Chapter 2 Theory and Literature Survey

In this chapter, the most primitive theory with respect to the transmission through a subwavelength aperture and the fundamentality of the surface plasmon polariton (SPP) are reviewed; besides, the numerical approach, finite-difference time-domain (FDTD) method, is briefly interpreted as well. The literature surveys including the special shaped apertures and the surface corrugations at the end of this chapter are the prelude of the next chapter.

2.1 Bethe's Theory

As widely known, the geometric optics is a macroscopic description of light when the calculation scale is much larger than the wavelength so that the wave-like property and the influences of boundary conditions are suppressed and ignored. This assumption will obviously fail as the scale approximated to or smaller than the wavelength; hence, wave mechanics should be employed to treat boundary conditions rigorously. Bethe's theory is the most fundamental one to figure out the transmitted power through a subwavelength aperture in a perfect conducting film.

2.1.1 Mathematical Formulation and Physical Assumption

Consider a perfect conducting film placed at plane z=0, dividing the whole space into two parts as the left one of z<0 and the right one of z>0, and the light source is normally incident with wavelength λ, as shown in Fig. 2.1-1. Assume E_0 and H_0 to be the fields on the left-hand side of the film as if there is no aperture; these fields then will fulfill the boundary conditions at z=0

$$\vec{n} \times \vec{E}_0 = 0 \quad \text{----(1a)}$$
$$\vec{n} \cdot \vec{H}_0 = 0 \quad \text{----(1b)}$$

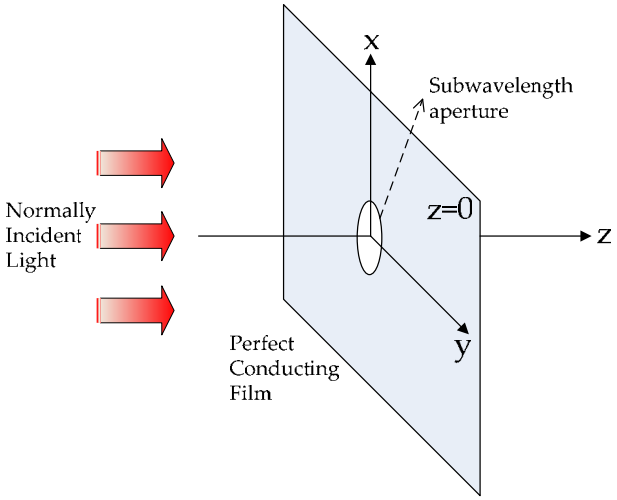


Fig. 2.1-1 Model of Bethe's theory

Due to the perfect conducting film, the field at the right-hand side vanishes identically and the boundary conditions are met everywhere on the film except within the aperture where the fields are discontinuous. The actual field is written as

$$\vec{H} = \vec{H}_0 + \vec{H}_1 \quad \text{and} \quad \vec{E} = \vec{E}_0 + \vec{E}_1 \quad \text{for } z > 0 \quad \text{----(2a)}$$

$$\vec{H} = \vec{H}_2 \quad \text{and} \quad \vec{E} = \vec{E}_2 \quad \text{for } z < 0 \quad \text{----(2b)}$$

and the boundary conditions are as follows:

$$\vec{E}_{1\,\tan} = \vec{E}_{2\,\tan}$$
$$\vec{H}_{2\,\tan} = \vec{H}_{1\,\tan} + \vec{H}_{0\,\tan} \quad \text{in the aperture ---- (3a)}$$

$$\vec{E}_{1\,\tan} = \vec{E}_{2\,\tan} = 0 \quad \text{for z=0 outside the aperture ---- (3b)}$$

A symmetric aperture allows the symmetric boundary condition imposed on x- and y-components of E as well as z-component of H with respect to z=0, and the rest components of E and H are anti-symmetric. Thus, two simple equations within the aperture are derived

$$\vec{H}_{2\tan} = \frac{1}{2}\vec{H}_{0\tan} \quad \text{----(4a)}$$

$$E_{2z} = \frac{1}{2}E_{0z} \quad \text{----(4b)}$$

Since the aperture size is extremely small compared to the wavelength λ, H_{0x}, H_{0y}, and E_{0z} can be considered as constants over the aperture; in addition, the aperture shape in the following calculation is supposed to be circular with radius a for simplicity.

Next, Stratton's formula [3], which represented E by full-vector theory, is introduced.

$$\vec{E}(r) = \frac{1}{4\pi}\int d\sigma\left[ik\vec{n}\times\vec{H}(r)\phi - \left(\vec{n}\times\vec{E}(r)\right)\times\nabla\phi - \vec{n}\cdot\vec{E}(r)\nabla\phi\right] \quad \text{----(5)}$$

Without electric charge or electric current, the first integral of above equation is omitted and a distribution of magnetic current density K and surface charge density η in the aperture play roles as the sources of the electric field accordingly. In the meantime, the continuity equation is fulfilled and the time dependence of all quantities is harmonic oscillation.

$$e^{i\omega t} \quad \text{and} \quad \omega = c\kappa \quad \text{----(6)}$$

The E and H can be expressed in terms of K and η with aid of scalar potential Ψ and vector potential F:

$$\vec{E} = \nabla\times\vec{F} \quad \text{----(7a)}$$

$$\vec{H} = \frac{1}{c}\frac{\partial\vec{F}}{\partial t} - \nabla\Psi \quad \text{----(7b)}$$

with F and Ψ represented by the integral of K and η over the plane of z=0, respectively,

$$\vec{F}(r) = -\int \vec{K}(r')\phi(|r - r'|)dr' \quad \text{----(8a)}$$
$$\Psi(r) = \int \eta(r')\phi(|r - r'|)dr' \quad \text{----(8b)}$$

where ϕ is the retarded radiation between the source vector r′ to the field vector r, as shown in Fig. 2.2-2.

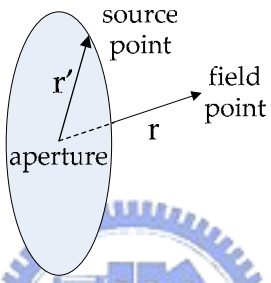


Fig. 2.2-2 Diagram of source point and field point

Finally, inserting eq.s (6)-(8) with aforementioned assumptions, E and H can be represented by the integral of magnetic surface current K and charge η over the aperture area.

$$\vec{E}(r) = \int \vec{K}(r') \times \nabla\phi d\sigma \quad \text{----(9a)}$$
$$\vec{H}(r) = \int \left[i\kappa \vec{K}(r')\phi - \eta(r')\nabla\phi \right] d\sigma \quad \text{----(9b)}$$

2.1.2 Determination of Surface Magnetic Charge/Current Density

Due to the much smaller aperture than the wavelength, H is approximated to be a static field corresponding to the surface charge density η. Thus, the first term in eq. (9b) can be neglected, resulting in a constant magnetic field component H_{tan} equal to ½H_0 over the aperture. In addition, the retardation term in eq. (8b) is omitted

because the field point is in the proximity of the aperture. Combining these statements, the scalar potential is condensed as

$$\Psi(r) = \int \eta(r') \frac{dr'}{|r - r'|} = -\frac{1}{2} \vec{H}_0 \cdot \vec{r} \quad \text{----(10)}$$

which had been solved in electrostatics [4]. Such a constant field is produced by a uniform distribution of dipoles in an ellipsoid, of which the direction is identical to that of the field; furthermore, the cross section of the ellipsoid in XY plane is equivalent to the aperture area. Therefore, the surface density of dipoles, μ, is proportional to the ordinate of the ellipsoid:

$$\mu = \left(a^2 - r'^2\right)^{\frac{1}{2}} \quad \text{----(11)}$$

and the surface charge density η is then proportional to

$$\eta = -C \frac{\vec{H}_0 \cdot \vec{r'}}{\left(a^2 - r'^2\right)^{1/2}} \quad \text{----(12)}$$

where C is a constant. Employing the same method in electrostatics [4] and elasticity [5], its value can be determined and the expression of magnetic surface charge density η and current K are derived:

$$C = \frac{1}{\pi^2} \quad \text{----(13a)}$$

$$\eta = -\frac{\vec{H}_0 \cdot \vec{r'}}{\pi^2 \left(a^2 - r'^2\right)^{\frac{1}{2}}} \quad \text{----(13b)}$$

$$\vec{K} = \frac{i\kappa}{\pi^2} \left(a^2 - r'^2\right)^{\frac{1}{2}} \vec{H}_0 \quad \text{----(13c)}$$

2.1.3 Boundary Condition for Normal Component of Electric Field (E_n)

According to eq. (9a), E can be derived by surface magnetic current density K; however, eq. (13c) implies that the surface magnetic current density K does not contribute appreciably. In eq. (13c), it suggests

$$\vec{K} \propto \kappa a \vec{H}_0$$

but

$$\nabla\phi \propto \frac{1}{a^2}$$

which leads to eq. (9a) over an area of order a^2. This contradicts that the normal component of E_0 shall be of the same order as H_0; thereupon, an additional magnetic current distribution K_E is introduced. This new current distribution, however, cannot give rise to any extra magnetic charge density or it will violate the matched boundary condition for the magnetic field; as a result, a gauge transformation of K_E is required:

$$\nabla \cdot \vec{K}_E = 0$$

Utilizing the trick in electrodynamics [3], the modified total magnetic current density is derived:

$$\vec{K} = \vec{K}_H + \vec{K}_E = \frac{1}{\pi^2}\left[i\kappa\left(a^2 - r'^2\right)^{\frac{1}{2}} \vec{H}_0 + \frac{\vec{r}' \cdot \vec{E}_0}{2\left(a^2 - r'^2\right)^{\frac{1}{2}}} \right] \quad \text{----(14)}$$

while the magnetic charge density is remained as eq. (13b).

2.1.4 Evaluation of Electric/Magnetic Fields and Total Radiation

With the surface magnetic charge/current density in eq. 13(b) and eq. 14, E and H can be evaluated by imposing them into Maxwell equations:

$$\vec{E} = \frac{1}{3\pi}\kappa^2 a^3 \phi_0 \hat{r} \times (2\vec{H}_0 + \vec{E}_0 \times \hat{r}) \qquad \text{----(15a)}$$

$$\vec{H} = -\frac{1}{3\pi}\kappa^2 a^3 \phi_0 \hat{r} \times (2\vec{H}_0 \times \hat{r} - \vec{E}_0) \qquad \text{----(15b)}$$

$\phi_0 = \frac{e^{i\kappa r}}{r}$: the first order approximation of ϕ

$\hat{r}$: unit vector in the direction of $\vec{r}$

Eventually, the Poynting vector, S, of the transmitted electro-magnetic field is:

$$\vec{S} = \frac{c}{4\pi}\vec{E} \times \vec{H} = \frac{c}{36\pi^3}\frac{\kappa^4 a^6}{r^2}\hat{r}\left(2\hat{r} \times \vec{H}_0 - \hat{r} \times \hat{r} \times \vec{E}_0\right)^2 \qquad \text{----(16)}$$

and the total transmitted intensity that integrates over all directions is:

$$S_{total} = \int_0^{\pi/2} \sin\theta d\theta \int_0^{2\pi} d\alpha r^2 \left|\vec{S}\right| = \frac{c}{27\pi^2}\kappa^2 a^6 \left(4\left|\vec{H}_0\right|^2 + \left|\vec{E}_0\right|^2\right) \qquad \text{----(17)}$$

2.1.5 Power Throughput and Discussion

Power throughput (PT) is defined as the ratio of the total transmitted power to the incident power impinging on the aperture. Based on Bethe's theory, PT of a subwavelength aperture is proportional to the fourth power of the aperture size, which is the basic theoretical foundation of the extremely low transmission through a subwavelength aperture.

Nevertheless, this analytical solution is supported by many premises. On the one hand, the screen is infinitely thin owing to supposing metal as a perfect conductor; on the other hand, the field expression is greatly simple due to the

symmetry of the aperture. Furthermore, this formula is valid with the much smaller aperture size compared to the wavelength and will fail while the aperture size is approximated to the wavelength. Afterwards, there were literatures trying to modify Bethe's formula [6, 7] for greater accuracy, the analytical solution is still almost impossible to obtain unless applying many assumptions.

2.2 Finite-Difference Time-Domain Method

Since the exact analytical solutions to the practical electro-magnetic problem are often formidably difficult, it is necessary to apply numerical approaches. Finite-difference time-domain (FDTD) method is one of the most powerful approaches owing to its rigorous calculation without any physical assumptions or premises. Its chief notion is to discretize Maxwell equations by using central difference operators to replace the differential operators in both time and space variables [8].

While **E** and **H** fields are represented by their discrete values on the spatial grid and are advanced in time steps of Δt, the components of the vector fields **E** and **H** are staggered so that every component of the **E** field is surrounded by the **H** field's four circulating components and vice versa, as shown in Fig. 2.2-1. The contour integrals of **E** (or **H**) along the cell's edges in Faraday's (or Ampere's) law circulate around the corresponding **H** (or **E**) field components at the cell face's center. Additionally, several commonly used material-dispersion models such as Debye, Drude, or Lorentz can be readily incorporated with the time-dependent formulation [9] via either a linear recursive relation [10] or auxiliary differential equation methods [11].

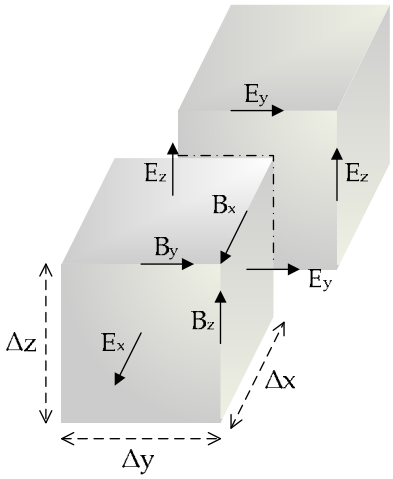


Fig. 2.2-1 Unit cell of FDTD mesh, components of **E** and **H** field are shifted by a half-pixel in x, y, z directions so that each **E** field component normal to the cell face is surrounded by the circulation of those **H** field components defined on the cell edges.

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

The boundary condition is a critical issue for FDTD method. In most of cases, low-reflection boundary conditions, such as Mur's absorbing boundary and the perfectly matched layer (PML) boundary condition [9], are popularly applied for those problems requiring open boundaries to simulate propagation into infinity. Furthermore, combining the total-field or scattered-field techniques [9] for the source-field distribution will allow accurate evaluation of the incident, reflected, and transmitted waves throughout the computational domain.

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

2.3 Surface Plasmon Polariton

A plasma is the quantum of the collective excitation of free electrons in solids, particularly in free-electron-like metals such as silver or gold, of which the dielectric constant can be expressed by Drude model as

$$\varepsilon_m(\omega) = 1 - (\omega_p / \omega)^2$$

where ω_p is the plasma frequency [12]. In the frequency range below ω_p, the electromagnetic properties related to the electron plasma effects are significantly different from those of ordinary dielectric materials owing to the negative real part of the dielectric constant. As a result of the imaginary wave vector, the propagation modes cannot exist in metals.

A surface plasmons is the electron plasma oscillation near a metal surface, stemming from the broken translational invariance in the direction perpendicular to the surface. On a flat surface of semi-infinite metal, the surface plasmon frequency, ω_{sp}, can be roughly determined by

$$\mathrm{Re}\,\varepsilon_m(\omega_{sp}) = -\varepsilon_i \text{ and } \varepsilon_i > 0$$

where ε_i is the dielectric constant of the adjacent dielectric medium; nevertheless, this formula needs to be modified in real metals due to the surface effects and the spatial dispersion relations [13].

A combined excitation consisting of a surface plasmon and a photon is called a surface plasmon polaritons (SPPs). The major parameters characterizing SPP are illustrated in Fig. 2.2-1, where a metal film with the dielectric constant ε_m is in the middle of two lossless dielectric media with dielectric constants ε_1 and ε_2.

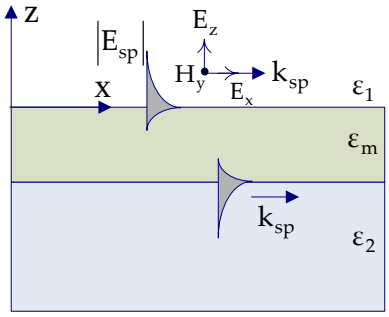


Fig. 2.3-1 SPPs on metal film interfaces with lossless dielectric materials

While the metal film is thick enough to neglect the interaction between SPPs on the opposite interfaces, the SPP modes existing on the film interfaces can be evaluated independently. The electric field of such an SPP mode propagating in X-direction can be expressed as [12]

$$\vec{E}_{sp}(x,z) = \vec{E}_0 e^{ik_{sp}x - k_z|z|}$$

$$k_{sp}^2 = \left(\frac{\omega}{c}\right)^2 \frac{\varepsilon_i \varepsilon_m}{\varepsilon_i + \varepsilon_m}$$

where ω/c is the wave vector of the incident light and ε_i (i=1,2) is the dielectric constant of the adjacent medium. This field configuration corresponds to an SPP mode, which propagates along the surface with wave vector k_{sp} and decays from the surface exponentially with the attenuation constant k_z. Because of the relation

$$k_z^2 = k_{sp}^2 - \left(\frac{\omega}{c}\right)^2 \varepsilon$$

where $\varepsilon=\varepsilon_m$ in metal and $\varepsilon=\varepsilon_i$ in dielectric materials and generally greater ε_m in comparison with ε_i, the SPP field will have a shorter penetration depth in the metal than in the adjacent dielectric material, as indicated in Fig. 2.2-1. As a result, the light near the interfaces can be enhanced by the SPP fields.

In symmetric surroundings $\varepsilon_1=\varepsilon_2$, the field distributions of the two modes at opposite interfaces are degenerate for a thick metal film, whereas for a thin one, they will split into two branches, symmetric (lower frequency) modes and antisymmetric (higher frequency) modes as a result of the interaction between SPPs on opposite interfaces. As implied by their names, these two modes possess symmetric and antisymmetric field distributions respectively; however, only antisymmetric SPP modes can facilitate field enhancement near the metal surface owing to its longer propagation length, which is inversely proportional to the square of the film thickness.

In order to excite SPPs by photons, illuminating on a smooth metal surface almost makes a futile effort; however, SPP modes can be stimulated on a metal surface with periodically distributed dielectric constant, which can provide an additional wave vectors to fulfill the matching conditions [14] , as illustrated in Fig. 2.2-2.

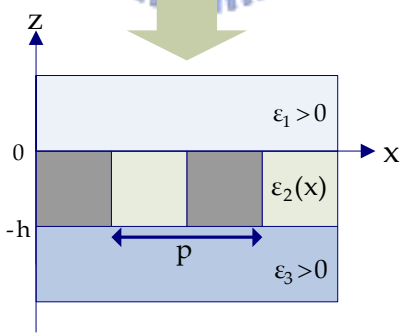


Fig. 2.3-2 Geometry of a periodically structured metal film

The electric field and the dielectric constant in such a periodically structured metal film can be expanded by Fourier series:

$$E = \sum_{n=-\infty}^{\infty} E_n \exp(ingx)$$

$$\varepsilon_2 = \sum_{n=-\infty}^{\infty} \varepsilon_n \exp(ingx)$$

$g = \frac{2\pi}{p}$, the vector of the reciprocal lattice related to the periodic structure

Imposing proper boundary conditions and employing analytical methods [6], three SPP eigenmodes are obtained:

$$\kappa_1^2 \sim -\varepsilon_0 k_0^2 - 2\alpha_1 \varepsilon_0 k_0^2 \left(\varepsilon_0 k_0^2 - g^2\right)/g^2$$
$$\kappa_2^2 \sim g^2 - \varepsilon_0 k_0^2 - 2\alpha_1 \varepsilon_0 k_0^2 \left(\varepsilon_0 k_0^2 - g^2\right)/g^2$$
$$\kappa_3^2 \sim -\varepsilon_0 k_0^2 + g^2 + 2\alpha_1 g^2$$
$$\alpha_1 = \varepsilon_1^2/\varepsilon_0^2$$

where κ_1, κ_2 and κ_3 are the wave vectors of the SPP fields.

For the conventional light tunneling without resonant SPP modes, the energy transferring between the photons on the opposite side are considerably inefficient and results in extremely low transmission. With the resonant SPP mode excited at one of the interfaces, it can enhance the transmission by augmenting more energy from SPP modes during tunneling process. Furthermore, with double resonances, there is no exponentially decreasing factor in the expression for transmission efficiency, which implies the strong field enhancement. These mechanisms are schematically depicted in Fig. 2.2-3

The tunnel energy exchange T in the double resonances, regardless of symmetric or asymmetric surroundings, is contingent on the overlapping between the different SPP modes [7]:

$$T \sim \int E_{spp1}(x,y) E_{spp2}(x,y) dxdy$$

where the lower indices 1 and 2 represent different interfaces. Applying more rigorous analytical treatment and modification of the metal film, the transmitted field is predicted to be enhanced up to 10^4 times [8].

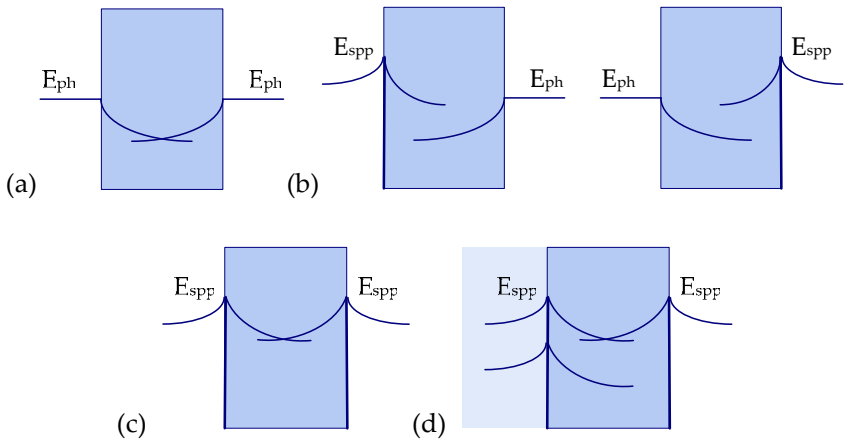


Fig. 2.3-3 Schematics of tunneling processes through a metal film: (a) no resonant SPP mode, (b) SPP modes excited on one of the film interfaces, and (c) and (d) SPP s excited on both interfaces in symmetric and asymmetric surroundings, respectively.

As a result of the considerably strong field enhancement stemming from the intrinsic properties of metals and electromagnetic waves, SPPs are greatly anticipating for optical recording, writing and detecting systems.

2.4 Special Shaped Apertures

Although Bethe's theory predicted the power throughput of a subwavelength aperture is proportional to the quartic of its size, this statement may fail if those supported premises are invalid. One of the used approaches is to reform the aperture shape so that the assumption of symmetric and circular apertures is subverted. This section will show several special-designed nano-apertures in sort of the published year.

2.4.1 Scatter-formed Aperture

A quite interesting innovation was to place a sub-micron sized metal scatterer in the aperture [18], as plotted in Fig. 2.4-1. The metal film functions as the suppression of background light and the metal scatterer which is used to excite surface plasmons acts as an extremely small light source.

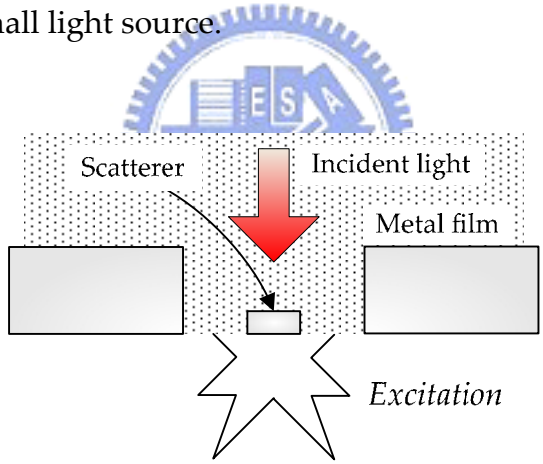


Fig. 2.4-1 Model of the planar aperture-mounted head with a minute scatterer (K. Tanaka *et al.* Jpn. J. Appl. Phys. Vol 40, P. 1542, 2001,)

The energy distribution along X- and Y-direction shown in Fig. 2.4-2 reveals the energy enhancements, but the edge effect at the scatter fringe is discernible and leads to two spots accordingly. Such a phenomenon also takes place with a conventional circular or rectangular subwavelength aperture at their edges perpendicular to the polarization of light. Moreover, the domain where the energy is enhanced is almost equivalent to the scatterer size, suggesting the spatial resolution is determined by the scatterer size rather than the aperture size. Nevertheless, this enhanced transmitted

power is still inadequate for practical application and the double-peak effect is a grave flaw for optical reading, writing and detecting systems.

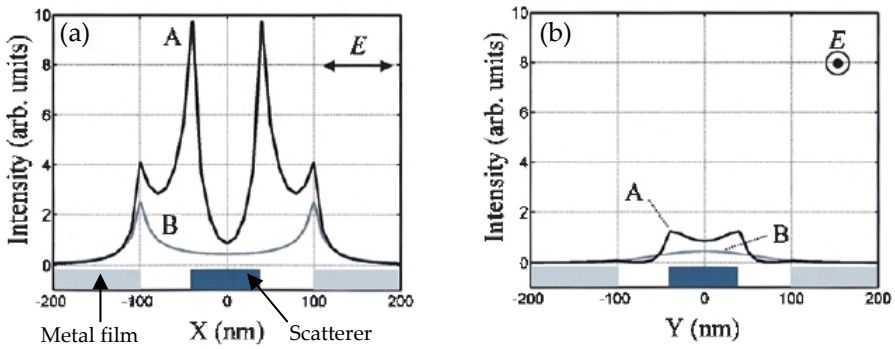


Fig. 2.4-2 Energy intensity profiles along (a) X-direction (at y=0) and (b) Y-direction (at x=0), the scatter and aperture area are 80*80 nm^2 and 200*200 nm^2, respectively, line A: with scatterer, line B: without scatterer. The insets describe the polarization of incident light (K. Tanaka *et al.* Jpn. J. Appl. Phys. Vol 40, P. 1543, 2001)

2.4.2 Triangular Aperture

A triangular aperture is designed as an equilateral triangular pyramid of which the bottom portion is truncated [19], as shown in Fig. 2.4-3.

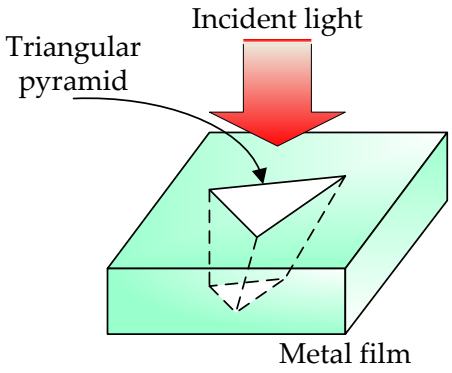


Fig. 2.4-3 Model of a near-field optical head with a triangular aperture (K. Tanaka *et al.* Jpn. J. Appl. Phys. Vol 42, P. 1114, 2003)

As shown in Fig. 2.4-4, the energy distribution is highly strengthened at the side perpendicular to the incident polarization; thereby, the double-peak effect can be eliminated. The transmitted power of a triangular aperture is several times than that

of a conventional circular aperture; however, this is insufficient for practical application, particularly for a dynamic storage system.

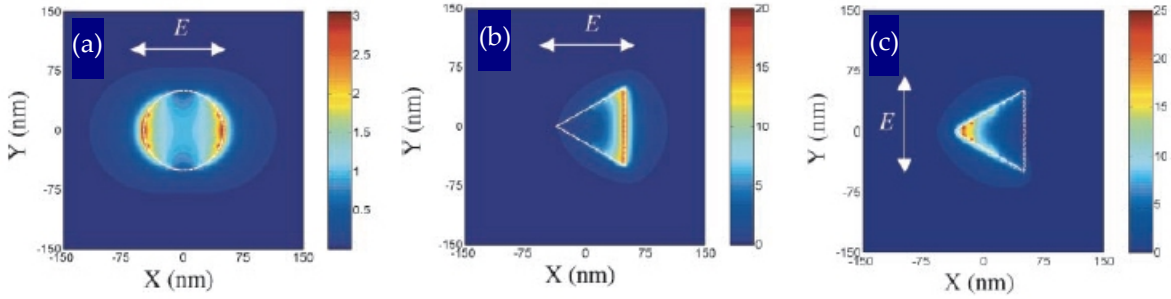


Fig. 2.4-4 Energy distribution of (a) a circular aperture with X-polarized incident light, and a triangular aperture with (b) X-polarized and (c) Y-polarized incident light. The aperture shape is outlined in each figure. Note the different color scaling in each figure. (K. Tanaka *et al.* Jpn. J. Appl. Phys. Vol 42, P. 1114, 2003)

2.4.3 C-shaped Aperture

The C-shaped aperture [20] is a novel design, which is modeled in a perfect conducting film with X-polarized incident light, as shown in Fig. 2.4-5.

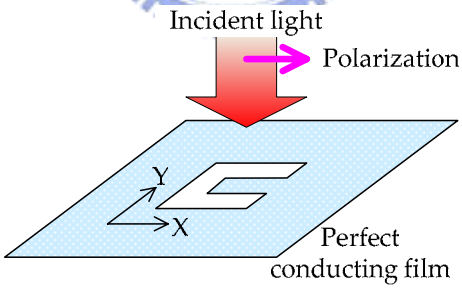


Fig. 2.4-5 Optical model of the C-shaped aperture (X. Shi *et al.* Opt. Lett. Vol. 28, P. 1320, 2003)

Compared with a 100*100 nm^2 square aperture, the power throughput of the C-aperture at the distance of 48 nm from the aperture is 1000 times stronger with a comparable near-field spot size, as displayed in Fig. 2.4-6. However, the incident light wavelength in this model is 1μm whereas most optical storage or detect systems are operated in the visible light regime where the natures of metals are

significantly different from perfect conductors. As a result, the behavior of such a C-aperture in visible light range is a pending problem to investigate.

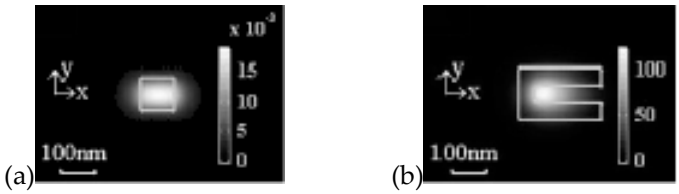


Fig. 2.4-6 Electric field intensity distribution at 48 nm from the two apertures: (a) square, (b) C apertures with their geometries outlined. (X. Shi *et al.* Opt. Lett. Vol. 28, P. 1320, 2003)

2.4.4. I-shaped (or H-shaped) Aperture

A more complicated aperture is I-shaped (or H-shaped) aperture as plotted Fig. 2.4-7. All the dimensions of the I-aperture and the metal film are expressed as the multiples of the incident wavelength λ [21].

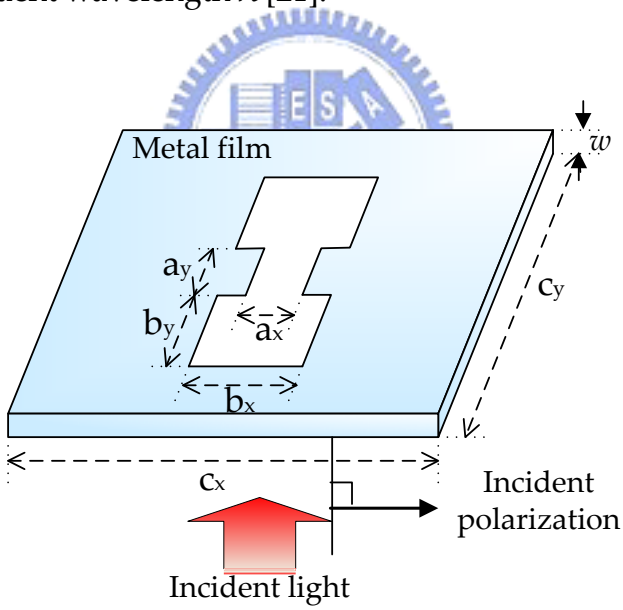
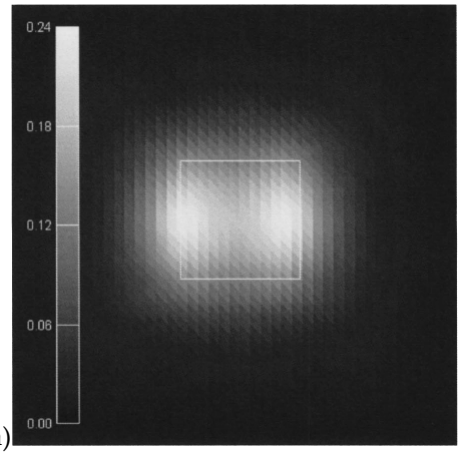


Fig. 2.4-7 Parameters of an I-shaped aperture (K. Tanaka *et al.*, Opt. Commun. Vol. 233, P. 232, 2004)

The comparison between a square aperture and an I-aperture, of which the gap area, ax*ay, is identical to the square aperture area, is displayed in Fig. 2.4-8, clearly showing the gap determines the transmitted power and the spot size principally. According to their simulation, the spot size is evidently reduced but the transmitted power is enhanced with a limited extent, around 6 times, which can compete with

the scatterer-formed and triangular apertures but is still inadequate for practical applications. Moreover, the I-aperture requires more complicated fabrication and will restrict its feasibility accordingly.

(a) 

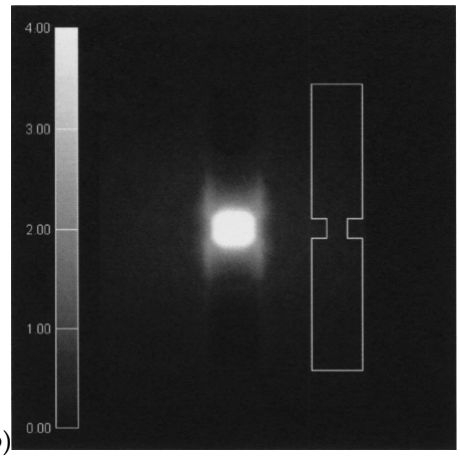
(b) 

Fig. 2.4-8 Electric energy distribution of (a) a square aperture of 0.064λ*0.064λ and (b) an I-aperture with gap area of 0.064λ*0.064λ, the film thickness of 0.224λ and the dielectric constant of the film material of -7.38-i7.18 (K. Tanaka *et al.* , Opt. Commu. 233, P. 240 2004)

2.5 Surface Corrugation

In 1998, Ebbesen *et al.* found that arrays of subwavelength apertures performed highly unusual zero-order transmission spectra [22], which is a sever challenge to Bethe's theory. In their configuration, a 200-nm-thick silver film was deposited on a quartz substrate and subsequently milled by focused-ion-beam (FIB) to fabricate aperture arrays. The distinct features of the transmission spectrum are the two phenomenal peaks taken place at two wavelengths; one is a little bit larger than the spacing between apertures and the other one corresponds to ten times aperture diameters. These two peaks show an evident overthrow of Bethe's formula that stated transmission is proportional to the quartic of the aperture size.

A further configuration attracting more attention is an aperture with surface corrugations [23], which allowed extraordinary transmission similar to aperture arrays as well. The structure was a free-standing Ni film of 300 nm thickness sandwiched in Ag layers, subsequently perforated an aperture and patterned grooves as the surface corrugation by utilizing FIB, as shown in Fig. 2.5-1.

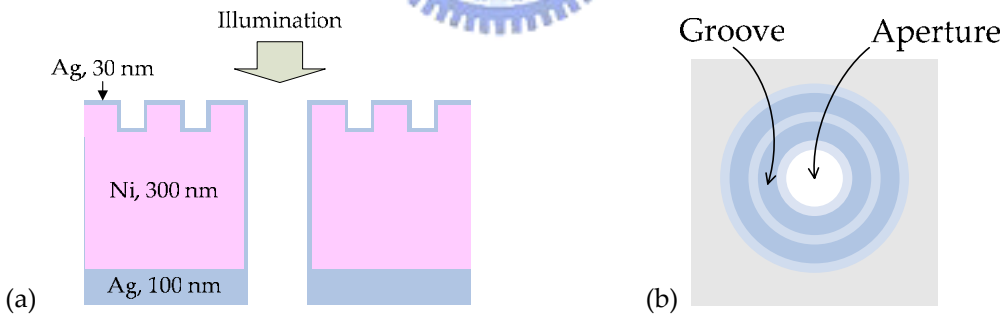


Fig. 2.5-1 The (a) side view and (b) top view of a single aperture surrounded by surface corrugations (T. Thio *et al*. Opt. Lett. Vol 26 P. 1972, 2001)

Transmission of this configuration reaches the maximum at the wavelength approximate to the pitch of the corrugations; in addition, the transmission efficiency (transmission normalized to the aperture area) at the peak can be up to 2.6, as shown in Fig. 2.5-2, indicating even those photons not impinging on the aperture are also captured and emitted through the aperture. This phenomenon results from the

resonance and coupling effect between the antisymmetric SPP modes and incident photons, which conduces the aperture acting as a novel probe of SPPs.

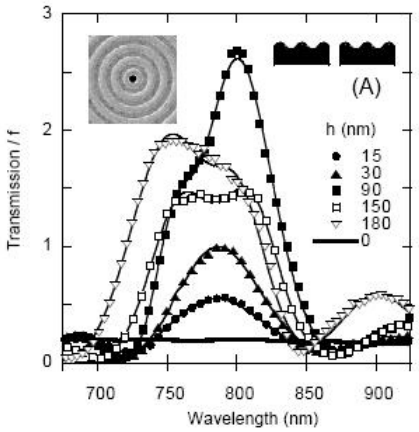


Fig. 2.5-2 Transmission spectra of a single aperture of diameter 440 nm surrounded by surface corrugation of pitch 750 nm with various corrugation depth h, the vertical axis is the ratio of transmission to aperture area (T. Thio *et al.*, Opt. Lett. Vol. 26, P. 1972, 2001)

Going a step further, if the exit side is also corrugated, the emitted light is highly directional [24], indicating the exit corrugation plays a role as a focusing grating [25]. A plot of the transmission intensity as a function of angle at the resonant wavelength of 660 nm displayed in Fig. 2.5-3 shows the divergent angle (defined by Full-Width Half-Maximum) is only 5°. Such a discovery is exhilarating because it implies the working distance between the pick up and the recording material can be enlarged.

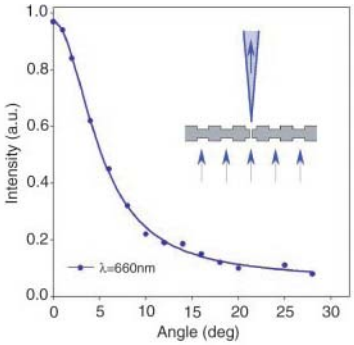


Fig. 2.5-3 Angular dependence of the transmission at the resonant wavelength of a double-corrugated configuration (H. J. Lezec *et al.*, Science Vol. 297, P. 820, 2003)

According to the aforementioned theory in section 2.2, the SPP modes can be excited on the exit surface as well so that the exit corrugations may not only function as a focusing grating but also contribute to the transmission enhancement by drawing supports from the coupling effect between SPP modes on opposite sides. In addition, the transmission process of this double-corrugated configuration can be divided into three independent procedures: coupling in, tunneling and coupling out [26]; therefore, the parameters of corrugations on the different sides can be optimized individually and enhancements can then be significant.

2.5 Summary

These literature surveys stimulate a thought of a special shaped aperture surrounded by the surface corrugations to come into being. Among those special shaped apertures, C-shaped aperture merits the favors because of the several orders of magnitude field enhancement. Consequently, optimizing C-aperture dimension to achieve satisfactory optical performance is the first step and the surface corrugations will be introduced sequentially for further field enhancement and spot size reduction.