Enhanced transmission through periodic hole arrays

Daniel Charles

Faculty advisor: Prof. David Tanner

Abstract

The enhanced transmission effect occurs in metals at certain wavelengths, where the transmission of light through a hole much smaller than the wavelength of light exceeds the expected value. A Zeiss MPM-800 microscope spectrometer was used to measure the reflectance and transmittance of three different periodic hole array samples. It was found that the positions of transmittance peaks and reflectance dips in these hole arrays can be explained by predictions from the theory of trapped modes, and the scaling of our results can be explained by both the surface plasmon theory and the theory of trapped modes.

Background

The samples that we have measured all have one element in common: "subwavelength holes," or holes smaller than the wavelength of the light used to analyze them. In 1944, H.A. Bethe [1] used Maxwell's equations and the appropriate boundary conditions to derive an equation for the "diffraction cross section" of a subwavelength circular hole:

$$A = \frac{64}{27\pi} k^4 a^6 (1 - \frac{3}{8}\sin^2\theta) \tag{1}$$

where k is the wavenumber of light ($k = 2\pi/\lambda$), a is the radius of the hole ($a = d/2$, where d is the diameter), and θ is the angle of incidence of the incoming light. The transmission T is equal to the diffraction cross section A divided by the hole's area ($T = A/\pi r^2$) [2]. By replacing every instance of the equation for the area of a circle with the equation for the area of a square ($A = D^2$, where D is the length of a side), we can rewrite Eq. (1) for a square hole. Making the appropriate substitutions and letting $\theta = 0$ (which is true for incident light), we obtain the equation for a square hole:

$$T \approx 18\left(\frac{D}{\lambda}\right)^4 \tag{2}$$

This equation shows that we would expect the transmission to decrease drastically as D becomes much less than λ. However, in 1998, T.W. Ebbesen [3] discovered that, for certain wavelengths, the amount of light transmitted through a periodic array of subwavelength holes was much higher than Eq. (2) predicts. Many others have since confirmed these findings. This greater-than-expected transmission of light through subwavelength holes has been called the "enhanced transmission effect." In some cases, the percentage of light transmitted actually exceeds the "open area fraction," or the percentage of the surface occupied by holes. This is interesting because, even if we use ray optics and disregard the effects of diffraction, the percentage of transmitted light should be no more than the open area fraction.

Several physical explanations have been proposed for the enhanced transmission effect. The original and most popular theory suggests that surface plasmons (which are collective oscil-

lations of the conduction electrons in a metal) are excited by incoming light, and then converted back into light at the site of a subwavelength hole. Many other theories have since been proposed, including CDEWs (composite diffracted evanescent waves), dynamical diffraction, and trapped modes. By measuring structures containing subwavelength holes, we can test the experimental predictions of these different theories, providing more insight on which theory best explains the enhanced transmission effect.

For all of the graphs in this paper, we will be using a unit convention commonly used in spectroscopy. Historically, spectroscopy plots were given with the independent variable as wavenumber (in units of cm^{-1}). This is not the same wavenumber as the one found in quantum mechanics ($k = 2\pi/\lambda$); for our purposes, wavenumbers are given as the inverse of the wavelength in units of centimeters. For example, to find the wavenumber of green light (500 nm), we would convert the wavelength to centimeters (500 nm = 500 $\times 10^{-7}$ cm) and take its inverse ($[500 \times 10^{-7}$ cm$]^{-1}$ = 20,000 cm^{-1}). More simply, we can convert between wavelength (λ, given in nm) and wavenumber ($\tilde{\nu}$, given in cm^{-1}) with the following equation:

$$\lambda\tilde{\nu} = 10^7 \tag{3}$$

Although it will not be used in this paper, there is another unit convention related to wavenumbers that is worth mentioning. We can relate the energy of a photon ($E = h\nu$) to the thermal energy and temperature of a material ($E = k_B T$). Setting the two equal and converting frequency to wavenumber, we get the following expression:

$$T = \frac{100hc\tilde{\nu}}{k_B} \approx 1.44\tilde{\nu} \tag{4}$$

where $h = 6.626 \times 10^{-34}$ J s (Planck's constant), $c = 3.00 \times 10^8$ m s^{-1} (the speed of light), $\tilde{\nu}$ is wavenumber (given in cm^{-1}), and $k_B = 1.381 \times 10^{-23}$ J K^{-1} (Boltzmann's constant).

Experimental details

All the measurements in this report were taken on a Zeiss MPM-800 microscope. This

microscope has the option to measure a sample's reflectance, transmittance, and photoluminescence. Reflectance and transmittance are simply the percentage of light reflected by or transmitted through a sample, respectively. The photoluminescence feature was not used for any of the data here.

When measuring reflectance on the microscope, light is emitted from a source and travels toward a beamsplitter. The light then shines down on the sample; some of this light is reflected, passes back through the beamsplitter, and travels to the detector. When measuring transmittance, light is emitted from the source and reflected by a mirror. The light then passes through a condensor and shines on the sample; some of this light is transmitted through the sample, and then proceeds through the same optical path as in reflectance. Note that the beamsplitter is removed for transmittance. (See Fig. 1.)

Two different light sources are used: a tungsten source for near infrared (NIR) measurements (λ = 600 nm to 2000 nm), and a xenon source for ultraviolet and visible (UV/VIS) measurements (λ = 200 nm to 800 nm). To isolate the wavelengths we wish to analyze, a monochromator is used for each source. A monochromator uses diffraction from a reflection grating to separate the

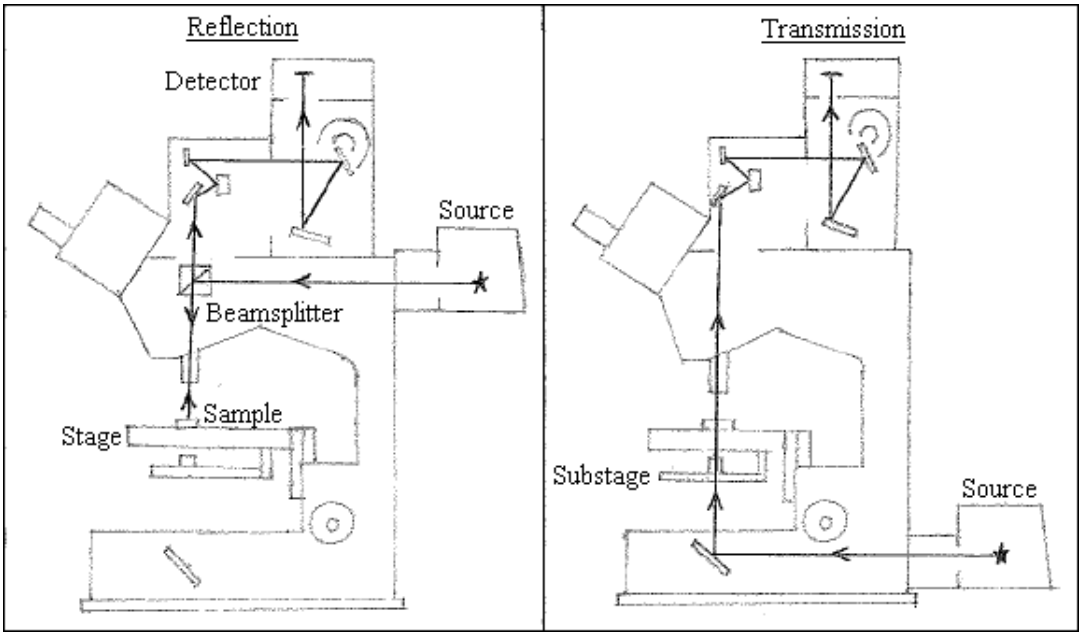

Figure 1: Simplified schematic of the Zeiss MPM-800 microscope used for our measurements. Based on Fig. 4-11 [4].

wavelengths of white light. For the tungsten source, the monochromator is located just before the detector; for the xenon source, the monochromator is just after the lamp.

The tungsten source has a variable voltage and can be adjusted from 1.5 V to 12 V. All our measurements were taken at approximately 8.00 V. The power supply for the tungsten source was somewhat unreliable; the voltage was sometimes perfectly stable for several hours, and sometimes fluctuated in a range as large as 0.5 V (for example, 7.60 V to 8.10 V). This instability may have been at least partially responsible for the noise observed in some of our measurements. Many of our measurements were taken as the average of multiple scans, so the noise should be minimal.

All the measurements taken with the microscope are relative. The microscope's sensitivity can be adjusted to the intensity of measured light for any set of conditions; therefore, intensities of individual measurements will always be arbitrary. In addition, the detector is not equally sensitive at all wavelengths, the monochromator may give some losses, and the source doesn't emit uniformly at all wavelengths, so any measurement will combine the actual properties of a sample with the properties of the detector, the monochromator, and the source. To find the actual reflectance or transmittance of a sample, then, we must compare the intensity of our sample with the intensity of some known standard. For reflection, our standard is an aluminum mirror of known reflectance. For transmission, our standard is simply empty space; we assume that 100% of all light should pass through empty space.

When calculating the absolute reflectance of a sample, the following formula is used:

$$R = \frac{O - P}{S - P} \cdot ref \tag{5}$$

where R is the absolute reflectance, O is the intensity of an object/sample, P is the intensity of parasitic light (that is, any stray light present in the system), S is the intensity of the standard, and ref is a correction factor for our standard. This correction factor is needed because our standards used for comparison are not perfect reflectors. For example, suppose that we have a highly reflective sample (for example, O = 0.98), a mirror that is 90% reflective (S = 0.90), and no stray light is

present ($P = 0$). If we used no correction, we would find that the absolute reflectance of our sample is over 108%, which is physically impossible. By multiplying by a correction factor equal to the absolute reflectance of the mirror ($ref = 0.90$), the correct value of 98% reflectance is obtained.

To calculate absolute transmittance, the following simpler formula is used:

$$T = \frac{O}{S} \tag{6}$$

The two differences between Eqs. (5) and (6) are that there is no parasitic light (P), and there is no reference correction (ref). There is no parasitic light because a beamsplitter is not used in transmission measurements, and there are no optical paths that repeat themselves. There is no reference correction because we are using empty space as a standard, and empty space transmits 100% of light.

Thin film mirror

Our first measured sample, which will be referred to as a thin film mirror, is a small cylin-

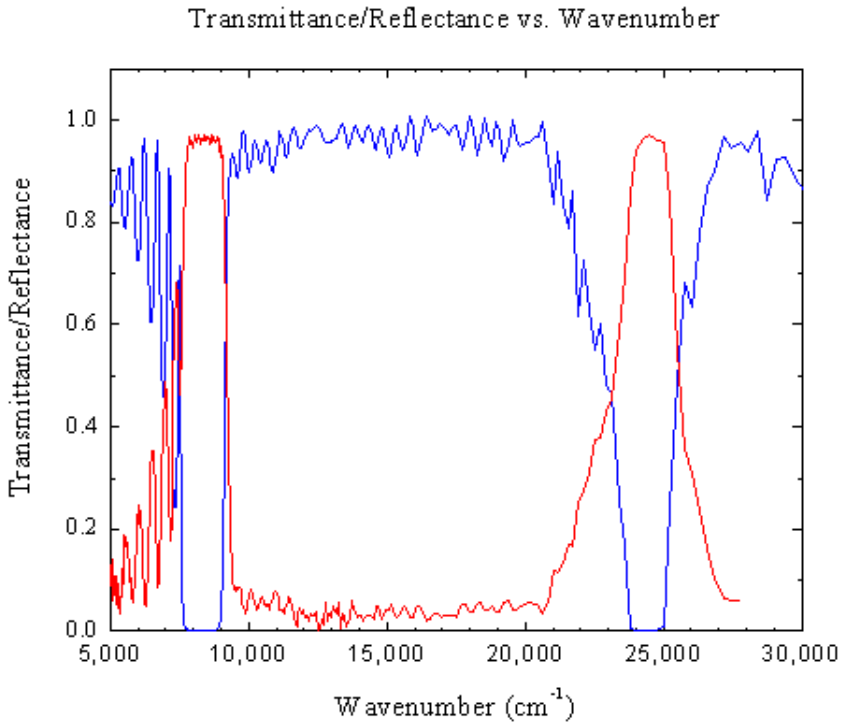

Figure 2: Transmittance and reflectance versus wavenumber for the multilayer thin film mirror sample. The blue curve is transmittance and the red curve is reflectance.

drical piece of substrate with a multilayered thin film deposited on one face. This film was known to be nearly 100% reflective in part of the NIR ($\sim$ 7,000-8,000 cm^{-1}) and was used to confirm that the microscope and the standard mirrors were working correctly. Transmittance measurements (see blue curve, Fig. 2) confirmed that the sample is highly reflective in that region of the NIR, as well as in part of the visible region ($\sim$ 24,000-25,000 cm^{-1}). Each dip in transmittance should (and does) correspond to a peak in reflectance, assuming absorption of light by the sample is negligible. The small fringes on each side of the large transmittance dips are not noise, but are due to interference between reflections from the multiple layers of the thin film.

The original reflection measurements with our first aluminum mirror, however, showed that the peak in the UV/VIS was around 120%, which is physically impossible. This suggests that the aluminum mirror was not performing ideally. Films of aluminum are known to react with oxygen and form a layer of aluminum oxide, which would reduce the aluminum's reflectance over time. For example, consider again the previous example. Suppose our thin film mirror is 98% reflective (S = 0.98), no stray light is present (P = 0), but our imperfect aluminum mirror now reflects at only 80% (O = 0.80). Using a correction factor of ref = 0.90, our final result will be R = 110%. This is similar to what we believe occurred in our measurements to give us unreasonable results.

In an attempt to correct this problem, we decided to create new aluminum mirrors using two different mirror fabrication processes. The first batch was made with argon ion etching, and the second was made with sputtering. To begin, two microscope slides are cut into six small rectangles per slide. These pieces are cleaned with acetone, cleaned again with methanol, and dried with compressed air.

For argon ion etching, the newly cleaned pieces are glued onto a platform with rubber cement and inserted into a Veeco etching machine. This machine accelerates argon ions toward a sheet of aluminum foil, spraying particles of aluminum towards the pieces of glass. For sputtering, a tightly wound coil of tungsten heats up and melts pieces of aluminum. The aluminum begins to evaporate, and the vapors are deposited onto the pieces of glass. In both cases, when the film of aluminum reaches the desired thickness, the mirrors are completed.

The reflection measurements of the thin film mirror were repeated. Using our first batch of mirrors (made with argon ion etching), the peak in reflectance in the UV/VIS was still well above 100%. Further examination of this batch of mirrors showed that they were actually less reflective than the older aluminum mirror that was used previously. This was most likely due to some flaw in the fabrication process.

Our second batch of aluminum mirrors (made with sputtering) proved to be much more successful than the first. Measurements of reflection confirmed that these mirrors were more highly reflective than both the old mirror and the mirrors from the first batch. The red curve in Fig. 2 shows reflectance results for the thin film mirror, using a mirror from the second batch as the standard. As we expect, the peaks in reflectance lie just below 100% in both the UV/VIS and the NIR. This confirms that our reflectance results will be reliable using these new mirrors.

Hole arrays

Our second sample, which is the main focus of this project, is a silver mirror with three 100 μm x 100 μm periodic arrays of subwavelength holes etched into it. Each hole array has square holes of different sizes (edge length a) with different spacings between them (D_g) (see Fig. 3). For sample A, a = 379 nm and D_g = 1200 nm; for sample B, a = 316 nm and D_g = 1000 nm; and for sample C, a = 253 nm and D_g = 800 nm. Figs. 4 and 5 show the results of our NIR reflection and transmission measurements of this sample.

These measurements used the new aluminum mirror as a standard. This result shows us that there are several dips in reflectance for each sample. In general, these dips in reflectance should correspond to peaks in transmittance. We see that many of the same peaks from the reflectance curve are present as dips in the transmittance curve.

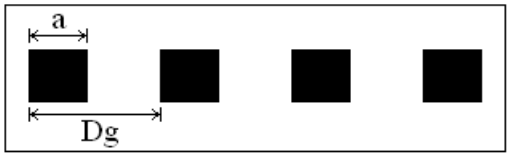

Figure 3: A drawing showing the hole size a and spacing D_g for a periodic hole array.

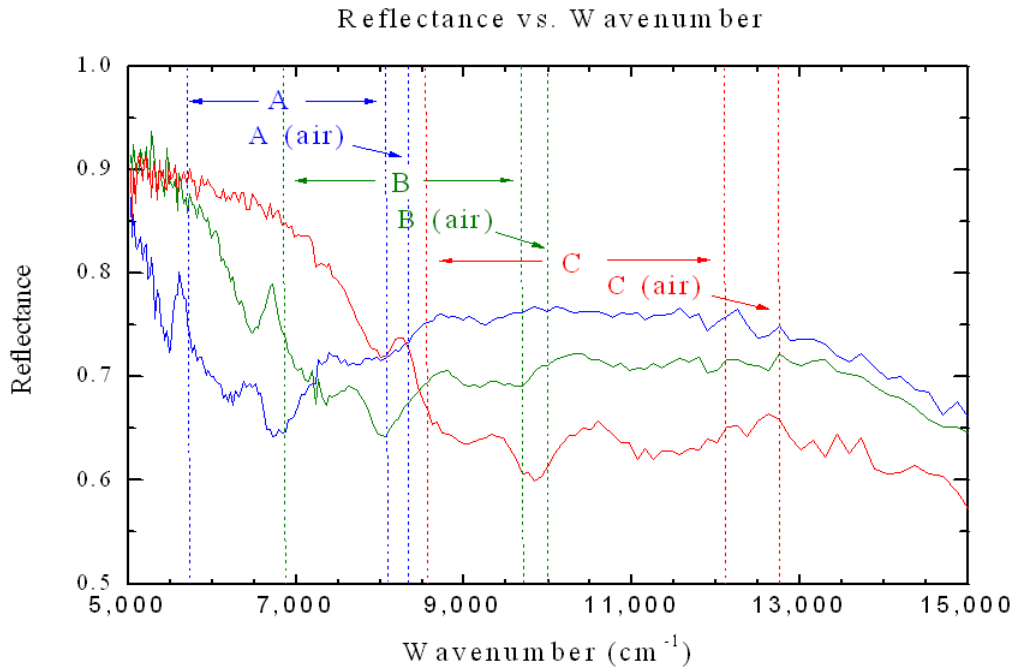

Figure 4: Reflectance versus wavenumber in the NIR region for the periodic hole array sample, for each of the three hole arrays. The blue curve is for sample A, green is for sample B, and red is for sample C. The dashed lines indicate the locations of diffraction thresholds for the substrate (labeled with bold letters) and for air (labeled in parentheses).

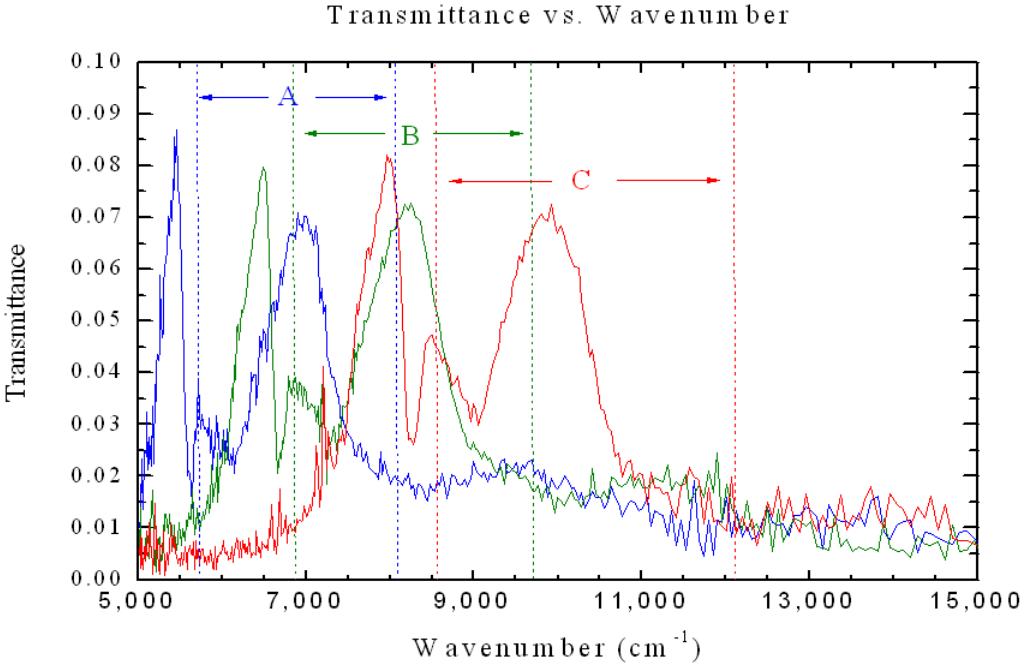

Figure 5: Transmittance versus wavenumber in the NIR region for the periodic hole array sample. Dashed lines indicate the locations of diffraction thresholds for the substrate.

It can be shown through diffraction theory [5] that transmission peaks from a metal-dielectric interface occur after the "diffraction threshold wavelength" λ_s given by the following equation:

$$\lambda_s = \frac{n_d D_g}{l_s} \tag{7}$$

where $l_s = \sqrt{l_x^2 + l_y^2}$, l_x and l_y are integers, n_d is the index of refraction of the dielectric, and D_g is the spacing between holes. The theory of trapped modes suggests that this concept can be used to explain enhanced transmission.

For our sample, there are two metal-dielectric interfaces: the silver-substrate interface and the silver-air interface. The refractive index n_d of the substrate is 1.46, and n_d of air is 1. Diffraction thresholds were calculated using Eq. (7) for both the substrate and air. Figs. 4 and 5 include dashed lines that indicate the positions of the diffraction thresholds due to the substrate; Fig. 4 also includes a few dashed lines indicating thresholds due to air.

These diffraction thresholds seem to predict the expected reflectance dips and transmittance peaks fairly well. Transmittance peaks (or reflectance dips) occur at wavelengths higher than the diffraction threshold wavelengths; if our plots used wavelength as the independent variable, the peaks (or dips) would lie to the right of the thresholds. However, our plots use wavenumber as the independent variable, so the peaks (or dips) will lie to the left of the thresholds. In Fig. 4, all six thresholds have a reflectance dip to their left; in Fig. 5, all thresholds have a transmittance peak to their left. Also, note that some of the smaller reflectance dips in Fig. 4 are explained by diffraction thresholds due to air.

Our periodic hole array sample also exhibits scaling properties. When we divide the wavenumber values of each curve by that sample's threshold wavelength for $l_s = 1$ (converted to wavenumbers), we find that all three curves have dips and peaks at roughly the same positions. (See Figs. 6 and 7.) This suggests that the most important feature of these curves is the diffraction threshold; by scaling these curves to the threshold (which depends only on the spacing between the holes and the index of refraction of the substrate), the curves take the same shape.

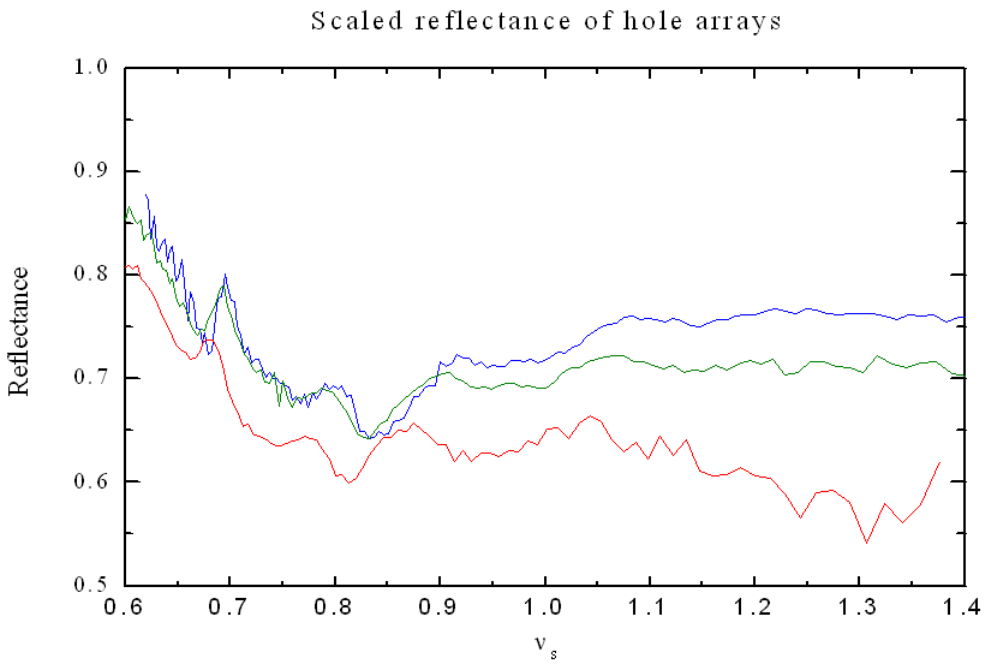

Figure 6: Scaled reflectance of the periodic hole array sample. The x-axis is given in wavenumbers divided by the diffraction threshold for $l_s = 1$.

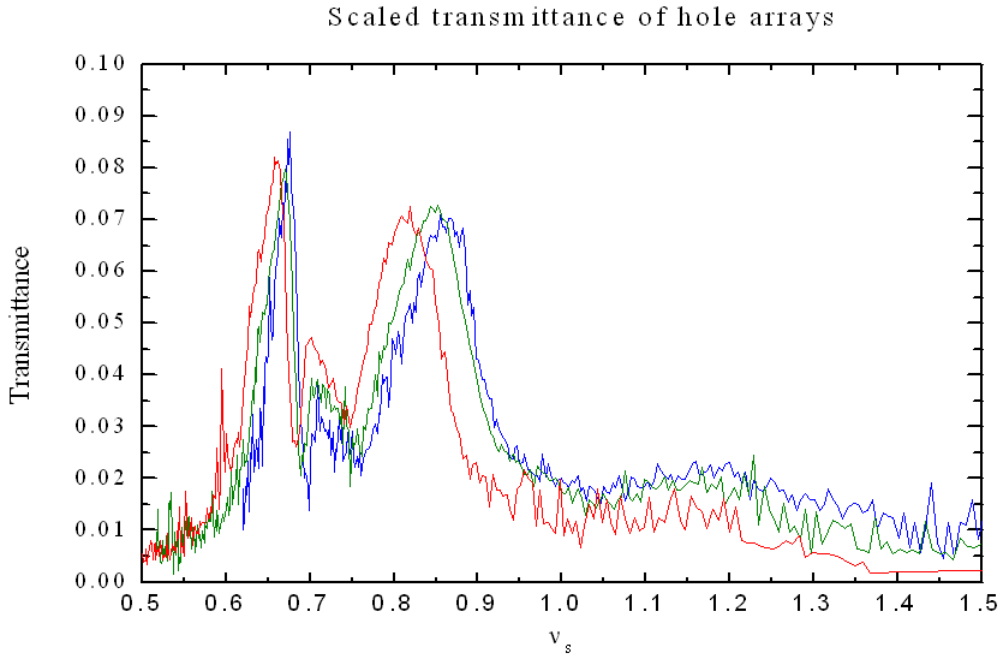

Figure 7: Scaled transmittance of the periodic hole array sample. The x-axis is given in wavenumbers divided by the diffraction threshold for $l_s = 1$.

Note that the scaling here is not perfect; there is some shifting to the left for smaller hole spacings. The surface plasmon theory predicts that surface plasmon excitations will occur at certain wavelengths λ_{sp} given by the following formula:

$$\lambda_{sp} = D_g \sqrt{\frac{\epsilon_d \epsilon_m}{\epsilon_d + \epsilon_m}} \tag{8}$$

where ϵ_d and ϵ_m are the permittivities of the dielectric and metal, respectively, and ϵ_m is defined as $\epsilon_m = \epsilon_\infty - \omega_p^2/\omega^2$. Using these equations, we can calculate the expected positions of the excitations from the surface plasmon theory and see whether they predict that this shifting should occur. Fig. 8 shows a plot comparing the observed peak positions (left) and the calculated excitations (right). The y-axis is the product of the hole spacing D_g and the wavenumber $\tilde{\nu}$ of the peak, and the x-axis is the hole spacing D_g. The actual and calculated curves seem to agree; both curves are relatively straight lines with an upward trend for higher D_g. Therefore, the surface plasmon theory seems to explain this shifting in our scaled curves.

Conclusions

In summary, we have shown that many of the reflectance dips and transmittance peaks of our periodic hole array samples relate well to the diffraction threshold wavelengths proposed by

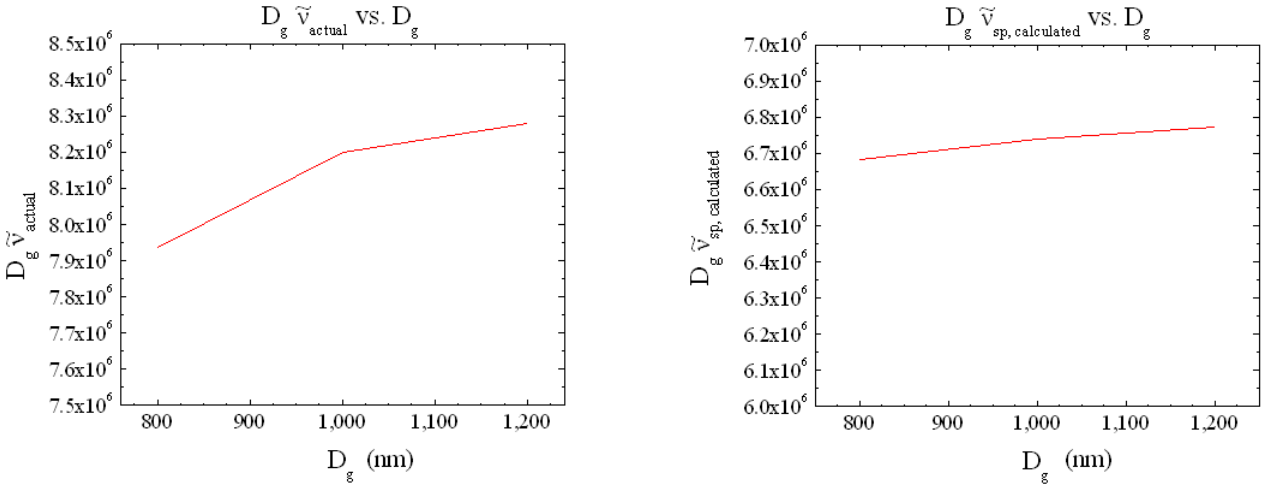

Figure 8: Actual (left) and calculated (right) plots of hole spacing D_g times wavenumber $\tilde{\nu}$ versus hole spacing D_g. This shows that the surface plasmon theory predicts the shifting in the scaled plots.

the theory of trapped modes. The scaling of our results is also predicted by the theory of trapped modes; however, the surface plasmon theory predicts this as well, and goes on to predict that the scaling will shift for different hole spacings. Our hole array results provide further evidence in support of both the surface plasmon theory and the theory of trapped modes; however, we have not narrowed it down to one, and it will likely be a matter of years until a concensus is reached on the true cause of enhanced transmission.

Acknowledgments

I would like to thank Dr. Tanner for being my advisor for this project. I am very grateful to the graduate students and the postdoctorate student that I worked with for their help and advice: Dimitrios Koukis, Xiaoxiang Xi, Kevin Miller, Zahra Nasrollahi, and Jungseek Hwang. Finally, I would like to thank the National Science Foundation for funding my stay at the University of Florida.

References

1. H.A. Bethe, Phys. Rev. **66**, 163 (1944).

2. K. Woo, Ph.D. thesis, University of Florida, 2006.

3. T.W. Ebbesen et al., Nature **391**, 667 (1998).

4. M. Nikolou, Ph.D. thesis, University of Florida, 2005.

5. S. Selcuk et al., Phys. Rev. Let. **97**, 067403 (2006).