Numerical method for high accuracy index of refraction estimation for spectro-angular surface plasmon resonance systems

Colin J. Alleyne,1 Andrew G. Kirk,1 Wei-Yin Chien and Paul G. Charette2

1Photonics Systems Group, McGill University
853 Sherbrooke St., Montréal, Québec, H3A 2T6, Canada
2Dept. de génie électrique et de génie informatique, Université de Sherbrooke
2500 Boul. de l’Université, Sherbrooke, Québec, J1K 2R1, Canada
*Corresponding author: Paul.Charette@USherbrooke.ca

Abstract: An eigenvector analysis based algorithm is presented for estimating refractive index changes from 2-D reflectance/dispersion images obtained with spectro-angular surface plasmon resonance systems. High resolution over a large dynamic range can be achieved simultaneously. The method performs well in simulations with noisy data maintaining an error of less than 10^-8 refractive index units with up to six bits of noise on 16bit quantized image data. Experimental measurements show that the method results in a much higher signal to noise ratio than the standard 1-D weighted centroid dip finding algorithm.

©2008 Optical Society of America

OCIS codes: (240.6680) Surface plasmons; (130.6010) Sensors; (050.1950) Diffraction gratings; (260.3910) Metal Optics.

References and links

#101454 - $15.00 USD Received 11 Sep 2008; revised 2 Nov 2008; accepted 5 Nov 2008; published 10 Nov 2008 (C) 2008 OSA 24 November 2008 / Vol. 16, No. 24 / OPTICS EXPRESS 19493
1. Introduction

Surface plasmon resonance (SPR) has many applications in biosensing, drug development, food, safety, environmental monitoring and medical diagnostics [1]. In these applications, SPR is used to measure quantities such as analyte concentration, affinity, kinetics and specificity, with a constant push towards increasingly accurate measurements [2,3]. Indeed, though SPR has the advantage that it does not require labeling of analytes, current SPR systems are not as sensitive as fluorescence-based systems in the case of small molecules [4,5].

A surface plasmon is an electron density wave that travels at a metal-dielectric interface such that the oscillating electrons in the metal have an associated electromagnetic field that is confined to the interface [6]. By tracking changes in the resonance conditions that excite surface plasmons (i.e. momentum and energy matching), changes in the refractive index of the dielectric sensing medium can be detected. Most existing SPR sensing systems are based on varying the angle of incidence (momentum) at a fixed excitation wavelength (energy) in order to track changes in the reflectance curve. Various algorithms have been published for identifying the minimum in the reflectance curve corresponding to the angle of resonance [7].

Currently, the highest accuracy SPR systems track optical phase changes at resonance (rather than reflectance): by combining SPR with interferometry, a resolution in index of refraction change of 4x10^-8 RIU (refractive index units) has been demonstrated [8]. Phase-based methods, however, suffer from very limited dynamic range since the optical phase changes occur only in the immediate vicinity of resonance. In contrast, the more common reflectance-based systems typically achieve an order of magnitude less in resolution, on the order of 10^-7 RIU, but with much greater dynamic range [3]. This work demonstrates a numerical technique for use with reflectance-based systems that is capable of achieving a measurement resolution (5x10^-9 RIU) superior to phased-based systems while retaining the large measurement dynamic range typical of reflectance-based systems. This is accomplished using linear algebra methods to analyze the 2-D reflectance/dispersion images of the surface plasmon resonance response, as outlined in Section 2.

The numerical analysis method described here was designed for use with spectro-angular SPR systems, where both angle and wavelength reflectance information are acquired simultaneously. The output of such a system is a 2-D image of the surface plasmon dip as opposed to the familiar 1-D curve containing the resonance dip (See Fig. 1). This multi-modal approach in SPR techniques, where multiple types of information are gathered simultaneously, is becoming more prominent in the literature [9,10]. Spectro-angular SPR configurations have been used to image the surface plasmon dispersion curve [11-13], indirectly in a biosensing application [14], and proposed for biosensing [15]. Though a number of papers in the literature have proposed methods for analyzing 1D SPR curves [16], to our knowledge, little work has been devoted to the analysis of multi-modal SPR results, such as that obtained with spectro-angular systems. The work presented here seeks to address this need.
The recent advent of surface plasmon photonic bandgap (SP-PBG) systems [17] provides interesting possibilities for new developments in SPR sensing, and for spectro-angular SPR systems in particular. A SP-PBG system is one where a surface grating is used to perturb the SP such that it cannot propagate on the interface due to destructive interference. A 1-D bandgap, for which SP propagation is forbidden in one direction, is considered here. It is generated using a 1-D grating since SP propagation is forbidden only along the grating vector. To generate a bandgap in the reflectivity image, the metal film on the sensing surface is patterned with periodic features having dimensions less than the wavelength of the incident light [18] (see “Grating” on the gold sensing surface in Fig. 2). In the case of small amplitude periodic corrugations, the Bragg period ($\Lambda_B$) is related to the surface plasmon wavelength ($\lambda_{sp}$) and free-space wavelength ($\lambda_o$) as follows:
where $\varepsilon_D$ is the dielectric constant of the dielectric and $\varepsilon_{mr}$ is the real part of the dielectric constant of the metal. The bandgap resulting from such a corrugated metal surface can be seen in Fig. 3, which compares 2-D reflectance/dispersion images for both smooth and patterned metal (bandgap) surfaces, generated with rigorous coupled wave analysis (RCWA).

SP-PBG systems generate 2-D reflectance/dispersion images with increased information density and “image texture” that we hypothesize can be used to improve measurement accuracy given the 2-D analysis that is presented here. In addition, recent work on SP-PBG systems based on corrugated metal surfaces suggests that such systems increase SPR sensitivity in the neighborhood of the bandgap; a phenomenon that may also benefit the 2-D analysis [19,20].

![Fig. 3. RCWA simulated 2-D reflectance/dispersion image from a spectro-angular SPR system having either a smooth 50nm thick gold surface (a) or a corrugated (period 300nm) gold surface (b). The structure for (b) is shown in (c).](image)

2. Method for refractive index estimation

The goal of this work was to develop a numerical method to estimate refractive index changes with spectro-angular SPR systems. Its basic principle is the creation of a set of orthogonal basis vectors, $b_j$, from a set of reflectance/dispersion images which span the refractive index range of interest. This basis, the ensemble of orthogonal vectors $b_j$, is thus used to model the reflectance/dispersion image data over the refractive index domain of interest. In order to estimate the effective index of refraction corresponding to a particular reflectance/dispersion image, the “inverse problem” is solved by projecting the image data against the basis using standard linear algebra techniques, as explained below.

The basis set, $b_j$, is calculated by performing an eigenvector analysis on a training set of reflectance/dispersion images corresponding to known refractive indices, $n_i$ ($i = 1..N$), spanning the range of refractive indices of interest, $n_{min} < n_i < n_{max}$. We chose to use singular value decomposition (SVD) for the eigenvector analysis [21]. As there is no numerical advantage to processing the image data in 2D, the linear algebra operations outlined below are performed in 1D. Therefore, the eigenvector analysis is performed on the set of normalized 1D vectors, $a_i$ ($i = 1..N$), formed by converting the training set of reflectance/dispersion images to vector form by concatenating the images row by row, yielding the basis vectors, $b_j$. Normalization of the input vectors ($a_i/||a_i||$) ensures that the method is insensitive to overall intensity fluctuations in the data.

Since the basis vectors are orthogonal, the following applies:

$$b_i \cdot b_j = 0, \text{ if } i \neq j$$

$$\neq 0, \text{ if } i = j$$

Once the basis vectors, $b_j$, are obtained, the training set vectors, $a_i$, are projected against the basis to form the matrix of weights, $A$, using the inner product operator:
\[ A_j = a_i \cdot b_j \quad (3) \]

Accordingly, the original reflectance/dispersion training set vectors, \( a_i \), can be reconstructed from linear combinations of the basis vectors, \( b_j \), according to the weights, \( A_{ij} \):

\[ a_i = \sum_{j=1..N} A_{ij} b_j \quad (4) \]

The “inverse problem”, by which the refractive index of a candidate reflectance/dispersion image, \( n_c \), is calculated, proceeds in two steps. First, the candidate reflectance/dispersion image is converted to vector form, \( v \), by row concatenation and then projected against the basis to form the weight vector, \( w \):

\[ w_j = v \cdot b_j \quad (5) \]

Note that the weight vector, \( w \), is similar to a row of the weight matrix \( A \) except that it corresponds to an unknown candidate image, rather than to one of the training set images. Secondly, the weight vector, \( w \), is projected against each row of the weight matrix \( A \) to obtain the solution vector \( s \):

\[ s = A w \quad (6) \]

whose elements are the result of the inner product between the candidate reflectance/dispersion weight vector, \( w \), and each of the weight vectors corresponding to the reflectance/dispersion images in the training set. The elements of \( s \) thus correspond to the relative degree of similarity between the candidate reflectance/dispersion image and each of the images in the original training set. Furthermore, the elements of \( s \) can be considered as sampled values of a continuous objective function to be maximized over the domain \([n_{min}, n_{max}]\). The abscissa of the interpolated maxima will thus correspond to the desired unknown index of refraction, \( n_c \).

The more experimentally relevant difference of refractive indices corresponding to a change of state in the dielectric sensing medium is obtained by subtracting refractive index estimates calculated from successive candidate reflectance/dispersion images. If desired, the dimensionality of the problem can be reduced by retaining only a subset of most significant

---

Fig. 4. Summary of refractive index estimation method.
vectors in the basis, \( \mathbf{b}_k \ (k < N) \), as in the method of principal component analysis (PCA). A summary of the process is shown in Fig. 4.

3. Results

In order to first validate the method using numerical simulations, a basis was constructed from a training set of reflectance/dispersion images at equi-spaced refractive indices over the range \( n_{\text{min}}=1.3 \) to \( n_{\text{max}}=1.38 \) generated using RCWA. A set of candidate test reflectance/dispersion images, at index values between that of the training set images, were also generated using RCWA. The algorithm outlined above was then used to calculate the refractive index differences between successive test candidate reflectance/dispersion images. The test candidate images were treated as unknowns when input to the algorithm and the error between the estimated refractive index differences and the known values constitutes the estimation error.

Figure 5 shows estimation error results, \( \varepsilon_{\text{rms}} \), for a known index difference of \( 1\times10^{-6} \) RIU across bases of varying sizes, \( N \), corresponding to training sets of varying index of refraction intervals, \( \Delta n_B \). A range of training set interval lengths was explored to determine the optimal value for accuracy and efficiency of computation. The interval of \( \Delta n_B = 0.002 \) RIU \( (N = 40) \) was found to be optimal: a smaller interval did not improve performance and only served to increase computation time. As seen in Fig. 5, for a basis set built from 40 images \( (\Delta n_B = 0.002) \), the \( \text{rms} \) error in the refractive index difference estimation is \( 5\times10^{-9} \)RIU (calculated over the range from \( n=1.305 \) to 1.375).

The wide dynamic range achievable using the proposed method can be easily seen from Fig. 5. There is a slight edge effect since performance degrades for \( n<1.305 \) and \( n>1.375 \), though it is not of great significance since the training set can be extended to cover a slightly larger index range if required. In cases of narrower refractive index ranges (i.e. a smaller dynamic range), the same accuracy can be achieved with a smaller basis set while maintaining the same index increment in the training set (see Fig. 6).

The robustness of the method to variations in experimental parameters was numerically evaluated with respect to random intensity additive noise, CCD camera spatial resolution and...
intensity level resolution (A/D quantization noise). Gaussian noise was added to the simulated images in such a way that the standard deviation of the pixel noise, $\sigma_{Nb}$, was equivalent to a given number of bits, $N_b=1,2,\ldots,n$, where $n$ is the number of bits in the CCD camera A/D:

$$\sigma_{Nb} = 2^{N_b} - 1 / 2^n - 1$$  

The $rms$ error in the refractive index difference estimation was then determined as a function of this bit noise. The noise injection process was repeated for all training images used to form the basis set, as well as for the set of candidate test images. The resulting $rms$ error in the refractive index difference estimation for each additive noise level is shown in Fig. 7, demonstrating that the method yields an error of less than $1x10^{-8}$ RIU for up to four bits of noise. It should be noted that this result was obtained without applying any of the standard noise filtering techniques that are typically used in SPR measurements (such as smoothing the curve).

Figure 7 shows results with respect to both the CCD spatial resolution and intensity level quantization. The upper set of curves is for an 8-bit CCD while the lower set is for a 16-bit CCD. Each set consists of three separate curves for pixel resolutions of 256x256, 512x512 and 1024x1024. While it is clear that high resolution CCDs outperform low resolution CCD’s, the gain in performance is modest. A more significant difference is the A/D quantization, where 16bit CCDs outperform 8bit devices by more than an order of magnitude. An 8bit CCD would only be capable of resolving an index difference of $\Delta n=10^{-6}$ RIU if the image noise was low ($N_b<2$) and the spatial resolution was high (1024 x 1024).
Fig. 7. Performance of the algorithm as a function of the number of bits, $N_b$, of additive Gaussian noise (Eq. 7) in the reflectance/dispersion images, calculated for a series of candidate test images over index values ranging from 1.305 to 1.375 and using $\Delta n_i=0.002$ RIU to build the basis ($N = 40$). Results are shown for both 8-bit and 16-bit cameras, at three different resolutions (256x256, 512x512, and 1024x1024). Note the graph labeled “Sum of squared diffs” which indicates the best-case performance (1024x1024 16-bit camera) of a variation of the algorithm that uses a simple sum of squared differences as the error measure between training set and candidate test images for building the solution vector $s$, rather than the full method involving the projection against the basis.

Finally, the graph labeled “Sum of squared diffs” indicates the best-case performance (1024x1024 16-bit camera) of a variation of the algorithm that uses a simple sum of squared differences as the error measure between training set and candidate test images for building the solution vector $s$, rather than the full method involving projections against the basis. The use of a less robust error measure such as the ubiquitous sum of squared differences thus downgrades the performance of a 16-bit camera to that of an 8-bit equivalent device. The comparative robustness of the basis projection method of building the solution vector, $s$, is expected to be even more pronounced for experimental data (compared to simulations) which will contain additional noise and image distortion.

A fundamental question impacting the performance of the proposed method involves the source of reflectance/dispersion training images used to create the basis. The natural inclination would be to use an actual spectro-angular SPR system with calibrated refractive index solutions to step through a range of refractive indices and capture the corresponding reflectance/dispersion images. This hypothesis was tested using numerical simulations by adding noise and random index errors ($\Delta n_i$) to a set of training images generated with RCWA, where the resulting indices, ($n_i\pm\Delta n_i$), have an average uncertainty, $\bar{\Delta n_i}$, similar to that of actual calibrated solutions due to solution purity limitations and temperature variations. Our results (not shown) clearly indicate that this random error would propagate through the calculations and limit the accuracy of the method to the same order of magnitude as the average refractive index uncertainty of the training set. Furthermore, the number of calibrated solution samples needed would depend on the desired dynamic range and could rapidly become unmanageable.

Perhaps surprisingly, a more accurate and practical choice for the training set image data is simulated data calculated by modeling the structure of the interface as closely as possible (this method is used throughout this paper). The advantage of this method lies in the fact that,
although the simulated reflectance/dispersion images will not exactly match that obtained from the actual system (due to the difficulty to perfectly model the experimental setup), the errors incurred due to this mismatch are systematic and are largely eliminated when taking differential measurements such as changes in refractive index.

A further advantage of this method is that the dynamic range can be arbitrarily extended simply by generating new simulations a posteriori (as opposed to measuring new training index samples under the exact same operational conditions). The resulting experimental error on index of refraction difference estimations using this method (simulated training images to form a basis), however, will depend on the cumulative effect of the differences between actual and modeled parameters (metal/dielectric interface, optics, etc.) and can only be truly assessed experimentally.

3. Bandgap images

In light of our experience with photonic bandgap structures [14], we hypothesized that adding additional “information density” to the reflectance/dispersion images would further improve the discrimination of candidate test images using our proposed method. For this purpose, gratings with a 10nm height and 300nm period on the sensing gold surfaces were simulated to create a bandgap in the reflectance/dispersion images, as shown in Fig. 3.

Surprisingly, however, our results (not shown) indicate that, for small index changes (Δn<1x10⁻⁶), the bandgap images conveyed no special advantage in terms of performance. There was a slight advantage when measuring a much larger index difference of 1x10⁻³RIU; however, the improvement was small, and more importantly, large index changes are not interesting since they are not difficult to measure with other, simpler, methods.

4. Experimental results

To validate the method, an experimental spectro-angular SPR system was built to monitor index changes in a liquid dielectric medium using a basis set formed from simulated training set images, as explained above. The setup monitored water and three salt solutions of different molarities (0.2M, 0.6M and 1M) in contact with a flat gold surface. The salt solutions were injected into a chamber (1.2cm³ volume) abutting the gold surface and images were taken every 60 seconds for 60 minutes. Each of the captured images was calculated from an average of ten frames taken with the CCD camera (480x640 pixels, 8bits/pixel). The background light was eliminated by subtracting a dark image taken without the source. The light source for this experiment was a 2mW super luminescent diode with a 42.5nm bandwidth centered at 850nm. The focal lengths of the cylindrical lenses were 19mm. An example of simulated and experimental candidate test images is shown in Fig. 8. The widening of the dip at the top and bottom of the image in Fig. 8(b) is due to the lack of source power at those wavelengths. The tilt of the dip is due to a slight rotational misalignment of optical components in the light path.

Normally, all reflectance/dispersion images would be normalized using the orthogonal polarization, as in standard SPR. With this particular experimental setup, however, fringes in the images (Fig. 8(b)) caused by aberrations in the large numerical aperture cylindrical singlet lenses (CL1 and CL2 in Fig. 2) forced us to use an alternative. Indeed the fringes in the orthogonal polarization images were out of phase by λ/2 and thus amplified the oscillations when used for normalization. Instead, we used the same polarization image acquired with air in the chamber, whose fringes were not shifted but of different magnitude, for normalization. In future experimental setups, using higher quality lenses, normalization will proceed in the usual way by using the orthogonal polarization reflectance/dispersion images.

The index change tracking in the experimental images was estimated using both the proposed 2-D method and a standard 1-D weighted-centroid dip finding algorithm performed on horizontal slices of the same data [22]. The results in Fig. 9 show the index change steps brought about by the salt solutions (starting with water) for both the proposed method (Fig. 9(a)) and the dip finding algorithm (Fig. 9(b)). Each solution was left in the chamber for approximately 15 minutes before a new solution was added.
To compare the performance of each method, signal to noise ratios (SNR) were calculated at each step in refractive index. Note that in the case of the 1-D weighted-centroid dip finding algorithm, the SNR for a particular index step was averaged over three slices obtained from the corresponding dispersion/dispersion image. The following expression was used to calculate a step-specific SNR using the mean index value of each step ($\bar{x}$) as the signal and the standard deviation of the values within each step ($\sigma$) as a measure of the noise:

$$SNR = 10 \log_{10} \left( \frac{\bar{x}}{\sigma} \right)$$

Figure 9 clearly shows that the proposed method has a far superior performance: the SNR averaged over the four index steps is 42.4dB for our proposed method compared to 15dB for the standard 1-D method, an almost 3 orders of magnitude improvement. The linearity of the response was good for both methods: the proposed method had a linear regression fit (coefficient of determination) of $R^2=0.9942$ for the response verses molarity curve while the 1-D method had $R^2=0.9917$.

The smallest index change measurable with this particular experimental setup can be estimated by assuming that a change in the order of one standard deviation of the noise can be determined: the average standard deviation of the noise across index change steps for the proposed method was $2.2 \times 10^{-5}$RIU. This correlates well with the limiting factor of this particular experimental system which used no temperature control: indeed, during the experiments, the temperature in the laboratory could easily change by 0.2°C over the course of 15 minutes, resulting in a $2 \times 10^{-5}$RIU change in water. A more advanced experimental setup is necessary to explore the true limits of the proposed method.
5. Conclusion

A method was proposed for estimating index of refraction changes from dispersion/reflectance images obtained with a spectro-angular surface plasmon resonance system. The method uses eigenvector analysis with simulated dispersion/reflectance images calculated with RCWA to determine refractive index changes from experimentally acquired image pairs. Numerical simulations indicate that this method is capable of achieving a precision of better than $10^{-8}$ RIU across a wide dynamic range if it were to be used in conjunction with a low noise, high resolution 16bit CCD camera. The method is insensitive to scaling of the data and is robust with respect to noise. Results obtained with a simple experimental system without temperature control showed the ability of the method to track index changes on noisy data with superior performance compared to a standard 1-D SPR data analysis method.

Acknowledgments

We would like to acknowledge funding from Le Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT). We would also like to thank the McGill Institute for Advanced Materials (MIAM) for microfabrication services.