.1 (0.1) 0.9

d/h= 0.1 (0.1) 0.9
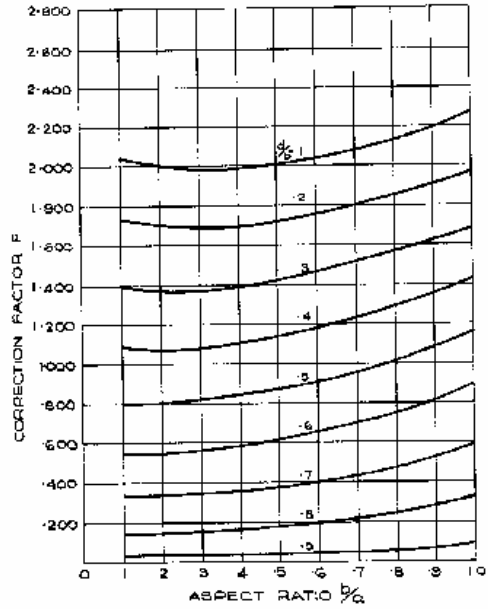
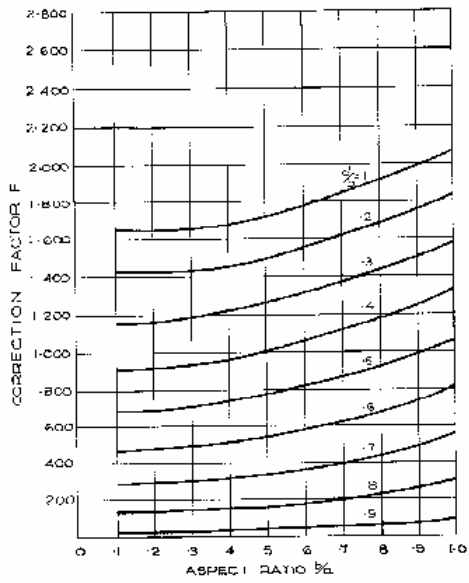Table. 2-1 Computed values of λ_c / a for c/a=0.45 single ridged waveguide

| s/a \ d/b	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.05	5.156	4.271	3.817	3.508	3.273	3.080	2.918	2.777	2.653	2.543	2.445	2.358	2.281	2.214	2.155	2.106	2.066	2.034	2.012
0.10	5.823	4.643	4.072	3.701	3.425	3.206	3.023	2.866	2.730	2.609	2.502	2.407	2.322	2.248	2.183	2.128	2.081	2.044	2.017
0.15	6.348	4.942	4.260	3.859	3.551	3.309	3.110	2.940	2.794	2.665	2.550	2.448	2.358	2.276	2.208	2.147	2.096	2.054	2.022
0.20	6.764	5.281	4.447	3.985	3.652	3.392	3.179	3.000	2.845	2.709	2.589	2.482	2.387	2.303	2.228	2.163	2.108	2.062	2.026
0.25	7.090	5.368	4.577	4.083	3.729	3.455	3.232	3.045	2.884	2.743	2.619	2.508	2.410	2.322	2.245	2.177	2.119	2.069	2.030
0.30	7.338	5.509	4.672	4.154	3.785	3.500	3.270	3.077	2.911	2.767	2.640	2.527	2.426	2.337	2.257	2.187	2.126	2.075	2.033
0.35	7.515	5.605	4.736	4.199	3.819	3.527	3.292	3.095	2.927	2.781	2.652	2.538	2.436	2.346	2.265	2.194	2.132	2.079	2.035
0.40	7.627	5.661	4.759	4.221	3.833	3.537	3.299	3.100	2.931	2.785	2.656	2.541	2.440	2.349	2.269	2.198	2.135	2.082	2.036
0.45	7.676	5.677	4.779	4.216	3.827	3.530	3.291	3.093	2.924	2.778	2.650	2.537	2.437	2.347	2.268	2.198	2.136	2.082	2.037
0.50	7.665	5.655	4.747	4.191	3.801	3.505	3.268	3.072	2.906	2.762	2.636	2.526	2.428	2.340	2.263	2.194	2.134	2.082	2.037
0.55	7.593	5.593	4.691	4.141	3.755	3.463	3.231	3.038	2.876	2.735	2.614	2.507	2.412	2.328	2.253	2.188	2.130	2.079	2.036
0.60	7.460	5.491	4.605	4.066	3.689	3.404	3.178	2.991	2.834	2.700	2.583	2.480	2.390	2.310	2.240	2.177	2.123	2.075	2.035
0.65	7.260	5.346	4.487	3.965	3.601	3.327	3.109	2.931	2.781	2.653	2.543	2.446	2.362	2.287	2.222	2.164	2.114	2.070	2.032
0.70	6.991	5.157	4.335	3.836	3.490	3.230	3.024	2.856	2.716	2.596	2.494	2.405	2.327	2.259	2.200	2.148	2.101	2.063	2.029
0.75	6.643	4.917	4.145	3.678	3.354	3.111	2.921	2.766	2.637	2.529	2.436	2.356	2.287	2.226	2.174	2.129	2.089	2.055	2.025
0.80	6.204	4.619	3.912	3.485	3.190	2.970	2.798	2.660	2.545	2.449	2.366	2.299	2.240	2.189	2.145	2.107	2.074	2.046	2.021
0.85	5.652	4.250	3.626	3.250	2.992	2.801	2.653	2.534	2.438	2.358	2.291	2.235	2.186	2.147	2.112	2.083	2.057	2.035	2.016
0.90	4.946	3.786	3.270	2.961	2.751	2.597	2.480	2.387	2.314	2.254	2.205	2.164	2.130	2.101	2.077	2.057	2.039	2.024	2.011
0.95	3.995	3.173	2.807	2.590	2.447	2.345	2.270	2.213	2.169	2.135	2.107	2.085	2.067	2.052	2.040	2.029	2.020	2.012	2.006

(a)

(b)

(c)

(d)

(e)

(f)

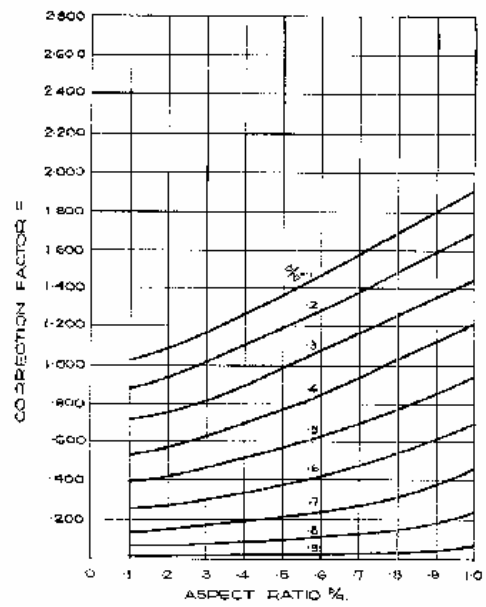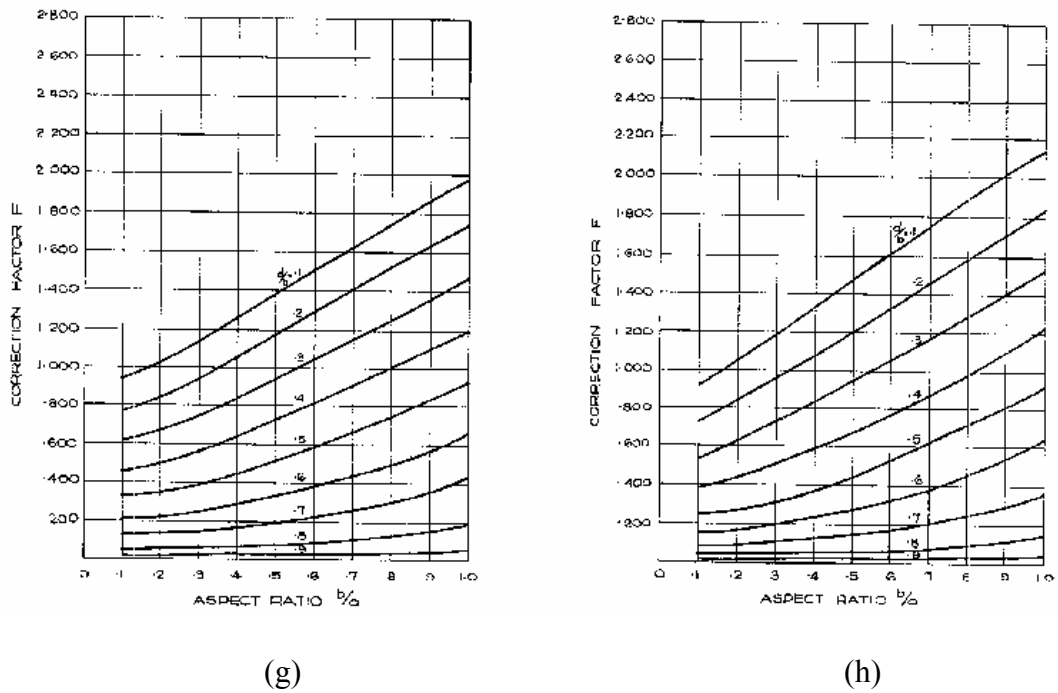| (g) | (h) |
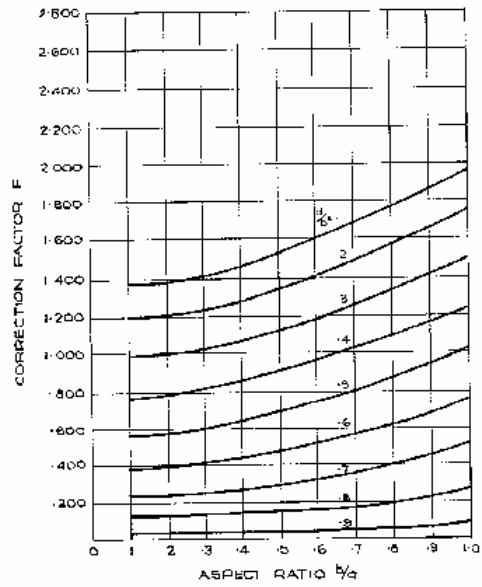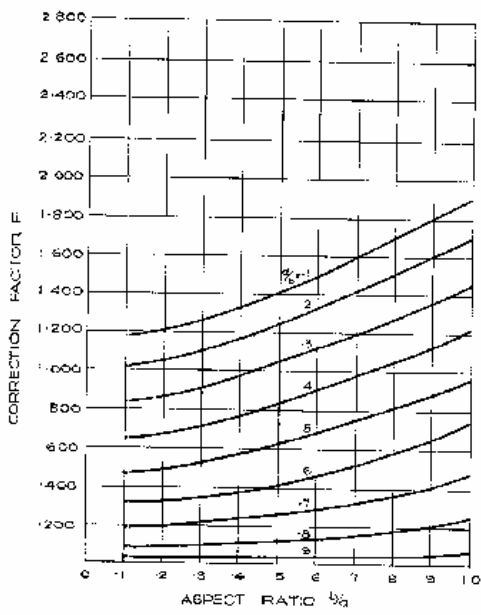
Fig. 2-6 Computed values of correction factor F against b/a at constant single width s/a, (a) s/a=0.1, (b) s/a= 0.2, (c) s/a =0.3, (d) s/a =0.4, (e) s/a=0.5, (f) s/a=0.6, (g) s/a=0.7, (h) s/a =0.8. F and $\left[\dfrac{\lambda}{a}\right]_{0.45}$ gotten from Tab. 1 are used in eq. 2-21 to get the cutoff wavelength of the different dimension of ridged waveguide.

The much larger bandwidth of ridged waveguide than that of square waveguide ensures that modes other than the lowest order mode attenuate fast. Ideally, the best coupling into the lowest order mode is obtained when the incident light closely resembles the lower order mode. As in the case of the infinitesimally thin and perfectly conducting planes, the intensity and P.T. are expected to be the largest at the cut-off wavelength at the specific mode because at cut-off the waveguide is like a resonator cavity as light does not propagate in the Z direction but keeps bouncing off the lateral walls of the waveguide. But in real case for metal taken into consideration,

the finite penetration into the metal is expected to shift cut-off wavelength to slightly higher wavelength. For the field confinement of the size of our relevance, the waveguide can be designed not to be in cut-off regime at optical frequencies.

2.4 Metal assumption

The discussions of above different shapes of aperture are based on the assumptions that the plates containing apertures are perfectly conductor with negligible thickness. An assumption for idealized perfect conductor is that conductivity is infinite. This is equivalent to saying that the electron, driven into oscillation by a harmonic wave, would simply follow the field's alternations [12]. There would be no restoring force, no natural frequencies, and no absorption.

In real case, properties of metals have to be taken into considered because of the finite conductivity and the finite thickness of metal plates. The characteristic feature of conducting media is the presence of a number of free electric charges. For metals, the charges are electrons, and their motion constitutes a current. The current per unit area resulting from the application of a field $\vec{E}$ is related to the conductivity of the medium σ by means of Eq. 2-7. For a dielectric there are no free or conduction electrons [11] and $\sigma = 0$ whereas for metals σ is nonzero and finite.

In real metals the conduction electrons undergo collisions with the thermally agitated lattice or with imperfections and in so doing irreversibly convert electromagnetic energy into joule heat. The absorption of radiant energy by a material is a function of its conductivity.

2.4.1 Penetration depth

The electromagnetic fields in any material can be obtained by the Maxwell's equations under appropriate boundary conditions. The Maxwell's equations are valid even in the quantum mechanical region. The metal under a strong laser field incident is visualized as a continuous medium working fairly well with all the detailed microscopic physics lumped in the constitutive material parameters: permeability, permittivity and conductivity. If we visualize the medium as being continuous for a particular temporal, spatial frequency, the materials are assumed as to be non-magnetic as well as external charge and current are absent [4], Maxwell's equations lead to

$$- j\vec{k} \cdot \vec{E} = 0 \qquad \text{(Eq. 2-22)}$$

$$- j\vec{k} \times \vec{E} = - j\omega\mu_0\vec{H} \qquad \text{(Eq. 2-23)}$$

$$- j\vec{\kappa} \cdot \vec{H} = 0 \qquad \text{(Eq. 2-24)}$$

$$- j\vec{k} \times \vec{H} = -\sigma\vec{E} + j\omega\varepsilon\vec{E} \qquad \text{(Eq. 2-25)}$$

and

$$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} = \sigma\vec{E} + j\omega\vec{D} = \sigma\vec{E} + j\omega\varepsilon\vec{E} = j\omega\varepsilon(1 + \frac{\sigma}{j\omega\varepsilon}) \qquad \text{(Eq. 2-26)}$$

$$\varepsilon_c = \varepsilon(1 + \frac{\sigma}{j\omega\varepsilon}) \qquad \text{(Eq. 2-27)}$$

$$\gamma \equiv jk_c = j\omega\sqrt{\mu\varepsilon_c} = j\omega\sqrt{\mu\varepsilon(1 + \frac{\sigma}{j\omega\varepsilon})} = \alpha + j\beta \qquad \text{(Eq. 2-28)}$$

$$e^{-(\alpha + j\beta)Z} = e^{\alpha Z} * e^{-j\beta Z} \qquad \text{(Eq. 2-29)}$$

$$\vec{E}_c = \hat{a}_x E_0 e^{-\alpha Z} * e^{-j\beta Z} = \hat{a}_x E_0 * e^{-\alpha Z} * \cos(\omega t - \beta Z) \qquad \text{(Eq. 2-30)}$$

$$1 = \alpha\delta \qquad \delta = \frac{1}{\alpha} \qquad\qquad \text{(Eq. 2-31)}$$

where α is called the absorption coefficient or the attenuation coefficient. If σ is not zero, ε_c will be complex and k will has a complex magnitude. The complex magnitude will drop by a factor of $e^{-1} = 1/2.7 \approx \frac{1}{3}$ after the wave has propagated a distance $y = 1/\alpha$, known as the skin depth or penetration depth. The energy is absorbed by the medium and results in the heating of the medium through different microscopic scattering process.
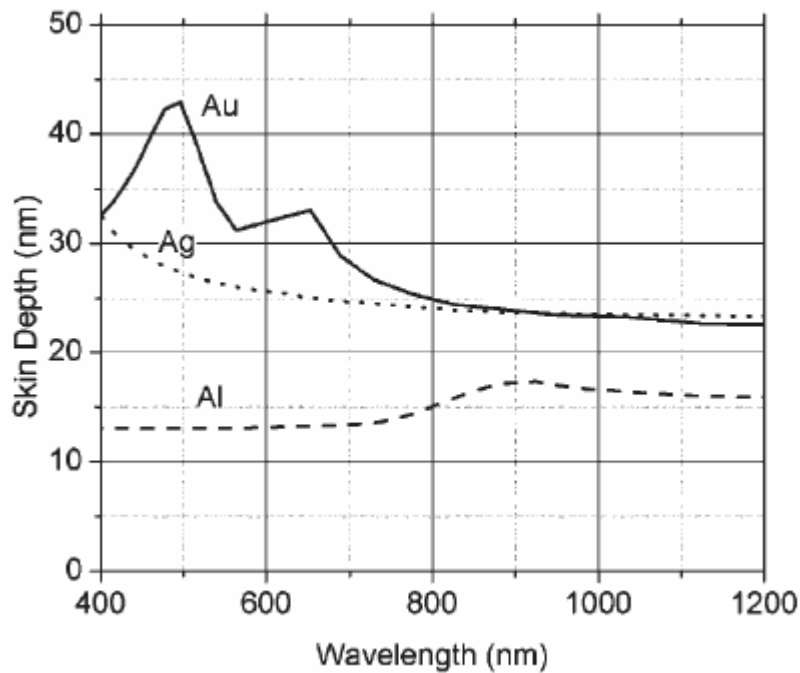
Fig.2-7 Penetration depth for silver, gold, and aluminum in the visible

computed from the optical constants in ref. 13. (Skin depth

$\delta \approx 25nm$ for Ag, $\delta \approx 14nm$ for Al, and $\delta \approx 34nm$ for Au)

2.4.2 Dispersion relation

As mentioned before, the materials are assumed as to be non-magnetic and

external charge and current are absent. The dielectric function is also assumed to be isotropic. For simplicity, we limit ourselves to media with the magnetic permeability is 1 and dielectric function independent of the wave vector in the absence of spatial dispersion effects. This is true for many materials. Substituting equations 2-23 and 2-25, the dispersion relation for transverse modes is given by [14]

$$\omega^2 = \frac{\vec{k} \cdot \vec{k}}{\mu_0 \varepsilon_c}$$
(Eq. 2-32)

In the microwave, lower frequencies domain of the infrared, a metal behaves like an ideal conductor. The penetration depth of electric and magnetic field into metals is negligible and current associated with metals is superficial. However, at optical frequencies, the losses in metals are considerably greater and the penetration depth of the fields can become comparable to the device dimensions in near-field optics.For a material to be transparent, the penetration depth must be large in comparison to its thickness.

The presentation of metal as a continuous medium works fairly well in the low-frequency, long-wavelength domain of the infrared. We might expect that as the wavelength of the incident beam decreased, the actual granular nature of matter would have to be reckoned with. Indeed the assumption fails at higher (non-optical) frequencies and plasma oscillations are possible. Indeed, the continuous model shows large discrepancies from experimental results at optical frequencies. And so we turn to the classical atomistic picture initially formulated by Drude model [10].

2.4.3 Dielectric function of metals

The simplest model for the permittivity of a metal is Drude model. In a real

metal, the electrons can undergo interband transition near the Fermi energy by photon absorption. This gives rise to additional features in the energy dependence of the permittivity [14]. The dielectric constant of a metal is given by the sum of a contribution from the free electrons of the metal and a contribution from the interband transitions. The contribution from the free electrons ε_f is given by the Drude free-electron model.

This model is similar to the Lorentz harmonic-oscillator model, but the spring constant is zero. The equation of motion of a free electron is [12]

$$m_e \frac{\partial^2 \vec{r}}{\partial t^2} + m_e \Gamma \frac{\partial \vec{r}}{\partial t} = e\vec{E}_0 e^{-j\omega t} \qquad \text{(Eq. 2-33)}$$

where m_e is the effective mass of the electron, e is the charge of the electron, $\vec{E} = \vec{E}_0 e^{-i\omega t}$ is the external electric field, and Γ is the damping constant, and $\vec{r}$ is the displacement of the electron from its mean position. The dipole moment is given by $\vec{p} = e\vec{r}$. Denoting the number of free electrons per unit volume by n, the polarization is $\vec{P} = n\vec{p}$. If we assume that the medium is isotropic ($\vec{P} // \vec{E}$), the dielectric function $\varepsilon_f(\omega)$ is

$$\varepsilon_f(\omega) = 1 + \frac{\left|\vec{P}\right|}{\varepsilon_0 \left|\vec{E}\right|} \qquad \text{(Eq. 2-34)}$$

Where ε_0 is the dielectric constant of a vacuum. From the equations described above,

$$\varepsilon_f(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\Gamma\omega} = 1 - \frac{\omega_p^2}{\omega^2 + \Gamma^2} + i\frac{\omega_p^2 \Gamma}{\omega(\omega^2 + \Gamma^2)} \qquad \text{(Eq. 2-35)}$$

where ω_p is the bulk plasmon frequency,

$$\omega_p = \left(\frac{ne^2}{\varepsilon_0 m_e} \right)^2 \qquad \text{(Eq. 2-36)}$$

$$\varepsilon(\omega) = \varepsilon_f(\omega) + \varepsilon_{ib}(\omega) \qquad \text{(Eq. 2-37)}$$

The simplest situation occurs in simple metals, in which the electron system in its properties is similar to a homogeneous electron gas of the same density. Among the simplest metals, Aluminum is one of the most interesting materials from the point of view of electron theory. This is due to the success of the nearly free-electron (NFE) model in interpreting Fermi-surface experiments. Since the band structure of Aluminum is expected to be described by NFE model, one would hope to in a position to account for its optical properties. A peak in σ near 1.6 eV is related to interband transitions [14]. Near the plasma peak there is a rather sharp drop in reflectivity.

2.5 Surface modes

After discussing the losses arise from the complex propagation constant, transverse magnetic surface modes which are known to exist at the interface of a metal and dielectric will be discussed. These modes are called surface plasmons or surface plasmon polarizations. Plasmon oscillations, i.e., the collective motions of an electron gas, are an inherent property of solid-state plasmas. A nonzero local charge density in a metal or semiconductor, whose conduction electrons from an electron gas, creates long-range Coulomb fields which organize the system into collective motion. The condition $\nabla \cdot \vec{E} \neq 0$ in the bulk material, where $\vec{E}$ is the electric field in the bulk, gives rise to one set of plasmon oscillations, commonly known as the bulk plasmon modes, which are described by Eq. 2-35. The bulk plasmon modes exist alone in the idealized situation where the system is unbounded by surfaces.

The presence of surfaces complicates the situation. Not only are the bulk plasmon modes modified but new modes called surface modes are created, which are absent in the bulk. These surface modes have relations, differing from the bulk plasmon modes and varying according to the characteristics of the surfaces. As with the bulk mode, the surface modes can be exited by incident electrons or photons.

A complex dielectric constant of metals obeys a plasmon dispersion law. At optical wavelength, the real part of the dielectric constant of metals is negative and has therefore a surface active character which leads at certain frequencies to plasmon excitations (resonant oscillations of the electron gas). These resonances depend on the material, frequency, and geometry. Surface modes, such as surface plasmon modes, are localized at the interfaces between different media. In our case, there are surface plasmon modes existed at the surface of the film and that of the sidewall of metallic aperture. The field distribution can be derived by a classical analysis of a multilayer system with the Fresnel formula and appropriate boundary conditions. The wavenumber of the surface plasmon wave is given by [11]

$$k_{sp}(\omega) = \frac{\omega}{c} \sqrt{\frac{\varepsilon_m(\omega)\varepsilon_s}{\varepsilon_m(\omega) + \varepsilon_s}}, \qquad\qquad \text{(Eq. 2-38)}$$

where ω is the angular frequency of the surface plasmon wave, $\varepsilon_m(\omega)$ is the complex dielectric constant of the metal, and ε_s is the dielectric constant of the surrounding medium. Figure 2-15 is a plot of the $\omega - k$ dispersion relationship. Since the surface plasmon polarization is associated with an evanescent field, it can be excited by irradiation with an evanescent light wave satisfying the dispersion relationship at the interface.

Because of the complexity of the propagation through the metallic aperture, here a simple picture of the surface plasmon polarization excited at sidewall of aperture is introduced. Diffraction is happened to the light during the transmission through the sub-wavelength aperture. Therefore the incident waves are departed into several directions to propagate, and the propagation with different wavenumber in the same direction induces the surface plasmon when the k_z is equivalen to k_{sp}. It is well known that the higher excitation of the surface plasmon occurred when the absolute value of the real part of the dielectric constant is much greater than the imaginary part [14], [15]. Therefore, the metals we choose should conform to the situation to excite higher surface plasmon. Under our conditions, the material should be selected to satisfy higher surface plasmon on the sidewall of apertures and shorter penetration depth to avoid too much loss converted into joule heat to induce some unpredictable defects.

Fig. 2-8 SPR dispersion curve

2.6 simulation method for subwavelength aperture

When a light field interacts with structures that have complex geometric features comparable in size to the wavelength of the light, it is not permissible to invoke the assumptions of the classical diffraction theory, which simplify the problem and allow for approximate solutions. For such cases, direct numerical solutions of the governing equations are sought through approximating the continuous time and space derivatives by the appropriate difference operators. The Finite Difference Time Domain (FDTD) method discretizes Maxwell's equations by using a central difference operator in both the time and space variables. The $\vec{E}$ and $\vec{B}$ fields are then represented by their discrete values on the spatial grid, and are advanced in time in steps of Δt. The numerical solution thus obtained from Maxwell's equations (in conjunction with the relevant constitutive relations) provides a highly reliable representation of the electromagnetic field distribution in the space-time region under consideration. The FDTD method is a rigorous solution of Maxwell's equations and does not have any approximation or theoretical restrictions. FDTD simulation method will be introduced. Here a commercial FDTD package (RSOFT FullWAVE v.3.0.1) on a 3GHz Pentium　 is used.

2.6.1 The FDTD method

The spatial unit cell used in three-dimensional (3D) FDTD simulations is shown in Fig.2-9. The components of the vector fields $\vec{E}$ and $\vec{H}$ are located at different position with respect to the cell center, so that every component of the electric fields

surrounded by four circulating components of the magnetic fields, and vice versa [17].

Such a staggered mesh is motivated by the integral form of Maxwell's curl equations.

The contour integrals of $\vec{E}$ ($\vec{H}$) along the edges of the cell in Faraday's (Ampere's

law) circulate around the corresponding magnetic (electric) field component at the

center of the cell face.

In 3D simulations at least six field components must be stored and updated at

each grid point, which leads to considerable memory and CPU requirements for

FDTD simulations. In this method, an incident electromagnetic wave at a particular

optical frequency and all the linear materials are assumed. All the excitations induced

in the materials namely, charges, currents and electromagnetic fields are all at the

same frequency as the incident field. The other assumption is that the fields resulting

from the material excitations do not induce other oscillating charges in the source of

the incident field that would effectively alter the incident field.
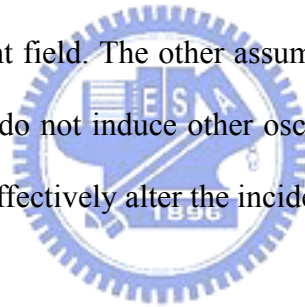
2.6.2   The YEE Algorithm

Yee algorithm provides a beautifully simple picture of three-dimensional space

filled by an interlinked array of Faraday's Law and Ampere's Law contours.

2.6.2.1 Central difference accuracy

It is well known that the highest accuracy occurs when the finite difference is

centered as

$$\frac{df}{dz}\bigg|_{z=z_0} = \frac{f(z_0 + \Delta z/2) - f(z_0 - \Delta z/2)}{\Delta z} + O(\Delta z^2) \qquad \text{(Eq. 2-39)}$$

resulted in second-order accuracy. Thus, discretization errors are reduced by a factor of 4 when the mesh size is halved. Here the central difference accuracy is used in YEE algorithm.

2.6.2.2 FDTD spatial discretization

The axes of our Cartesian coordinate axes are assumed to be parallel to the edges of the objective space meshed up into a regular Cartesian grid. The spacing along each axis could be different. As illustrated in Fig 1, the Yee algorithm also centers its $\vec{E}$ and $\vec{H}$ components in time in a leapfrog arrangement. All of the $\vec{E}$ components in the model space are completed and stored in memory using the E data just computed. The cycle repeats with the computation of the $\vec{E}$ components based on the newly obtained $\vec{H}$. This process continues until time-stepping is concluded.
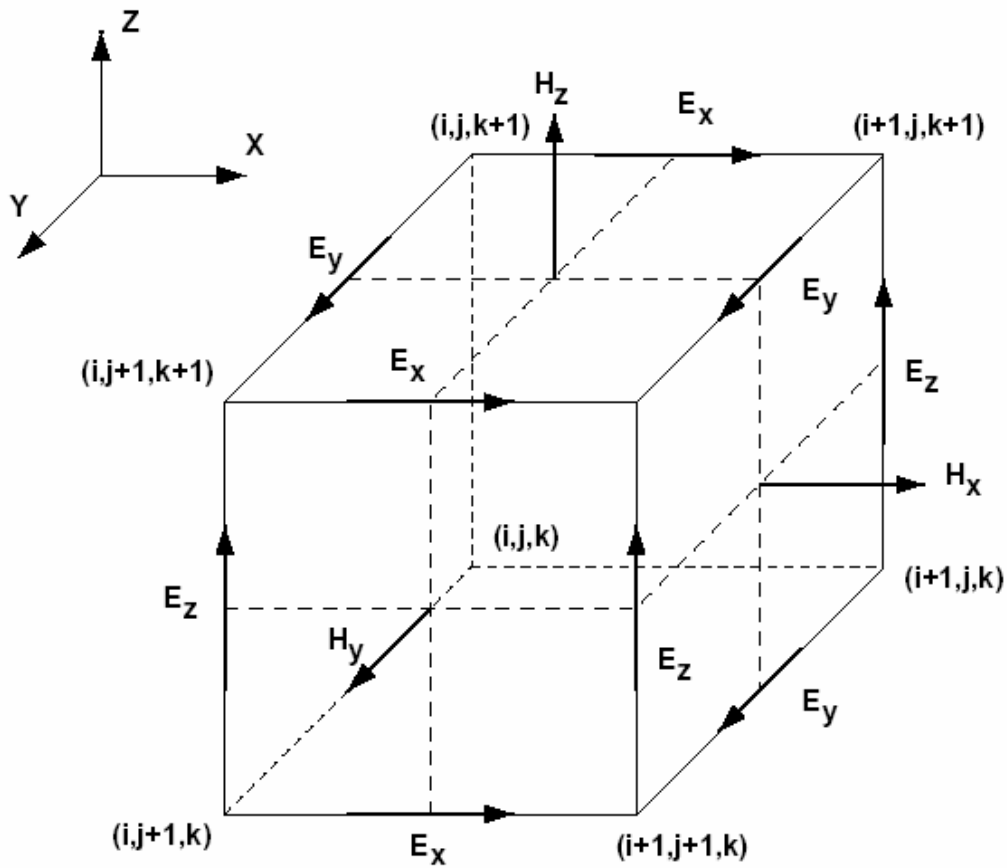
Fig. 2-9 Kane Yee, IEEE Trans. Antennas and Propagation, May 1966

From Ampere's Law and Faraday's Law

$$\frac{\partial \vec{H}}{\partial t} = -\frac{1}{\mu}\nabla \times \vec{E}, \qquad \frac{\partial \vec{E}}{\partial t} = -\frac{1}{\varepsilon}\nabla \times \vec{H} - \frac{\sigma}{\varepsilon}\vec{E} \qquad \text{(Eq. 2-40)}$$

these equations can be expanded into an equivalent system of six couple scalar

equations for all electric and magnetic field components,

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu}\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right), \qquad \frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - \sigma E_x\right) \quad \text{(Eq. 2-41)}$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu}\left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right), \qquad \frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - \sigma E_y\right) \quad \text{(Eq. 2-42)}$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu}\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right), \qquad \frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - \sigma E_z\right) \qquad \text{(Eq. 2-43)}$$

Yee's basic scheme is to use a series of finite-difference approximations to transform Eqs. (2-41) to (2-42) into a system of six algebraic equations. These equations can then be solved with a simple leap-frog algorithm. For illustration we provide here the FDTD approximation for the E_x field component,

$$\frac{\vec{E}^n - \vec{E}^{n-1}}{\Delta t} = -\frac{\sigma}{\varepsilon}\frac{\vec{E}^{n-1} + \vec{E}^n}{2} + \frac{1}{\varepsilon}\nabla \times \vec{H}^{n-\frac{1}{2}} \qquad \text{(Eq. 2-44)}$$

$$E_x\big|_{i,j,k}^{n+1} = \frac{1 - \dfrac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}}{1 + \dfrac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}} E_x\big|_{i,j,k}^{n}$$

$$+ \frac{\Delta t/\varepsilon_{i,j,k}}{1 + \dfrac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}}\frac{1}{\Delta y}\left(H_z\big|_{i,j+1/2,k}^{n+1/2} - H_z\big|_{i,j-1/2,k}^{n+1/2}\right) \qquad \text{(Eq. 2-45)}$$

$$- \frac{\Delta t/\varepsilon_{i,j,k}}{1 + \dfrac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}}\frac{1}{\Delta z}\left(H_y\big|_{i,j,k+1/2}^{n+1/2} - H_y\big|_{i,j,k-1/2}^{n+1/2}\right)$$

where i, j and k indices represent the x, y, and z spatial locations, respectively. n indicates the current time index; Δy and Δz represent the y and z increments, respectively; and Δt is the time step.

In one-dimensional FDTD code, $c\Delta t$ Δx is required for stability. This constrain is known as the Courant stability condition and is easy to interpret physically. When the Courant stability condition is satisfied, the FDTD grid is casually connected; the speed of light bounds the rate at which information can be transmitted across the mesh.

In higher dimensions, say three, the stability criterion becomes

$$c\Delta t \leq \frac{1}{(1/\Delta x^2 + 1/\Delta y^2 + 1/\Delta z^2)^{1/2}} \qquad \text{(Eq. 2-46)}$$

The time step must be reduced in size accordingly to ensure stability on this smaller effective mesh.

2.6.3 Materials and material interfaces

Since all the materials are non-magnetic, $\bar{B} = \mu_0 \bar{H}$. Here μ_0 is the permeability of free space. On the other hand, a constitutive relation has to be defined between $\bar{D}$ and $\bar{E}$ depending on the material. For the linear materials in our assumption, in the frequency space [13],

$$\bar{D}(\omega) = \varepsilon(\omega)\bar{E}(\omega) \qquad \text{(Eq. 2-8)}$$

the velocity of propagation wave varies with its sinusoidal frequency at low intensities of the wave's $\bar{E}$ and $\bar{H}$. This phenomenon is caused by a variability of the material's dielectric permittivity or magnetic permeability with frequency. As with all FDTD techniques, these approaches rigorously enforce the vector field boundary conditions at interfaces of different media at the time scale of a small fraction of the impinging pulse width or carrier period. In order to model such materials, we need to include a dispersion relation for permittivity that has the damping into it. The Drude relaxation and the Lorentzian resonance are used for dispersion models [17].

Drude

$$\varepsilon(\omega) = \varepsilon_\infty + \frac{\omega_p^2}{2j\omega v_c - \omega^2} = \varepsilon_\infty + \chi(\omega) \qquad \text{(Eq. 2-47)}$$

The Drude model does not accurately describe the real and imaginary indices of refraction of noble metals over a wide frequency range. Also, the Fourier transform of

41

$\chi(\omega)$ is often not numerically stable for the Drude model. In order to avoid some problems with the Drude model, the Lorentzian model is utilized in some cases.

Lorentz

This model is similar to the Drude model, except the nonzero spring constant.

$$\varepsilon(\omega) = \varepsilon_\infty + \frac{(\varepsilon_s - \varepsilon_\infty)\omega_p{}^2}{\omega_p{}^2 - 2j\omega\delta_p - \omega^2} \qquad \text{(Eq. 2-48)}$$

At the interface between materials, the tangential components of $\vec{E}$ and $\vec{H}$ are continuous. On the other hand, while the normal component of $\vec{H}$ is continuous (since all materials are non-magnetic), the normal component of $\vec{E}$ is not. Thus, we choose the interface to correspond to a plane in the E grid. This ensures that there is no normal E edge intersecting the interface. The normal $\vec{E}$ edges are on either sides of the plane. Due to the difference in the update equations for $\vec{E}$ from $\vec{D}$ on either side of the plane, the discontinuity is established. This also accounts for the charges induced at the interface. The tangential $\vec{E}$ components are supposed to be shared between the two regions. Thus, the complex permittivity used in the update of $\vec{E}$ from $\vec{D}$ is set to be the mean of the permittivity of the materials on the either side of the interface. Thus, in the case of interface between Lorentz or Drude materials, the interface edges have to be modeled as belonging to a new material. If one of the materials is a perfect electrical conductor, we set the interface edges also to be perfect electrical conductor. This approach can be generalized to the case where there are corners with two or more materials.

2.6.4 Boundary conditions

One of the greatest challenges of the FDTD method is to get the efficient and accurate solution of electromagnetic wave interaction problems in unbounded regions. For such problems, an absorbing boundary condition (ABC) must be introduced at the outer lattice boundary to simulate the extension of the lattice to infinity [19]. A number of analytical techniques have been introduced to achieve this goal. An alternate approach to realize an ABC is to terminate the outer boundary of the space lattice in an absorbing material medium. This is analogous to the physical treatment of the walls of an anechoic chamber. Ideally, the absorbing medium is only a few lattice cell thick, reflectionless to all impinging waves over their full frequency spectrum, highly absorbing, and effective in the near field of a source or a scatterer. Berenger's introduction of a highly effective absorbing-material ABC designated the perfectly matched layer, or PML. The innovation of Berenger's PML is that plane waves of arbitrary incidence, polarization, and frequency are matched at the boundary. The idea is that the surface impedance of the surface adjacent to the problem space is matched to the problem space to prevent reflection. Moreover, loss is introduced in the layer so that the fields decay to negligible value till the outer surface of the boundary layer is reached. Berenger derived a novel split-field formulation of Maxwell's equations where each vector field component is split into two orthogonal components of first-order partial differential equations. By choosing loss parameters consistent with a dispersionless medium, a perfectly matched planar interface is derived. The outer surface, for convenience, is assumed to be a perfect electrical conductor. The medium in the boundary layer is hypothetical in the sense that standard Maxwell's equations do not apply. A modified set of equations is used for fields update. For the special case of normal incidence ($\theta=0$) of the impinging wave, we have

$$\Gamma = \frac{\eta_1 - \eta_2}{\eta_1 + \eta_2} \qquad \text{(Eq. 2-49)}$$

where the wave impedances in Regions 1 and 2 are given by

$$\eta_1 = \sqrt{\frac{\mu_1}{\varepsilon_1}} \; ; \qquad \eta_2 = \sqrt{\frac{\mu_2(1+\sigma^*/j\omega\mu_2)}{\varepsilon_2(1+\sigma/j\omega\varepsilon_2)}} \qquad \text{(Eq. 2-50)}$$

then if we set $\varepsilon_1 = \varepsilon_2$ and $\mu_1 = \mu_2$ and further enforce the condition

$$\sigma^*/\mu_1 = \sigma/\varepsilon_1 \qquad \longrightarrow \qquad \sigma^* = \sigma\mu_1/\varepsilon_1 = \sigma(\eta_1)^2 \qquad \text{(Eq. 2-51)}$$

then $k_1 = k_2$ and $\eta_1 = \eta_2$. This yields Γ=0, a reflectionless Region-1 /Region-2 interface for the normally impinging wave.

For a normal-incidence case, the transmitted wave in lossy medium attenuates exponentially along the normal direction. A major difficulty is that this medium is matched to the interior of the lattice only for normally incident waves. Therefore the absorber must be placed sufficiently far from any source such that impinging waves are nearly plannar and normally incident. This is not practical for large or elongated calculation domains. Berenger found it possible to quickly attenuate waves entering the PML region by simultaneously matching the wave impedance at any angle of incidence (i.e., nearly zero reflectance). Conceptually, PML boundary conditions can be thought of as a numerical implementation of an anechoic chamber. That reflection coefficient below 10^{-6} can be achieved by applying 12-16 PML layers.

Chapter 3

Fabrication and measurement equipment

3.1 Nano-aperture fabrication process

To verify the simulation results, fabrications and experiments are carried out.
Here small apertures on Al film with 80nm thickness are fabricated by focused ion
beam (Micrion 5000 FIB) which will be introduced below.

3.1.1 Focused ion beam (FIB)
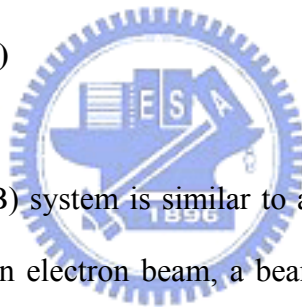
A focused ion beam (FIB) system is similar to a scanning electron microscope,
except that instead of using an electron beam, a beam of ions is scanned across the
sample. The ion beam is ejected from a liquid metal ion source (usually Ga), with a
spot size on modern systems of <10nm. Imaging using secondary electrons at low
beam current provides surface information with similar resolution to that obtainable
from an SEM; however the main applications arise from the use of ions as the
scanned species. These include compositional imaging via secondary ions, direct
etching of material in selected regions for in-situ sectioning and imaging,
microfabrication, transmission electron microscopy specimen preparation, and
localized deposition and implantation of metal and insulator structures.

A FIB system is shown in Fig. 1. By applying a voltage different both the
extractor and the suppressor, gallium ions are extracted from the liquid metal ion

source (LMIS). Lens1 turns the beam into a parallel current. The automatic variable aperture (AVA) then can be set for the desired current to the sample. The deflector plate gives the option to blank the beam so no ions can hit the sample. Finally, lens 2 focuses the beam to the substrate. A pattern can be drawn by using CAD facilities available in the software.
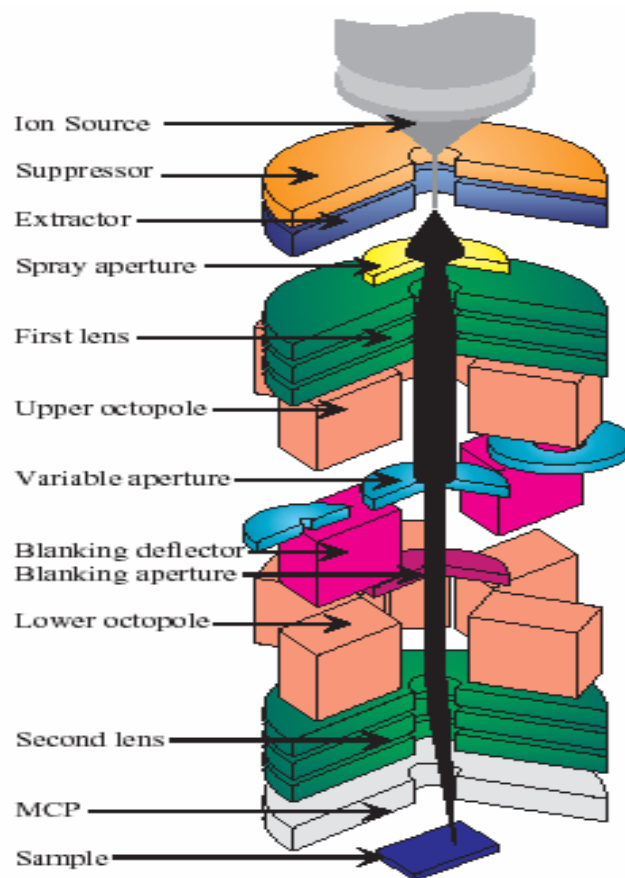
Fig. 3-1 A schematic view of FIB system

At low primary beam currents, very little material is sputtered; modern FIB systems, such as our Micrion systems, can achieve 5 nm imaging resolution. At higher primary currents, a great deal of material can be removed by sputtering, allowing precision milling of the specimen down to a sub micron scale.

In addition to primary ion beam sputtering, the system permits local "flooding" of the specimen with a variety of gases. These gases can either interact with the primary gallium beam to provide selective gas assisted chemical etching or selective deposition of either conductive or insulating material by decomposition of the deposition gas by the primary ion beam.

Until recently, the overwhelming usage of FIB has been in the semiconductor industry. Such applications as defect analysis, circuit modification, mask repair and transmission electron microscope sample preparation of site specific locations on integrated circuits have become commonplace procedures. The latest FIB systems, such as our Micrion systems, have high resolution imaging capability; this capability coupled with in situ sectioning has eliminated the need, in many cases, to examine FIB sectioned specimens in the SEM.
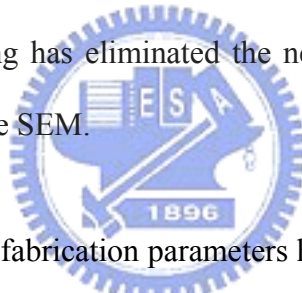
Here are some important fabrication parameters have to be chosen about our FIB system in CARNEGIE MELLON UNIVERSITY:

1. Acceleration voltage of the ion beam
2. Aperture size
3. Dose

3.2 Measurement system

After the fabrication of the component, the inspection will be performed to make sure that the performances agree with the simulation results. Here, NSOM will be introduced.

3.2.1 The Near-field Scanning Optical Microscope (NSOM)

Near-field scanning optical microscopy (NSOM) has been extensively studied owing to the capabilities for optical imaging and spectroscopy with spatial resolution beyond the diffraction limit of the light, for subwavelength photolithography, and next generation optical data storage using the near-field light. NSOM is mostly based on the cantilevered probe where the field of light has to be confined and controlled at a subwavelength size aperture. It is a powerful tool for the study of nanometer features with spatial resolution of 50~100nm [1].

The NSOM instrument must contain a light source, collection optics which is an efficient far-field optical design for the low light signal levels, detector, feedback mechanisms. Laser light is focused onto the sample by objective lens. The signal is collected by fiber tip and transferred into a photomultiplier (PMT) by a far-field microscope objective and a high NA lens, a charge-coupled device (CCD), or a photodiode. Here the feedback system and modes of operation will be demonstrated.

3.2.2 Modes of operation

There are four possible modes of operation with NSOM [2] :

(a)   (b)   (c)   (d)
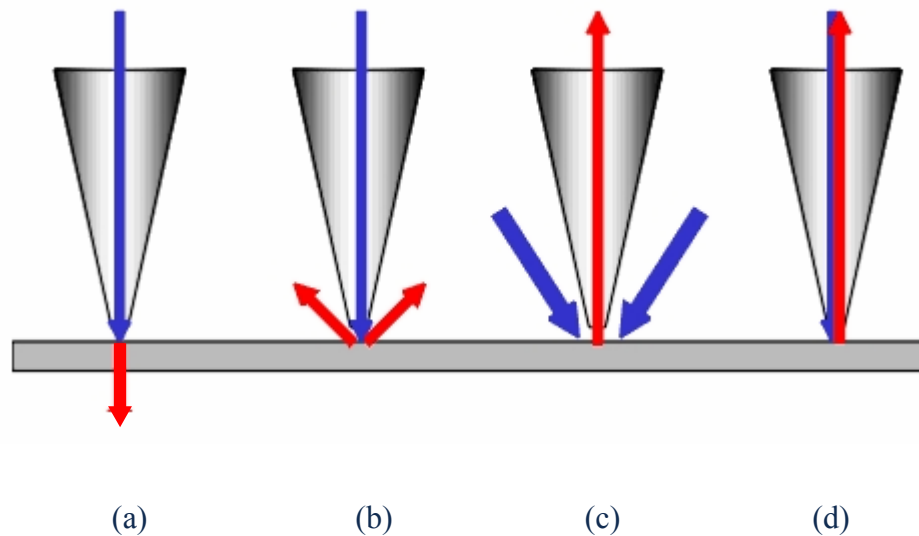
Fig.3-2 Modes of operation, (a) Transmission, (b) Reflection, (c) collection , and

(d) Illumination/collection mode imaging

- Transmission mode imaging. The sample is illuminated through the probe, and the light passing through the sample is collected and detected.

- Reflection mode imaging. The sample is illuminated through the probe, and the light reflected from the sample surface is collected and detected.

- Collection mode imaging. The sample is illuminated with a macroscopic light source from the top or bottom, and the probe is used to collect the light from the sample surface.

- Illumination/collection mode imaging. The probe is used for both the illumination of the sample and for the collection of the reflected signal.

Detecting the collected light can be achieved with a wide variety of instruments: an Avalanche Photo Diode (APD), a Photomultiplier Tube (PMT), a CCD, or a spectrometer. The signals obtained by these detectors are then used to create an NSOM image of the surface.

Here the mode of operation with NSOM used is collection mode imaging and the schematic view of our NSOM setup is shown in Fig. 3-3.
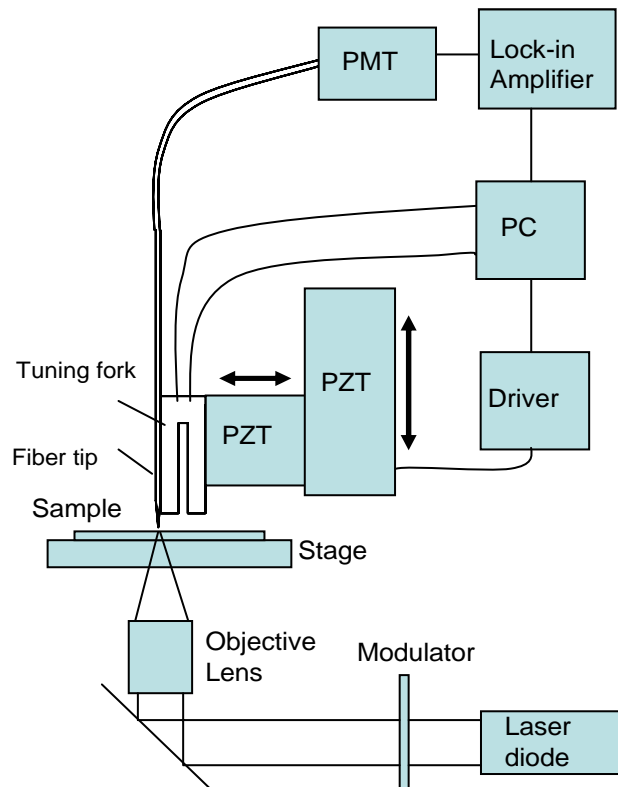
Fig. 3-3 Schematic view of NSOM setup

3.2.3 Feedback

In a NSOM experiment, the fiber tip must be placed closed to the sample surface of interest because that the behavior of electromagnetic near-field of a subwavelength aperture shows it to be strongly diverging, except very near the aperture.

The aperture-sample distance is normally kept constant by utilizing shear-force feedback which includes alternative optical and nonoptical methods or optical deflection techniques. Here alternative optical method of shear-force feedback based

on piezoelectric detection of tip-sample interaction is used in our NSOM system to avoid the low-contrast resulting from alternative optical method. The probe is mounted on an AFM noncontact stage on which the cantilever can be vibrated by a piezo element. The near-field signal modulated by the vibration signal is lock-in amplified using the vibration signal applied on the piezo element as a reference.

Chapter 4

Simulations and discussions

4.1 Introduction

Based on our system and the principle in chapter 2, we begin to establish a simulation model used to characterize the features of the field distribution below the subwavelegth aperture, such as P.T., spot size and intensity. Because of the complicated electromagnetic interaction of the nano-scale aperture, the simulation was departed into three parts to obtain the results more correctly and efficiently. First, ray tracing method was used to determine the proper parameters of the fiber lens and the distance between fiber lens and subwavelength aperture for longer working distance and smaller spot size. Then the output field from the fiber lens was caculated using Beam Propagation Method and was used as launched field into FDTD to simulate the field distribution below the subwavelength aperture. The flow chart of our simulation was shown below.

2. Beam propagation Method (BPM) ----get the corrective

632.8 μm

3. Output the field will be launched into the subwavelength component for FDTD calculation

$200\,\mu m$ $500\,\mu m$

1. The proper parameters are decided by OSLO by the group member
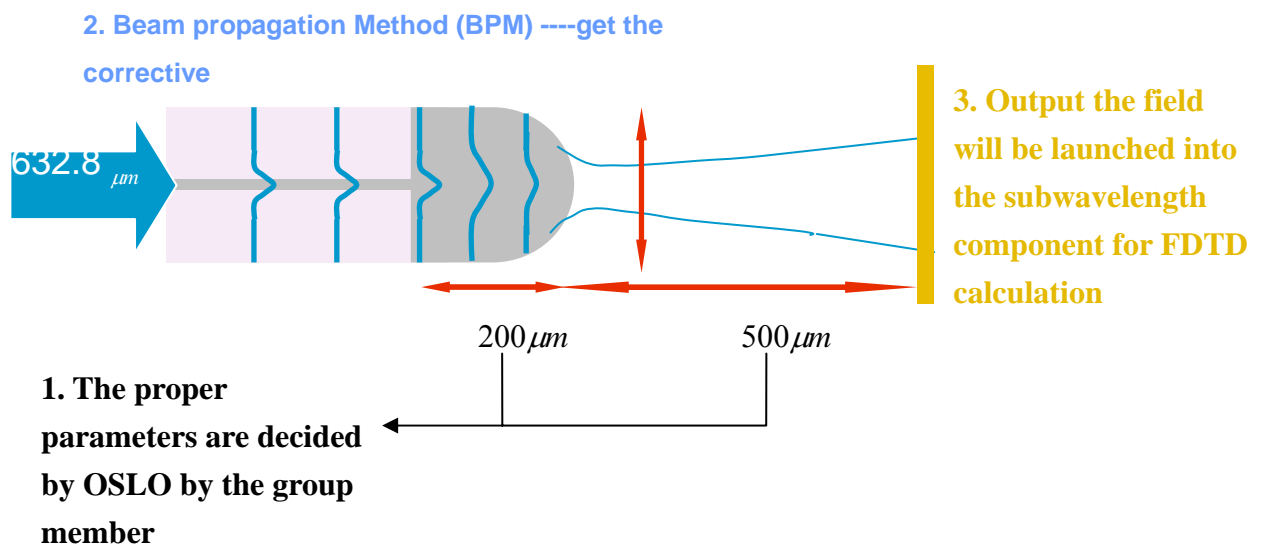
Fig. 4-1 The flow chart of our simulation

Here the aperture size was much smaller than the incident spot size; therefore the field incident was treated as a plane wave incident. The peak intensity of light was assumed to be the intensity of the plane wave for P.T. calculation.

A central challenge with the use of near-field apertures is the transmission of adequate power through the aperture, while confining the field intensity to a region much smaller than a wavelength. The transmittance field distribution through the subwavelength apertures with the infinitesimally thin and perfectly conducting planes of different shapes and dimensions were demonstrated. The combination of SIL and aperture under the same condition were also simulated and discussed. Finally, the properties of materials will be taken into considered to get optimal design of the nano-structure.
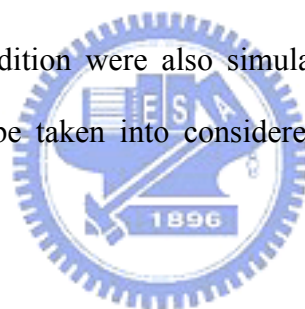
4.2 simulation software

The simulator BeamPROP 5.1.1 using BPM method and FullWAVE v3.0.1 using FDTD method are all developed by RSOFT design group. BeamPROP was used to compute the propagation of light waves in arbitrary waveguide geometries. FullWAVE was used to simulate the electromagnetic interactions with materials to further design to do further design the nano structure.

4.2.1   BeamProp

The computational core of the program is based on the finite difference beam propagation method. The fundamental physical limitation of the above approach

results from the parabolic approximation to Helmholtz equation, which implies a paraxiality condition on the primary direction of propagation. These limitations can be reduced using more accurate approximations to the Helmholtz equation.

A second limitation of the above approach results from the assumption of scalar waves, and prevents polarization effects from being considered. The third key limitation of the BPM approach is that it can not account for backward reflections since the one-way wave equation on which it is based does not admit both positive and negative traveling waves. However, BeamPROP introduced a new bidirectional BPM algorithm which considers coupled forward and backward traveling waves, and can account for reflection phenomenon, including resonant effects in grating structures.

4.2.2   Fullwave

The FDTD method as described in chapter 2 is a rigorous solution to Maxwell's equations and does not have any approximation or theoretical restrictions. This method is widely used as a propagation solution technique in integrated optics, especially in situations where solutions obtained via the Beam Propagation Method (BPM) cannot cope with the structure geometry or are not adequate solutions. Here the field output from BeamPROP can be launched into FDTD to do further simulation.

4.3   Start with infinitesimally thin and perfectly conducting planes

Owing to the effects of thickness and material are less apparent than that of

abnormity shape, the bare aperture based on perfect conducting and negligible thin sheet provides a good starting point for further optimization. Here aperture designs start with negligible thin perfectly conductor and then properties of metals will be taken into considered in next paragraph.

4.3.1 bare square aperture
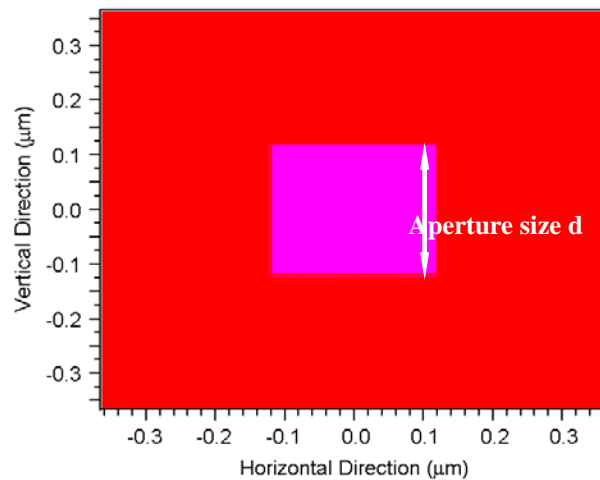
Fig. 4-2 Cross section of square aperture

According to Bethe's theory [1] mentioned before, the hole is supposed to be very small compared with the wavelength, the field near a subwavelength aperture in a perfectly conducting sheet illuminated with a normally incident plane wave is suffered from extremely low power throughput decays as the forth power of the aperture size. It was verified as shown in Fig.4-3. When the ratio of aperture size to incident wavelength is 0.1, the P.T. is closed to 0.007. The results got from our simulations agreed Bethe's theory.
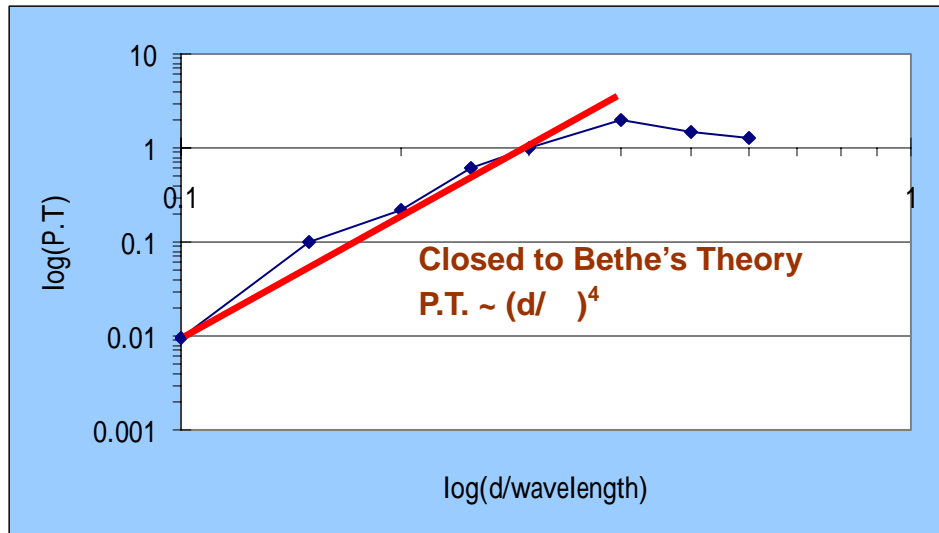
Fig. 4-3 Power throughput of a square aperture in a perfectly conducting metal screen as the aperture size decreases. The incident light is a plane wave and the thickness of the metal screen is one grid size in the simulation, which is negligible small compared with the light wavelength

The thickness of the perfectly conductor plate is set to be one grid size in our simulation under TE wave incidence. For the perfectly conductor plate, the imaginary part of refractive index is much larger than real part. The imaginary part of refractive index k is 1000 and the real part of refractive index n is 1 in simulation.

4.3.2 C-shape aperture
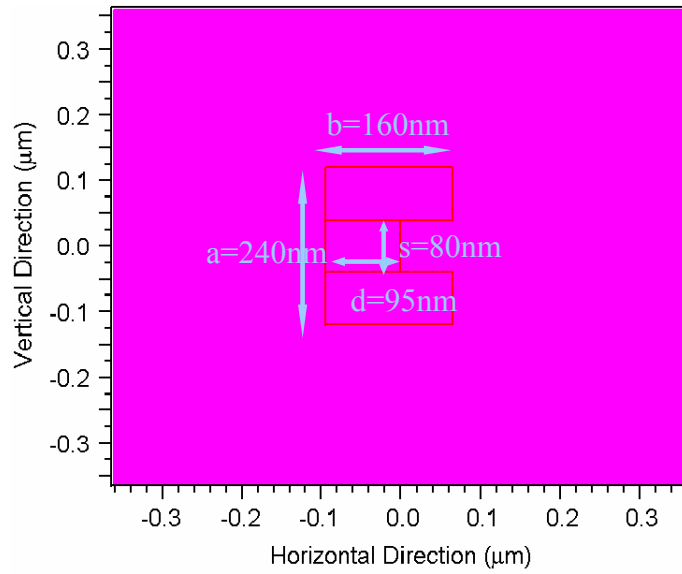
Fig. 4-4 Cross section of the c-shape aperture based on negligible thin perfectly conductor
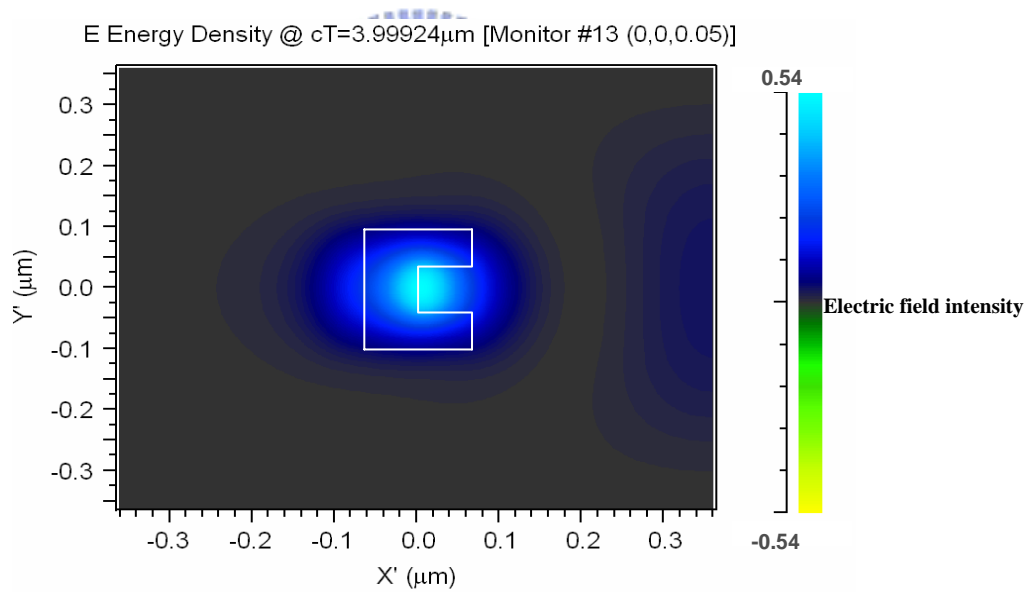
Fig. 4-5 Electric field intensity distribution 50nm from the c-shape aperture

(a=240nm, b=160nm, s=80nm, d=95nm; + and − mean the different

directions for electric field to propagate)

The quantitative comparison as performed by the researchers strongly supports the idea that the propagation TE_{10} mode leads to the greatly enhanced power

throughput from the c-shaped aperture. The cut-off wavelength of the corresponding ridge waveguide and the optimal power throughput wavelength for the aperture are much closed to each others.

Owing to that, when the incident light wavelength is 633nm, the dimension of C-shape aperture was designed based on the design of ridge waveguide with equation 2-21. The correction value of F can be obtained according to the design curves of ridge waveguide in Fig. 2-6 and computed values of λ_c / a for a single rectangular ridge in a rectangular waveguide of b/a ratio equal to 0.45 are shown in Table 2-1. And from the observation the spot size depended on the length of the center of the inner side as shown in Fig. 4-5 and 4-6. The optimal power throughput for C-shape aperture is 6.3 and the maximum electric field intensity is 0.55 observed at 50nm from the aperture as shown in Fig. 4-5.
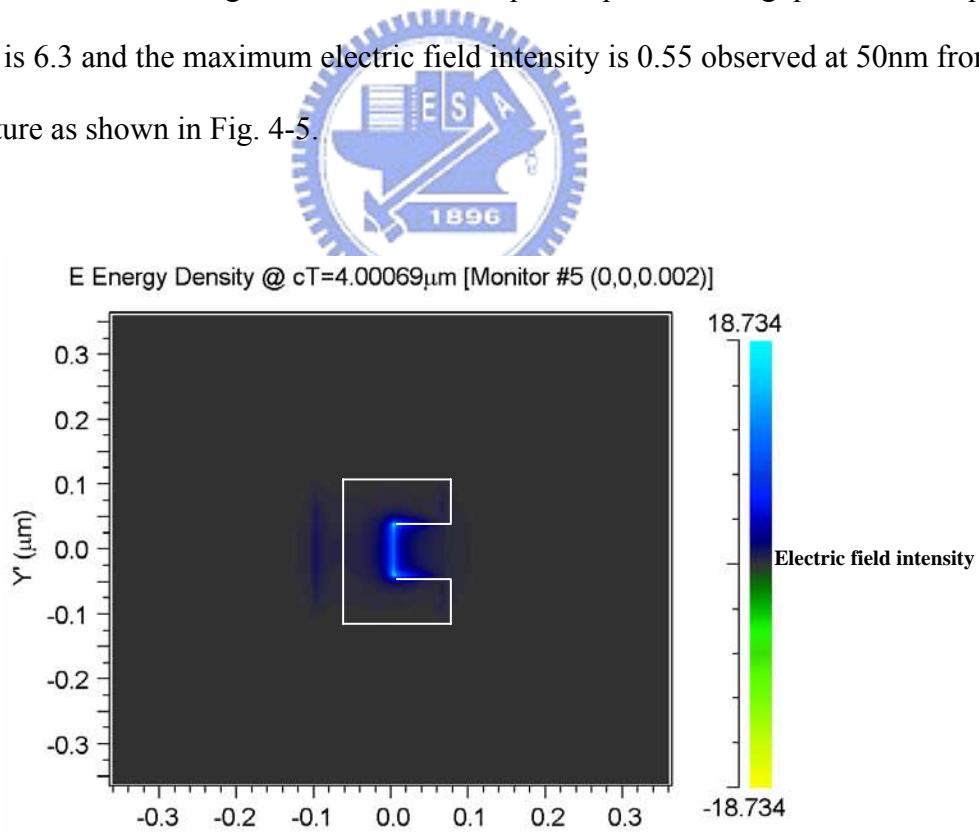
Fig. 4-6 Electric field intensity distribution 2nm from the c-shape aperture (a=240nm, b=160nm, s=80nm, d=95nm; + and − mean the different directions for electric field to propagate)

4.3.3 Combination of SIL and aperture

As mentioned before, SIL was used to reduce the spot size and also raise the intensity. The cross section of the combination was shown in Fig. 4-7 where the radius of SIL was optimized by OSLO to 8 μm. P.T. was 0.015 and the Max. electric field intensity is 0.0108.
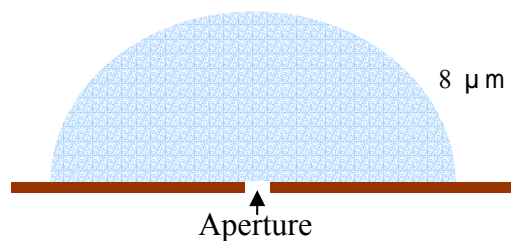
Fig. 4-7 Cross section of a SIL-aperture combination

4.3.4 Comparison

The comparison between the three cases with infinitesimally thin and perfectly conducting planes is shown in Fig. 4.8, the C-shaped aperture compared with a 120 nm square aperture provides ~900 × P.T. enhancement while maintaining a comparable near-field spot size and that compared with a combination of SIL and a 120 nm square aperture provides 250 × P.T.. For the c-shaped aperture, 120 nm-square aperture and combination of SIL and square aperture, P.T. was 3.6, 0.01 and 0.015 respectively; the near-field spot size (FWHM) at 50nm was 90 nm × 140 nm, 85 nm × 138 nm, 75 nm × 130 nm. Respectively, the maximum intensity was 0.54, 0.0006 a.u. and 0.0138 a.u..
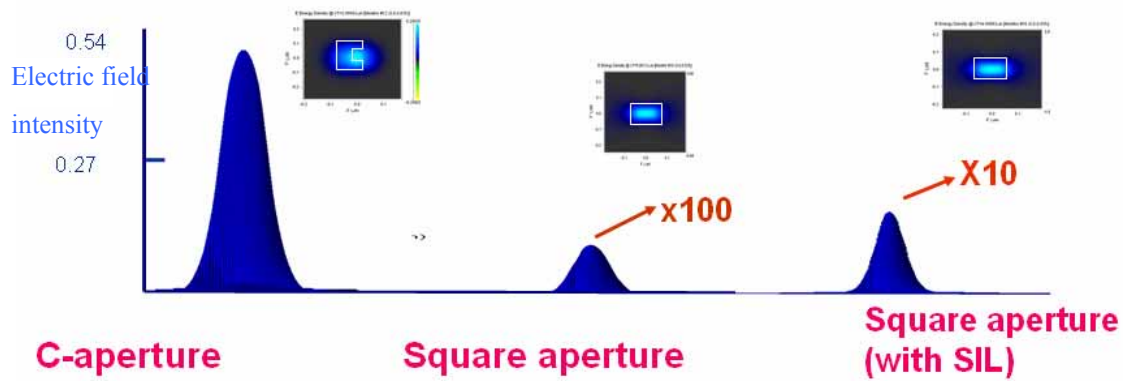
Fig. 4.8 comparison between the three cases: c-aperture, square aperture, and square aperture combined with SIL

4.4 Properties of metals taken into considered in real case

The properties of metal were taken into considered for the real case. Since we were interested not only in optical heating of materials but also that whether the metal film was thick enough to completely block the incident beam, it was important to interpret the above dispersion relation correctly. Under our conditions, the material should be selected to satisfy that higher surface plasmon can be excited and shorter penetration depth is found to avoid too much loss arisen from thick metal screen. If the metal film was thin, aperture transmission is high. On the other hand, if the metal film was too thin, light leaks through the portions of the metal surrounding the aperture.

4.4.1 Bare square aperture surrounding metal screen

In order to optimize metal screen surrounding the aperture, the simplest situation taken into account was the square aperture surrounded with different metals.

To obtain sufficient contrast between light that propagates through the aperture and that which leaks through the metal surrounding the hole, however, the metal must be sufficiently thick. For the apertures in thick metallic screens the light propagating through the aperture is attenuated strongly with increasing thickness. So, although the screen must have sufficient thickness to block the light leaking through the metal, it cannot be so thick that the light propagating through the aperture is attenuated to an unacceptable level. Thus, there is a finite range of metal thickness that provide sufficient contrast between the light coming through the aperture and that leaking through the surrounding metal while maintaining a reasonable intensity of the light propagating through the aperture.

Here the thickness of metal films made of aluminum, copper, silver, and gold surrounding the square aperture is optimized shown in Figs. 4-9 to 4-12.
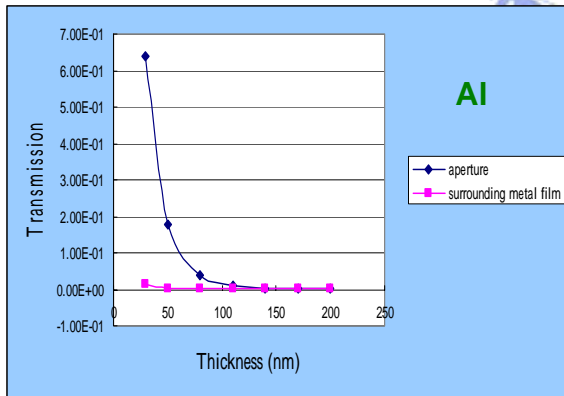
Fig. 4-9 Power throughput of bare square aperture and surrounding Al film V.S. thickness
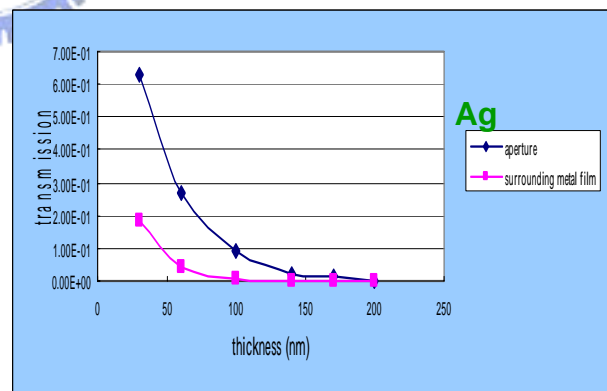
Fig. 4-10 Power throughput of bare square aperture and surrounding Ag film V.S. thickness
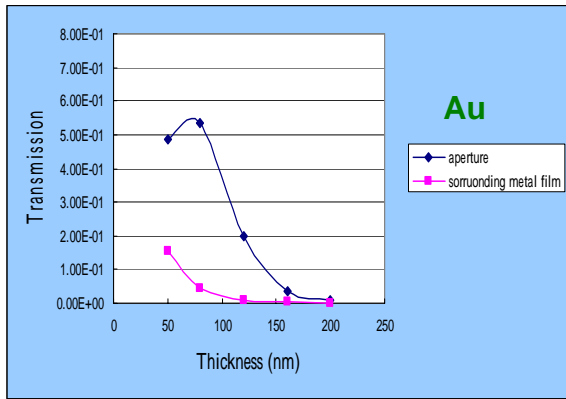
Fig. 4-11 Power throughput of bare square aperture and surrounding Au film V.S. thickness
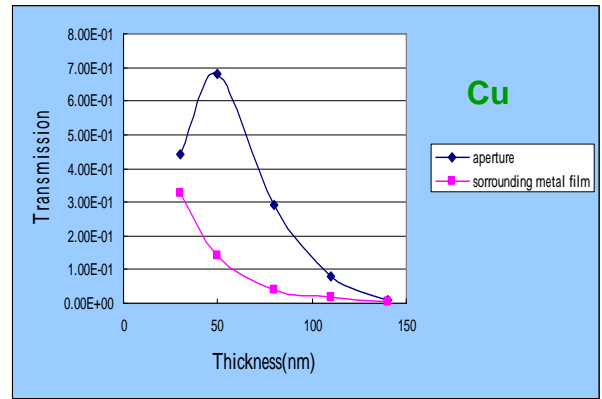
Fig. 4-12 Power throughput of bare square aperture and surrounding Cu film V.S. thickness

The optimum thickness of aluminum, silver and gold are found to be between 70 to 90nm, 80 to 100nm and 90 to 110nm. Finally, aluminum was chosen to be the material because that it can be easily fabricated, not easily to be oxidized and good performance.

In order to understand in detail the role of the dielectric properties of the metal, the dielectric constants of these four metals are compared under the 633 nm wavelength incident. To compute the dielectric function, $\varepsilon = \varepsilon_1 + j\varepsilon_2$, from the index of refraction $N = n + jk$, use $\varepsilon_1 = n^2 - k^2$ and $\varepsilon_2 = 2nk$. At $\lambda = 633$ nm,

Table 4-1 Dielectric constants for materials at the wavelength= 630 nm [1]

Material	ε_1	ε_2
Aluminum	-57	21.4
Copper	-10	3.7
Silver	-16	2
Gold	-13.3	0.73

$\varepsilon_{Ag} = -16 + j2$, $\varepsilon_{Cu} = -12 + j1.7$, $\varepsilon_{Au} = -13.3 + j0.73$ whereas $\varepsilon_{Al} = -57 + j21.4$.

This implies that Al is much closer than the other metals to the perfect conductor ($-\varepsilon_1 >> \varepsilon_2$), and is, therefore, expected that higher excitation of the surface plasmon is occurred. And skin depth $\delta \approx 25nm$ for Ag, $\delta \approx 14nm$ for Al, $\delta \approx 34nm$ for Au and $\delta \approx 48nm$ for Cu. The max. intensity of the field distribution is reduced as shown in fig.4-13 due to the finite thickness and absorption of metal.
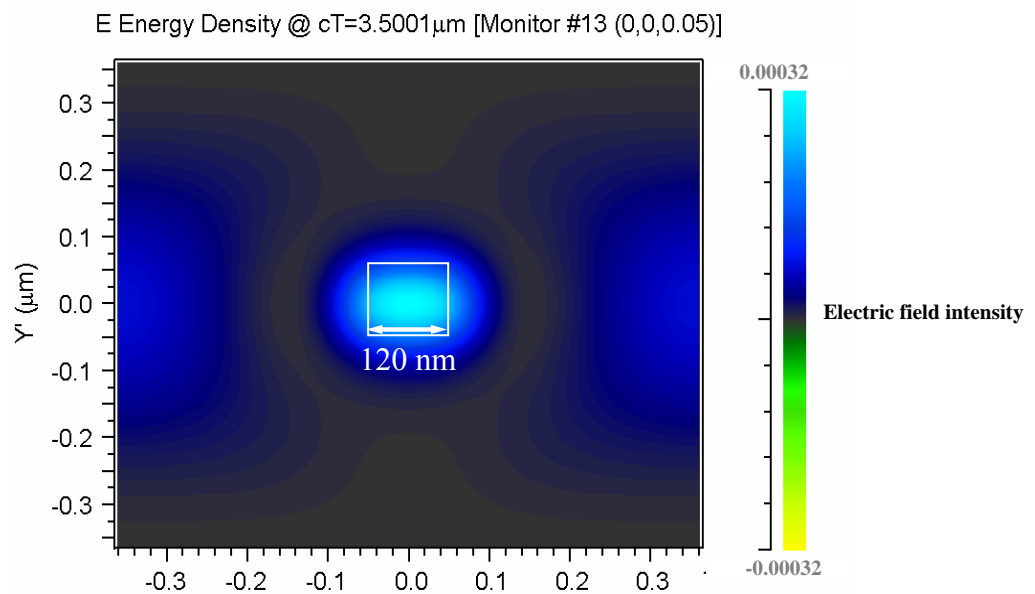
Fig. 4-13 electric field intensity distribution 50nm from the square aperture d=120 nm with thickness of 80 nm-Al; (+ and – mean the different directions for electric field to propagate)

4.4.2 c-shaped aperture surrounding metal screen
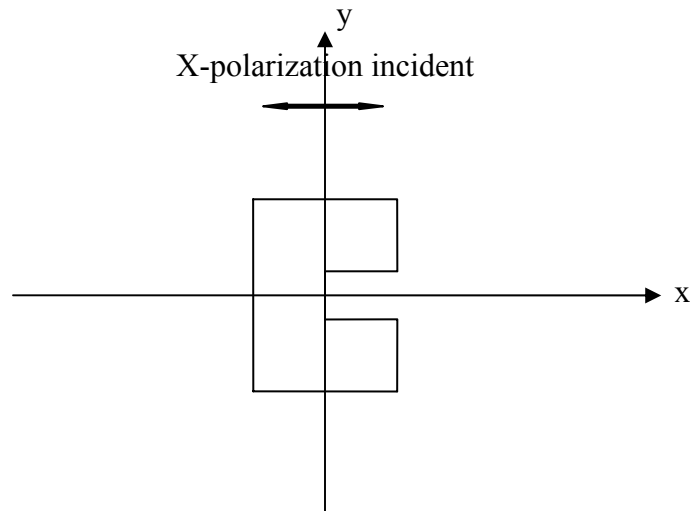
Fig. 4-14 The cross section of the c-shaped aperture

Aluminum is chosen as the waveguide material and the Drude model is chosen to describe the dispersion relation because that it is treated as a free electron model. The input beam is a linearly x-polarized plane wave with unit power. In addition this field enhancement is a function of wavelength and peaks at about 900 nm. The waveguide was designed to resonant at 633nm. The shift in resonance towards higher wavelength is due to the field penetration into the metal as a result of the loss. By replacing aluminum with a perfect conductor, this resonant wavelength should be designed to be lower than 633nm.

From the observation on the distribution of electric field intensity, the energy is observed almost gather at the ridge that has to be verified through simulation. Here for a dimension designed for a 560 nm-cutoff wavelength ridged waveguide, a is fixed at 150 nm and b is fixed at 150 nm and 100 nm with different s varied from 30 nm to 120 nm to observe the relation between spot size and s as shown in Fig. 4-15. We found that the y-spot size is mainly determined by s. Then the shift toward 633 nm has to be characterized.
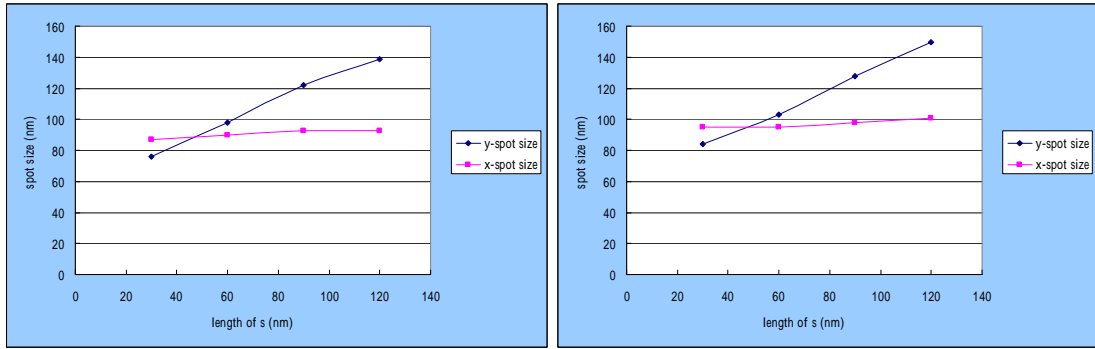
(a)   (b)

Fig. 4-15 The relation between length of s and spot size with fixed a, b and cutoff

wavelength 560 nm (a) a=150 nm, b=100 nm (b) a=150 nm, b=150 nm
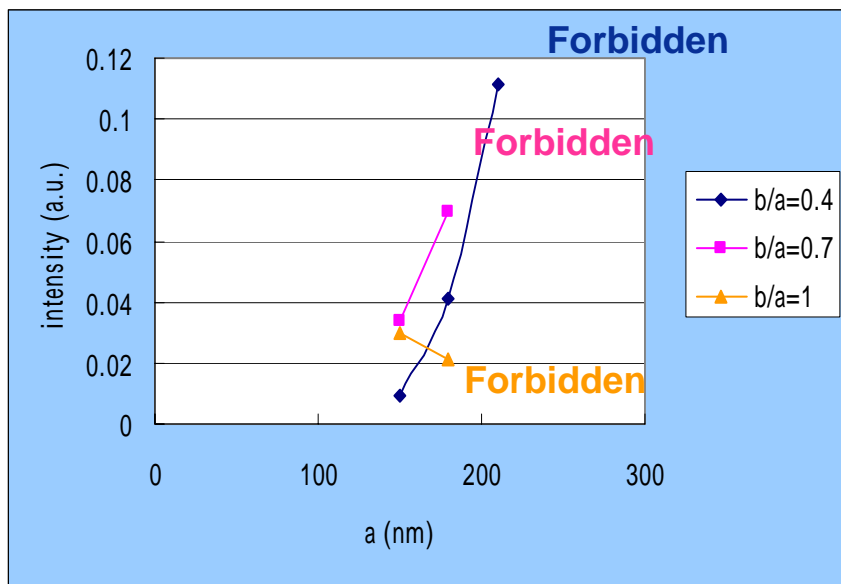
Fig. 4-16 Intensity vs. different dimension of ridged waveguide designed for 560 nm

cutoff wavelength (a=150 nm, 180 nm, 210 nm, 240 nm)

As shown in Fig. 4-16, based on the spot size compared with that of square

aperture, a is chosen from 150 nm to 240 nm with fixed s at 60 nm whereas b/a of the

c-shape aperture varied from 0.4 to 1 with various dimensions designed for 560

nm-cutoff-frequency ridged waveguide.

We found that the resonant wavelength shift toward 633 nm decreased when a increased from 150 nm to 240 nm because of the decreasing loss arisen from the larger dimension. It is a good starting point to decide the way to continue the optimization to find the optimal case.

Different d is chosen whereas a is chosen from 150 nm to 240 nm with fixed s at 60 nm and b/a of the c-shape aperture varied from 0.4 to 1 to get the optimization results. As fig. 4-18 shown below, the optimization dimension is that a=210 nm, b=85 nm, s=60 nm and d=50 nm.

If the cutoff wavelength of the c-shaped aperture is smaller than 450 nm, there is no propagation mode exited in the aperture. When the cutoff wavelength is chosen to be around 1400 nm, there will be other propagation modes exited. The transmitted position of the light is different from that of TE_{10} mode for other propagation modes with increasing spot size as shown in Fig. 4-17.

E Energy Density @ cT=2.60005μm [Monitor #13 (0,0,0.04)]
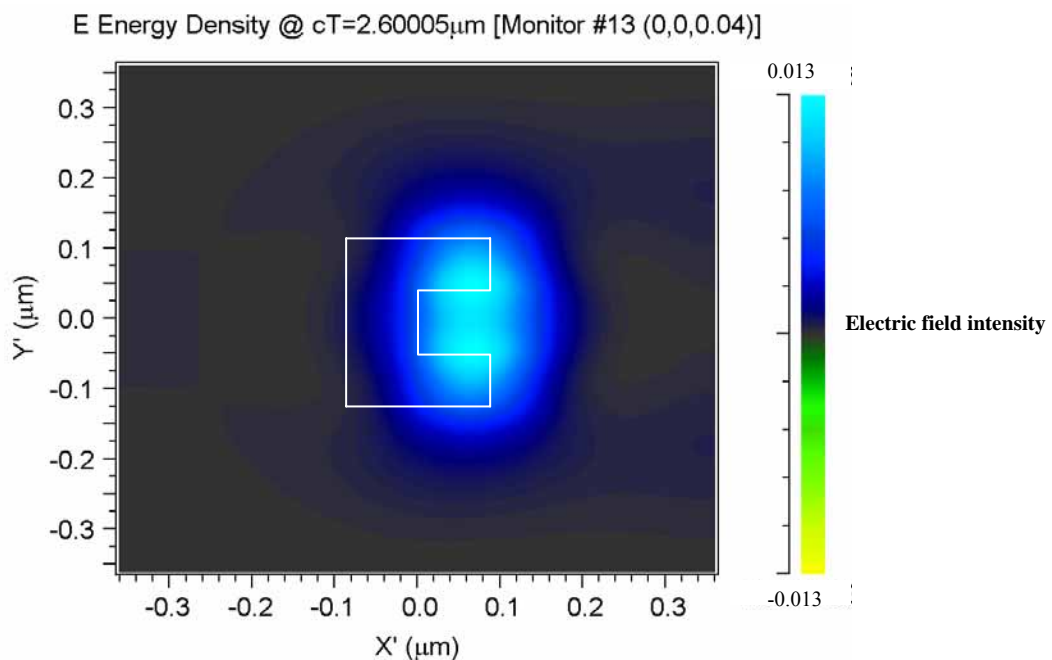
Fig. 4-17 electric field intensity distribution 50nm from the c-shape aperture for

1400 nm cut-off wavelength (+ and − mean the different directions

for electric field to propagate)

The purpose of the simulation is to find a dimension for high intensity, large P.T.

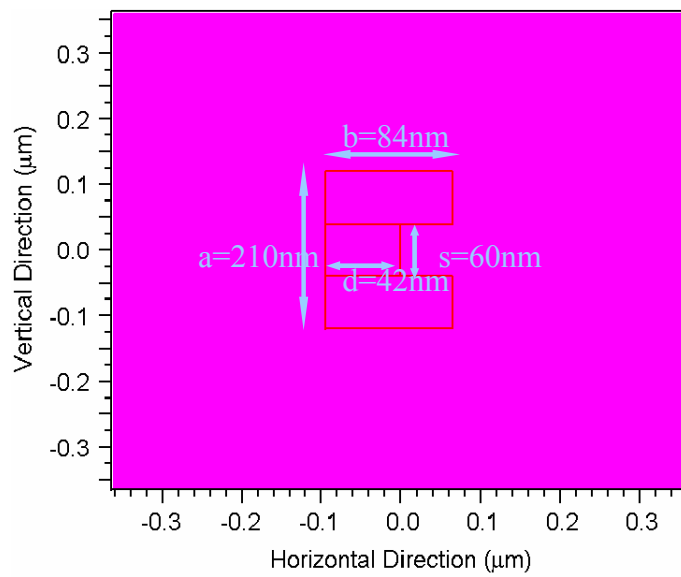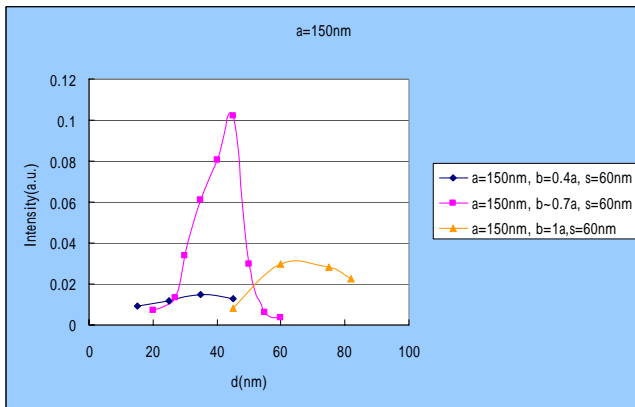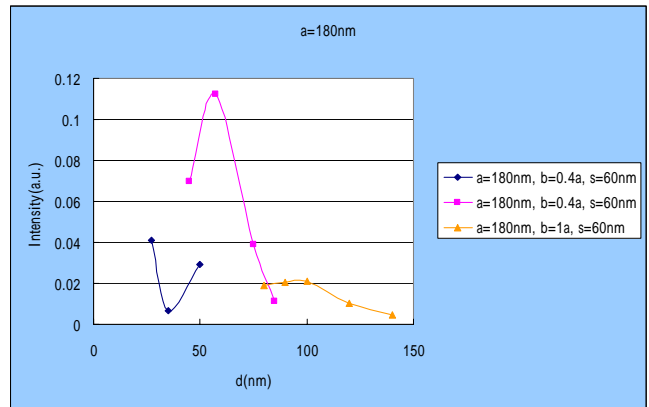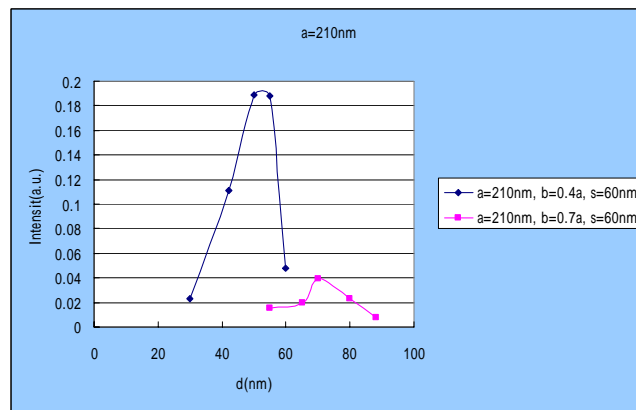and small spot size compared with that of 120 nm-square aperture.

Fig. 4-18 Cross section of the c-shape aperture with Al-80nm

Fig. 4-19 Different d is chosen whereas a varies from 150 nm to 240 nm

fixed s at 60 nm and b/a of the c-shape aperture varied from 0.4 to 1
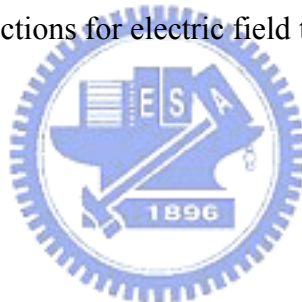
(a)for a= 180 nm, (b) a= 150 nm, (c) a=210 nm
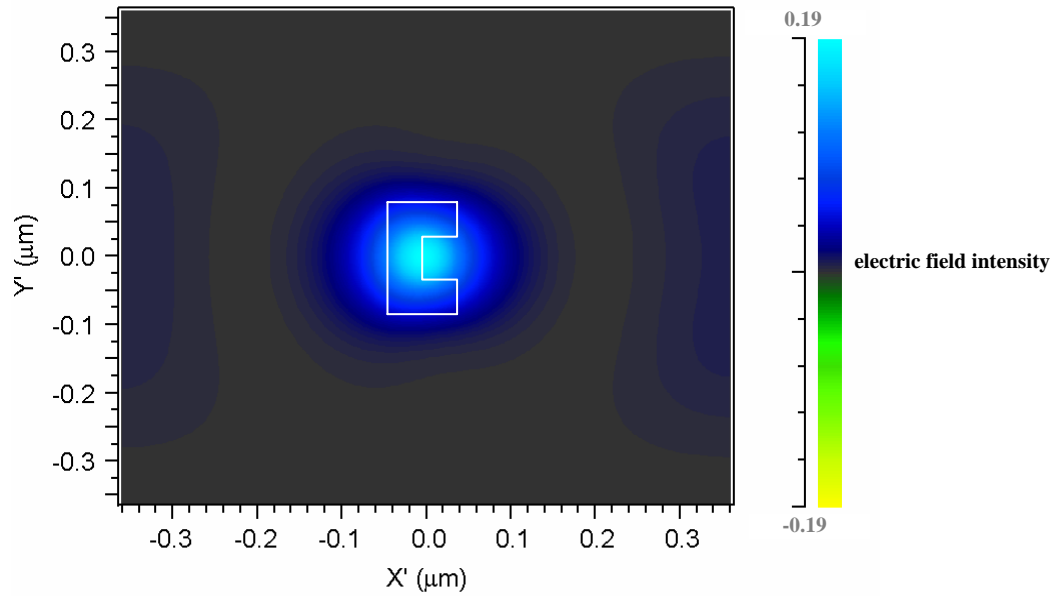
Fig. 4-20 Electric field intensity distribution 50nm from the c-shape
aperture (a=210nm, b=84nm, s=60nm, d=42nm; + and − mean the
different directions for electric field to propagate)

4.4.3 Comparison

For the c-shaped aperture and 120 nm-square aperture with Al 80 nm-thick P.T. was
1.6 and 0.017, respectively; the near-field spot size (FWHM) at 40nm was 108 nm ×
108 nm and 146 nm × 118 nm. Respectively, the maximum intensity was 0.11 and
0.0003. The P.T. and maximum intensity from the c-shaped aperture was therefore
~100 times and ~400 times higher than those of square aperture.

Chapter 5

Fabrication and measurement

5.1 Fabrication (Focused Ion Beam)

The features of the performance of abnormity shape apertures have been demonstrated by the simulation. Therefore, we start to fabricate a square and c-shape aperture to confirm the simulation results and to compare between the performances of square apertures and c-shaped apertures. However, in order to confirm the performances of the whole system, the apertures were fabricated on the end surface of the fiber and laser was directly coupled into the fiber to do the measurement. The detail steps to produce the sample were listed below:

(1) Substrate preparation: Because that the spot size illuminated onto the aperture is very large compared with the aperture, the wave can be viewed as a plane wave incident to the apertures. So the distributions of the field output from the apertures are not changed except of the intensity under the different situation of spot illuminated.

(2) Aluminum deposition: After the fiber is prepared and cleaned, an aluminum film with thickness of 90 nm was sputtered on it.

(3) FIB micromachining: The apertures with desired dimension according to the simulation results were fabricated by Focused Ion Beam (FIB) in the central of the end surface of the fiber sputtered by Aluminum. The fiber is fixed on the holder and the remained free section should be as short as possible to

avoid the vibration of the fiber tip because any vibration may cause in correct pattern due to the high resolution of FIB. The cross-section of the fiber and the holder were shown in figs. 5-2 and 5-3.

Fig. 5-1 Schematic diagram of Micrion 2500 Focused Ion Beam

FIB Specification:

Micrion 2500 FIB with Ga liquid metal ion source (LMIS) and charge
 neutralizatuion unit.

Max. sample size: 4" wafer (75 mm) can be loaded

Magnification: 100x – 250.000 xs

Resolution: 5 nm

Acceleration voltage: 5 kV-50 kV

Beam diameter: 5 -1100 nm (FWHM)

Beam current: 1 pA -20 nA

Available metal depositions: W, Au, Fe

The gallium-ion beam current was varied, depending on the particular task, from approximately 800 pA down to 1 pA. The stronger the ion current is, the worse the resolution is. Therefore, there is a trade-off between the etch rate and the resolution. The acceleration voltage is chosen to be 50 kV and the blanking aperture is set to be 15 μm and the ion current is 1 pA. The pattern was drawn by CAD software and can be read and transferred into the operation commands directly.

After milling, the cross-sections of the c-shaped and square aperture were shown in figs. 5-3 and 5-4.

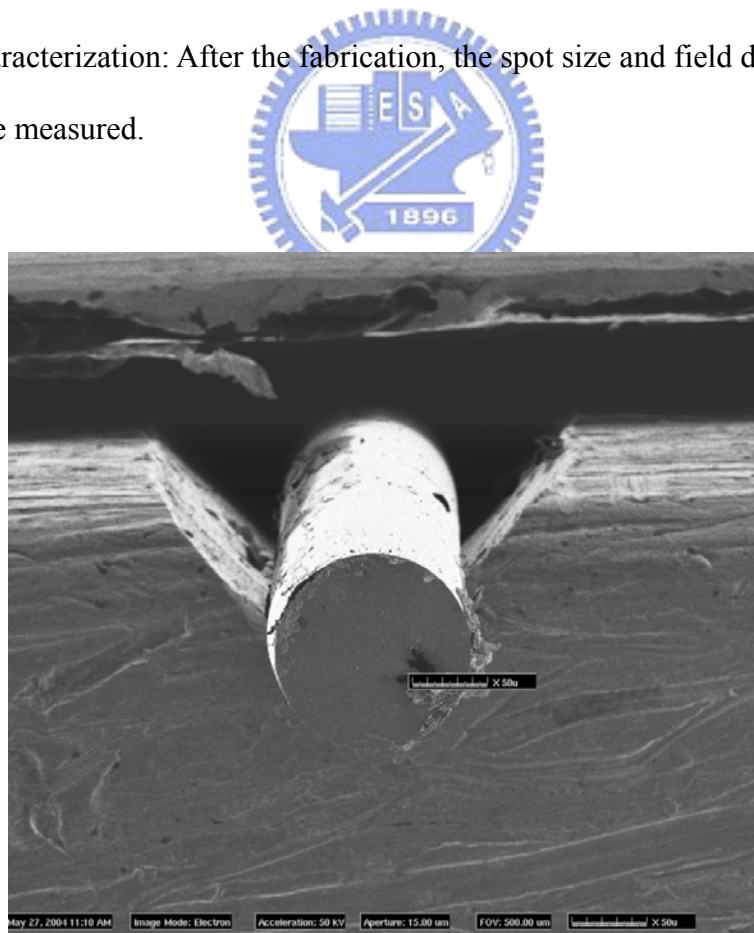(4)Characterization: After the fabrication, the spot size and field distribution were measured.
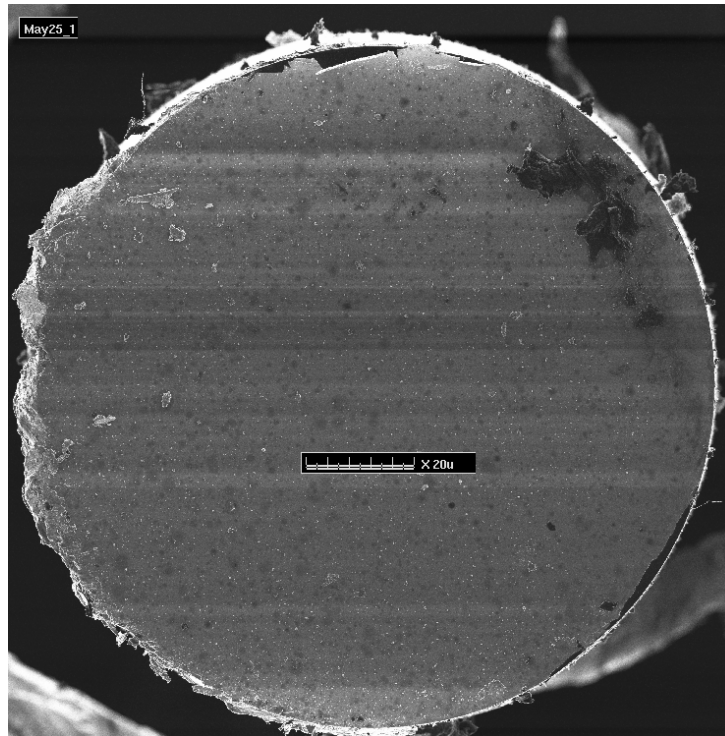

Fig. 5-2 The fiber is fixed on the holder
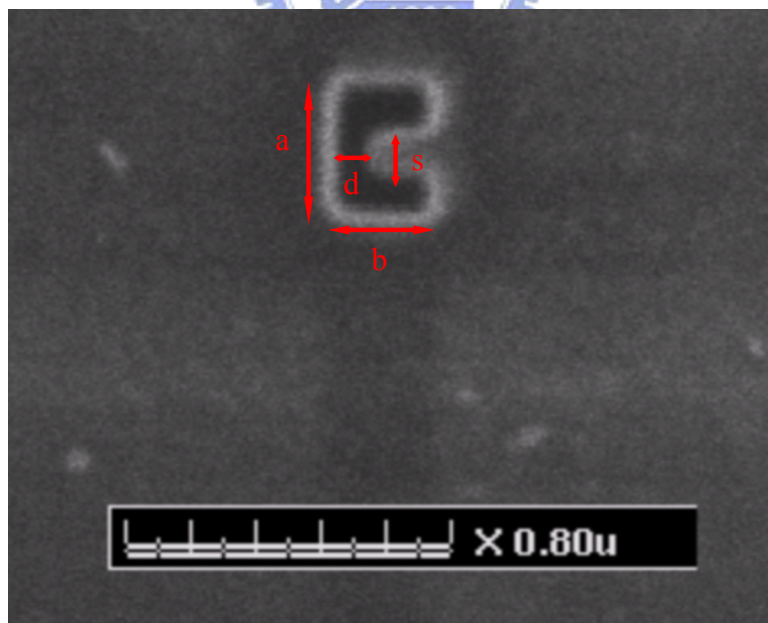
Fig. 5-3 The cross-section of the fiber

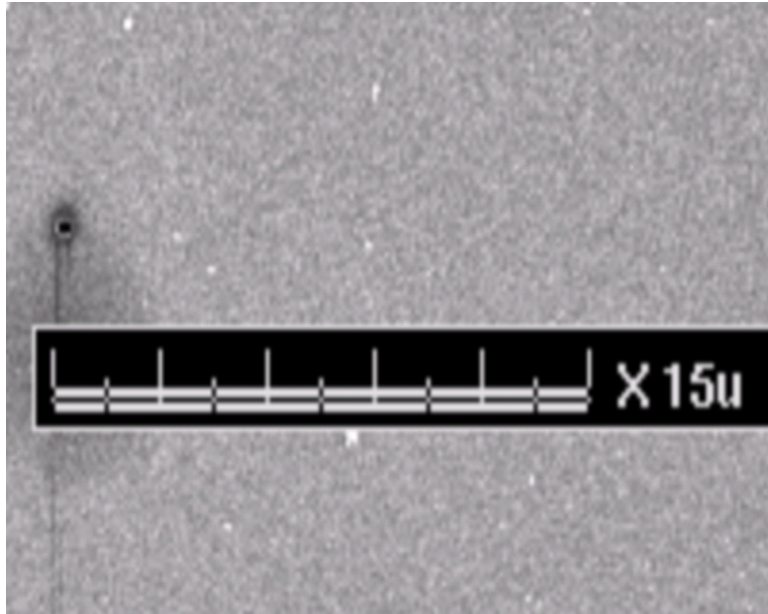Fig. 5-4 C-shaped aperture (~a=300 nm, b=200nm, s=90 nm, d=70 nm)

Fig. 5-5 square aperture (~250 nm)

5.2 experiment (Near-field Scanning Optical Microscope)

The whole setup of NSOM was shown in fig. 5-6. The fiber sample which 633 nm-laser is coupled into is mounted on a XYZ piezostage. An aluminum-coated silicon tip is brought to proximity over the aperture and oscillated parallel to the sample surface.

The bare tuning fork resonance is usually at 34 kHz, but this is typically increased to 30-40 kHz when the fiber is bonded to one tuning fork arm due to stiffening of the mechanical oscillator. The probes were mounted with approximately 2 mm extending beyond the mount. The resonance frequency of the fork/fiber can decrease due to poor glue joint or too much fiber extending over the arm of the tuning fork.

As described before, an aluminum-coated tapered optical fiber is rigidly attached to one of the prongs of a high Q tuning fork driven by a piezoelectric actuator. The fiber probe used is fabricated by HF-etching method. The fork is mechanically

dithered at its resonance frequency so that the tip amplitude does not exceed 400 nm. A corresponding piezoelectric signal is measured on electrodes appropriately placed on the prongs. As the tip approaches within 20 nm above the sample surface, a 0.1 nN drag force acting on the tip causes the signal to reduce. This signal is used to position the optical fiber tip to about 0 to 25 nm above the sample. Shear forces resulted from the tip-sample interaction can be quantitatively deduced. In each case, the microscope is controlling to sustain the same pre-set amplitude to avoid collision arisen from the defect of the surface or large particle. First the resonant frequency for the feed-back detection system is operated at 24 kHz to scan over the sample roughly and then the resonant frequency of the detection system is reduced to shorten the distance between the sample and fiber-tip.

The field distribution from the c-shaped aperture and square aperture were shown in figs. 5-6 and 5-7. Because the gray scale is normalized to the incident power, it can not easily to distinguish which one provides the largest P.T.. However, the field distribution of the near-field spot size is in an agreement with the simulation results.
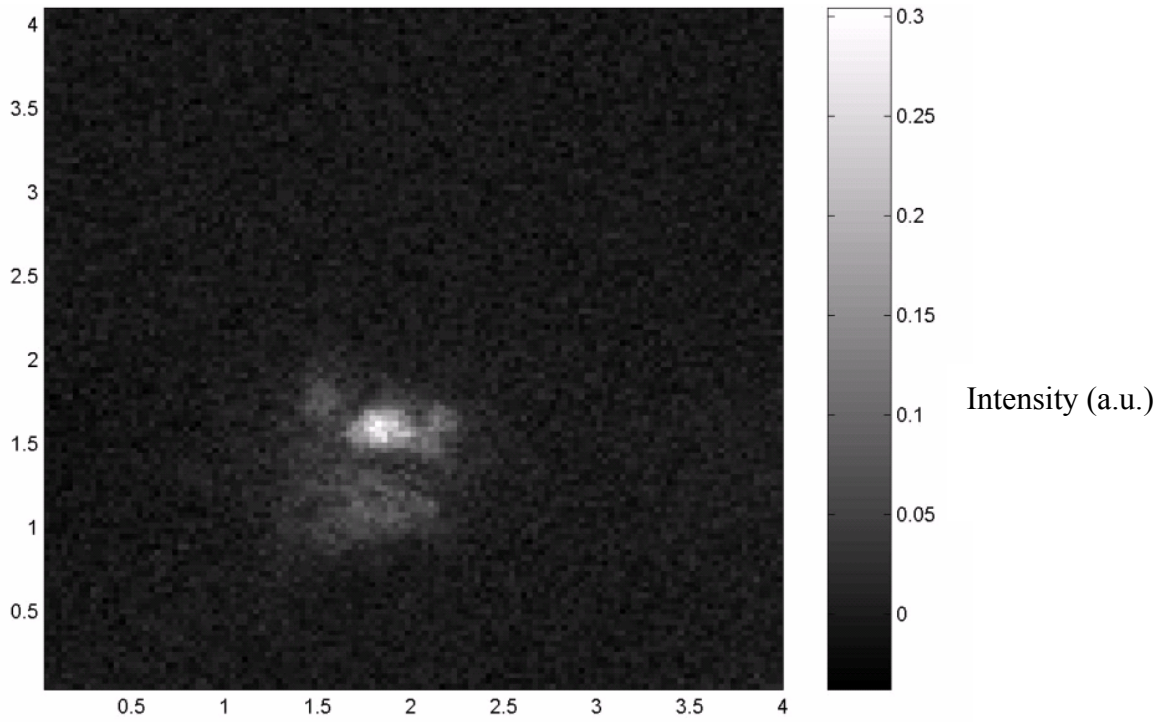

Fig. 5-6 The schematic diagram of NSOM

Fig. 5-7 Spot diagram of c-shaped aperture. The gray scale is normalized to the incident power
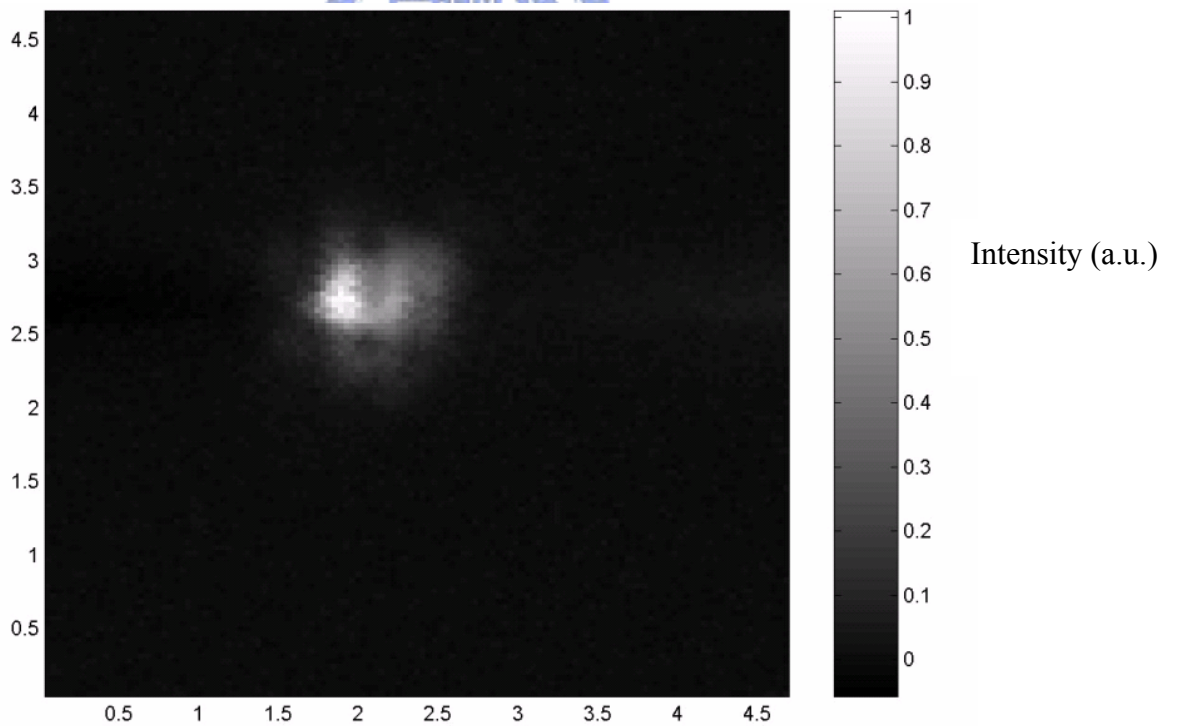

Fig. 5-8 Spot diagram of square aperture. The gray scale is normalized to the incident power.

Chapter 6

Conclusions

6.1 Conclusion

Near-field optical data storage systems are very important for achieving ultrahigh -density data storage as the minimum spot size for recording and readout is no longer limited by the optical diffraction limit. The fiber-based hybrid pickup system of our design utilizes planar nano metallic aperture to define the spot size. However, the problem encountered is that the extremely power throughput decays to unacceptable level while reducing the spot size.
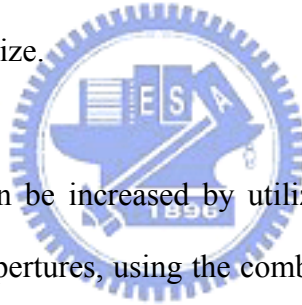
The power throughput can be increased by utilizing various methods such like employing abnormity shape apertures, using the combination of SIL [1] and aperture, and utilizing the surface plasmon excited modes to enhance the resonant transmission [2]. C-shaped aperture was reported to circumvent the extremely low P.T. due to the TE_{10} propagation mode exiting within the aperture.

The study of field distribution and optimization of the aperture were carried out by the numerical simulation finite-difference time domain (FDTD) method because the electromagnetic interaction and complex boundary condition taking place on the subwavelength scale are relevant to the optical properties measured.

As the simulation results shown, the power throughput decaying as forth power of the aperture size of square apertures in a perfectly conducting sheet with negligible

thickness is in an agreement with Bethe's theory. For a 120 nm square aperture, the power throughput is less than 0.01 and spot size (FWHM) is 145 nm × 95 nm at 50 nm. Under the same circumstances, the power throughput of c-shaped aperture was enhanced to be 6.3 because of the TE_{10} propagation mode exiting within it and spot size (FWHM) was is 115 nm × 105 nm at 50 nm.

In real case, the characteristic features of metal were taken into consideration. Compared with a perfectly conducting sheet, an aperture in an aluminum plate gave a larger spot size and lower P.T. because of the skin depth. Therefore, there was a resonant wavelength shift toward 633 nm for C-shaped aperture. Thickness of 80 nm aluminum was chosen to be the screen material to block the light leaking through the metal sufficiently and to induce a great contrast between the light propagating through the aperture and that leaks through the metal surrounding the aperture. A C-shaped aperture was designed with a P.T. of 1.9 and spot size (FWHM) of 100nm×110nm.
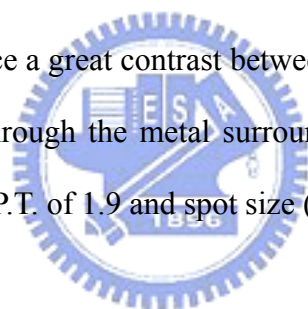
In the fabrication, in order to confirm the results of the whole system, the aperture was fabricated on the end surface of the fiber. FIB was used to carry out the square and c-shaped aperture on the end surface of the fiber. Due to that the fabrication of the c-shaped aperture was restricted by the resolution of FIB, the dimension of the aperture were not perfectly manufactured as same as the optimal case. For a convincingly measurement results, Near-Field Scanning Optical Microscope (NSOM) was utilized to confirm the simulation results. Finally, the field distribution measured of C-shaped agreed with the simulation results.

6.2 Future works

In the thesis, the enhancement of P.T. of C-shaped aperture have been analyzed and examined. In order to enable recording densities of $1\,Tb/in^2$ or higher, the optical spot must to be confined to less than 50 nm. There are several ways to improve the results to conform to the demands.

(a) The optimal C-shaped aperture should be designed under the restriction limited by the resolution of FIB, and be successfully fabricated on the end surface of the fiber.

(b) To combine SIL with nano metallic aperture is convinced to increase NA and maximum intensity of light illuminated on the aperture. However the new challenge encountered is that the wave incident on the aperture can not be viewed as a plane wave anymore. Therefore, the phenomenon as localized surface plasmon and propagation characteristics around the aperture will be changed. It is a significant subject to discuss and make it confirmed.