

PHOTONICS Research

Neuromorphic metasurface

ZHICHENG WU,  MING ZHOU,  ERFAN KHORAM, BOYUAN LIU, AND ZONGFU YU*

Department of Electrical and Computer Engineering, University of Wisconsin, Madison, Wisconsin 53706, USA

*Corresponding author: zyu54@wisc.edu

Received 20 August 2019; revised 31 October 2019; accepted 9 November 2019; posted 12 November 2019 (Doc. ID 375886); published 18 December 2019

Metasurfaces have been used to realize optical functions such as focusing and beam steering. They use subwavelength nanostructures to control the local amplitude and phase of light. Here we show that such control could also enable a new function of artificial neural inference. We demonstrate that metasurfaces can directly recognize objects by focusing light from an object to different spatial locations that correspond to the class of the object. © 2019 Chinese Laser Press

<https://doi.org/10.1364/PRJ.8.000046>

1. INTRODUCTION

Optical neuromorphic computing offers an alternative approach to realize artificial neural computing. It has several potential advantages compared with digital neural computing such as ultrafast speed and ultralow energy consumption. Several architectures have been demonstrated based on integrated silicon photonics [1], diffractive optics [2], and nanophotonic random structure [3]. In this paper, we introduce another platform to realize artificial neural computing based on metasurfaces. Metasurfaces were developed to perform arbitrary phase front engineering [4]. Their optical functions are realized by the resonant scattering of arrays of nanoscale scatterers fabricated on a flat surface. It is compatible with today's nanofabrication and can be mass-produced at low cost [5]. Here, we use these nanoscale scatterers to perform neural computing. It leverages the platform of flat optics to realize high-density integration. We describe the design procedures and demonstrate direct image recognition of handwritten digits.

The concept is illustrated in Fig. 1. An object, such as a handwritten digit, is illuminated by a plane wave. The scattered light is then processed by a multilayer neuromorphic metasurface, which consists of arrays of nanoribbons. By changing the size of the ribbons, we can control the amplitude and the phase of scattered light as shown in Fig. 1, which leads to strong interference of light waves passing through the metasurface. With a few layers of metasurface, the output light becomes a focused beam and is directed toward a spatial location corresponding to the value of the handwritten digit. The widths of the nanoribbons are the trainable parameters, which are learned through a training process similar to stochastic adjoint optimization [3].

This work is related to the diffractive neural network demonstrated by Lin *et al.* in 2018 [2], where they use the thickness of the material that light passes through to modulate the

phase. Changing the thicknesses is not easily compatible with nanofabrication for large-scale integration. By using metasurfaces, we can tune the phase delay using the lateral dimension so that the device can be made easily with today's lithography. In order to account for the phase delay caused by lateral structures, full-wave electromagnetic modeling must be used. Such full-wave modeling can be extremely expensive. Here we describe the approaches to reduce the computational load. Also related to the work is Ref. [3], where continuous media are used for neural computing. Here the metasurface can be fabricated on flat surfaces, greatly simplifying the fabrication process.

2. STRUCTURE DESIGN

We use a specific example to illustrate how to design neuromorphic metasurfaces. The goal is to recognize handwritten digits such as the one shown in Fig. 1. We use the database MNIST [6], which contains 60,000 different handwritten digits. We use 50,000 examples for training and 10,000 examples for the test stage. The neuromorphic metasurface should correctly recognize the value of the digits despite their different handwriting styles. We divide the dataset into two groups. The first group, the training set, is used to train the metasurface. The second group, the test set, is used to test the utility of metasurface. A plane wave illuminates the handwritten digits and then passes through the metasurface, which scatters the light in a way that is equivalent to artificial neural computing. The output light will focus on one of 10 different spatial locations that correspond to different digits. Below, we will use two-dimensional (2D) metasurfaces to illustrate the design process. The three-dimensional design follows the same procedure. The 2D design can be done on a personal desktop in 13 h. The three-dimensional metasurface design is computationally feasible on a computer cluster. The computational resource required will be proportional to the area of metasurface.

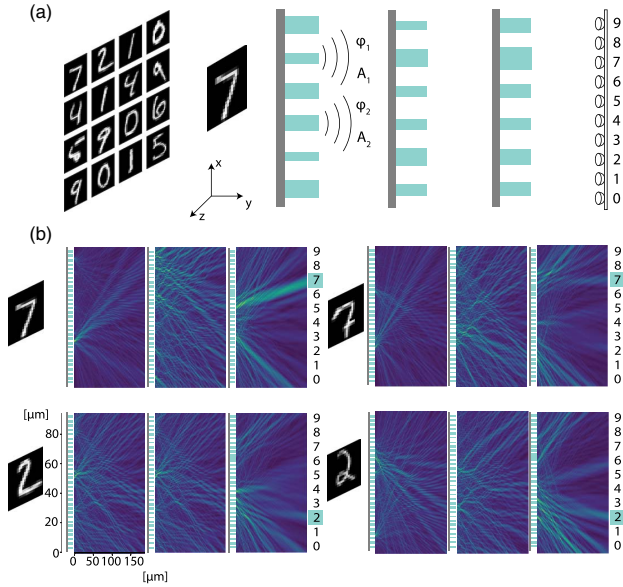


Fig. 1. (a) Schematic of the neuromorphic metasurface. The neuromorphic metasurface consists of multiple layers of nanostructures, which are composed of an array of nanoribbons on top of a dielectric substrate. A handwritten digit is illuminated by a plane wave, and the scattered light then is processed by the neuromorphic metasurface. By changing the sizes of ribbons, the phase and amplitude of the transmitted light after each layer can be modified. After multiple layers, the transmitted field can be focused on specific photodetectors, which are labeled by the values of the handwritten digits, i.e., 0 to 9. (b) Intensity distribution of the transmitted light after each layer in a three-layer neuromorphic metasurface. Handwritten digits of 7 and 2 with different writing styles are used as examples. Despite the different writing styles, the transmitted light is always focused on the spot corresponding to the value of the handwritten digit. Here, we normalize the intensity of the transmitted light after each layer to its maximum for clarity.

The metasurface consists of a large area of subwavelength scattering elements. Full-wave simulation tools such as the finite-difference time-domain method are too computationally expensive for this type of multiscale problem. To obtain the full-wave electromagnetic properties without losing speed, we use locally periodic approximation [7–17]. It assumes the metasurface is locally periodic: the transmitted field in any small region is approximately the same as the transmission from a periodic surface. The field amplitude and phase immediately after a scattering element are calculated by a full-wave simulation assuming a periodic boundary condition, as shown in Fig. 2.

By using a small full-wave simulation to obtain the local field for each element, we can assemble the field along the plane right after the metasurface. Then, we can use near-to-far-field transformation to calculate far-field distribution. Compared to the Rayleigh-Sommerfeld diffraction equation used in Ref. [2], the local periodic approximation takes into account the wave effect of structured scatterers. Compared to the finite-difference full-wave method used in Ref. [3], this method is much faster. The comparison of this method with full-wave simulation can be found in Ref. [7]. Here we use TiO_2 pillars on a SiO_2

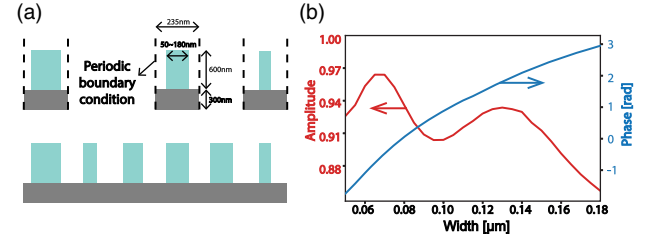


Fig. 2. (a) Schematic of locally periodic approximation. The metasurface consists of an array of TiO_2 pillars on top of a SiO_2 substrate. For plane wave incidence from the bottom side of the substrate, we set up a periodic boundary condition around each pillar. The local field of the transmitted light above each pillar then is approximated by that of the corresponding periodic array. (b) The phase (blue) and amplitude (red) of transmitted light as a function of the width of the pillar under normal plane wave incidence. The results are obtained from a full-wave simulation of a periodic array of pillars, which only takes a few minutes.

substrate to construct the metasurface [13]. As shown in Fig. 2(a), the thickness of the substrate is 300 nm. The height of the pillar is fixed to 600 nm, and the pitch is fixed to 235 nm. We vary the pillar's width from 50 nm to 180 nm. The phase $\phi_n(w)$ and amplitude $A_n(w)$ of the transmitted light are shown in Fig. 3, where w is the width of the pillar and the subscript n represents the normal incident direction. The operating wavelength is 700 nm.

The input wavefront to neuromorphic metasurfaces is generally much more complex than plane waves. Since we have to use a plane wave as the incident condition when applying the locally periodic approximation, we first decompose the incoming wavefront $E(x)$ using Fourier basis $E_k = \sum_x E(x) e^{ikx}$ and then simulate the response of metasurface under each individual plane wave E_k . Then, we sum all the contribution of plane waves together. We could also safely neglect plane waves with large wave vector k because of the large distances between different metasurfaces and between the object and the metasurface.

The phase delay and amplitude modulation change for plane waves incident from different angles. Figure 3 shows the response of the pillars for different incident angles. When we only consider small k components, which correspond

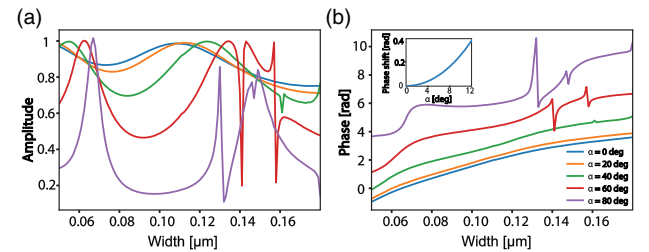


Fig. 3. (a) Amplitude and (b) phase of the transmitted field as a function of the width of the pillar for the different incident angle α . The incident wavelength is 700 nm. When the incident angle is small, the phase response curve shifts horizontally as we increase the incident angle, while the amplitude response does not vary significantly with the incident angle. When the incident angle continues to increase, some resonances appear. The inset shows the shift of the phase response as a function of the incident angle. The shift increases nonlinearly with increasing incident angle.

to small incident angles, the phase response curve shifts horizontally but the amplitude does not vary significantly. This observation allows us to further accelerate the computation by approximating the angular response with $E_c(x) = \sum_k E_k e^{-ikx + i\theta_k}$. The phase compensation θ_k accounts for the difference of phase delay compared to the normal incident wave $k = 0$ [Fig. 3(b) inset]. Now we can calculate the scattering field using transmission of normally incident plane wave with the corrected wavefront compensated for the different incident angles. The transmitted wave is calculated by the convolution $E_c(x) * A_n[w(x)]e^{i\phi_n[w(x)]}$, where $w(x)$ is pillar width at position x .

We now discuss the training process. The output of the neuromorphic metasurface is defined by the distribution of electric field intensity on a plane behind the last layer of the metasurface. In the 2D case, the output is $y(x) = |E_f^L|^2$. Here we use subscript f to indicate the far-field distribution of light after passing the L th metasurface layer. The training target for the output is

$$y_t(x) = I_0 \exp\left[-\frac{(x - x_t)^2}{2\sigma^2}\right], \quad (1)$$

where t is the value of the handwritten digits. x_t is the location where we would like output light to focus on. Locations for different digit values are evenly distributed on the output plane. One can also choose other training targets as long as it serves the purpose of classification. In our 2D case, the peak positions of the target intensity x_t for different digits are equally spaced by $9.4 \mu\text{m}$, and the variance σ of the target intensity is $2.35 \mu\text{m}$.

Training the metasurface is a gradient descent process that minimizes a loss function J . Here we use the L2 distance between the metasurface output and the target output $J = \sqrt{[y(x) - y_t(x)]^2}$. Unlike typical optimization used in nanophotonics and metasurfaces [7,17], the gradient descent used here is stochastic, which comes from the input data. Here, a stochastic optimization method Adam is used [18].

Next, we discuss how to compute J and its gradients. First, we try to formulate the relation between the far-field outputs of the l th layer and the $(l-1)$ th layer. This relation depends on the width of pillars w_l^i in the l th layer. The far-field output is calculated from the near-field $E_n^l(x)$ through a near-to-far-field transformation [19],

$$E_f^l(x) = - \int_{\text{surface}} G(x, x') E_n^l(x') dx', \quad (2)$$

where G is a Hankel function, $G(x, x') = -\frac{ik}{4} H_1^{(1)}(kr) \hat{n} \cdot \frac{r}{r}$. Here $k = \frac{2\pi}{\lambda}$, $r = x - x'$, and $r = |r|$. The near-field is obtained through local periodic approximation,

$$E_n^l(x) = \sum_k E_k e^{-ikx + i\theta_k} * A_n[w^l(x)] e^{i\phi_n[w^l(x)]}. \quad (3)$$

Here E_k is the Fourier component of E_f^{l-1} , the far-field output of the $(l-1)$ th layer. This series of calculation that connects $E_f^l(x)$ and $E_f^{l-1}(x)$ can be represented as matrix operations and implemented in TensorFlow. For example, the integral is changed to summation and can be expressed as a matrix multiplication, $E_f = G \cdot E_n$, where $G_{ij} = G(x_i, x'_j)$,

$E_n(j) = E_n(x'_j)$, and $E_f(i) = E_f(x_i)$. We neglect the reflection of the metasurfaces as the low-index substrate used here results in weak reflection.

We now are ready to calculate the derivative of the loss function with respect to the pillar widths $\frac{\partial J}{\partial w(x)}$. The calculation can be divided into two steps because $\frac{\partial J}{\partial w(x)} = \frac{\partial J}{\partial E_n^l(x)} \frac{\partial E_n^l(x)}{\partial w_l(x)}$. The first term is the derivative of the loss function with respect to each layer's near-field output, which is calculated by following the chain rule of derivative $\frac{\partial J}{\partial y} \frac{\partial y}{\partial E_n^l} \frac{\partial E_n^l}{\partial E_n^{l-1}} \dots \frac{\partial E_n^{l-1}}{\partial E_n^1}$ in TensorFlow. The second is the derivative of the output field with respect to the pillar widths $\frac{\partial E_n^l(x)}{\partial w_l(x)} = \frac{\partial E_n^l(x)}{\partial \phi(x)} \frac{\partial \phi(x)}{\partial w_l(x)} + \frac{\partial E_n^l(x)}{\partial A(x)} \frac{\partial A(x)}{\partial w_l(x)}$. The phase $\phi[w^l(x)]$ and amplitude $A[w^l(x)]$ as a function of pillar width are shown in Fig. 3, which allows us to easily calculate $\frac{\partial \phi}{\partial w}$ and $\frac{\partial A}{\partial w}$. One difference from the conventional deep learning is that the learnable parameters here are also constrained by the physical limit of pillar sizes.

Generally, the input of neuromorphic metasurface is the light scattered by an object. In the simulation, the input is replaced by the image of the object. For the 2D case, we vectorize the image of the handwritten digit number. The original image is resized to 20 by 20 pixels and converted to a 1 by 400 vector, and the intensity is normalized from 0 to 1. Then, we can set the intensity of the vectorized image as the amplitude of the input field. The phase of the input field is set to be the same. The input field is polarized in the z direction such that field can be treated as a complex scalar in simulation and the wavelength is 700 nm. At this frequency, the response of periodic TiO_2 structure changes smoothly when the width of pillar changes. To match the size of the input vector, each layer of neuromorphic metasurface also contains 400 elements. The pitch is 235 nm wide, the total length of the metasurface is $94 \mu\text{m}$, and the distance between the two adjacent layers is $188 \mu\text{m}$. The distance between adjacent layers is chosen based on two criteria. First, the distance should be large enough so that only the far-field from one layer of the metasurface reaches the next layer. Second, as we discussed earlier, we approximate the far-field wavefront by plane waves with only small k vectors. The distance should be large enough so that the contributions from plane waves with large k vectors to the wavefront can be neglected. For any distance between adjacent layers that satisfies the above two criteria, the calculation process is the same. However, the system needs to be retrained after changing the distance between the adjacent layers, and the accuracy will decrease if the layers are too far apart.

3. RESULTS AND DISCUSSION

The training process of the five-layer neuromorphic metasurface is shown in Fig. 4(b). Each data point is the averaged L2 loss over 100 training samples. The computation took about 13 h on an Intel Core i5-4430 CPU 3.00 GHz \times 4.

The neuromorphic metasurface starts to show its utility even with just two layers of metasurfaces, where it can achieve 80% accuracy for MNIST classification. It means that eight out of 10 times, this double-layer metasurface can focus the light on the right location based on the meaning of the handwritten digit. It is a remarkable focusing effect compared with

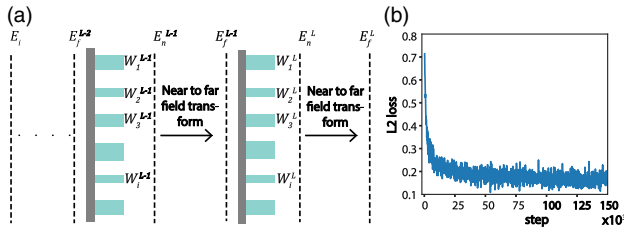


Fig. 4. (a) Forward propagation in a neuromorphic metasurface that has L layers. At the l th layer, the pillars take the far-field from the $(l-1)$ th layer E_f^{l-1} as input and process it. The output near-field E_n^l is obtained by using locally periodic approximation. The corresponding far-field E_f^l at the l th layer then is obtained by using near-to-far-field transformation. The final output of the neuromorphic metasurface is the intensity of light $|E_f^L|^2$. (b) The training loss of a five-layer neuromorphic metasurface as a function of training steps. Each step, we use one training sample to update widths of the pillars. The training dataset is reshuffled every 50,000 steps.

Table 1. Accuracies of the Neuromorphic Metasurface for Different Number of Layers

Layer	2	3	4	5	6
Accuracy	80%	85%	88%	89%	90%

traditional metasurfaces that focus all light to a single spot. The accuracy can be further improved when more layers are used. These results are shown in Table 1. However, more layers lead to more energy loss, which leads to the difficulty of detecting in practice. The output intensity of a multilayer structure decreases with increasing number of layers as $I = I_0(T)^L$, where I_0 is the incident intensity, T is the transmission efficiency of each layer, and L is the number of layers. In practice, the transmission efficiency should also be optimized during training if

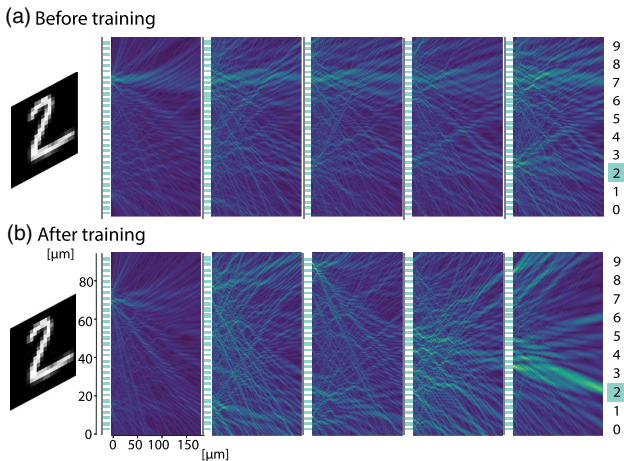


Fig. 5. Comparison of light propagation (a) before and (b) after training for a five-layer neuromorphic metasurface. The input object is a handwritten digit of 2. Before training, the widths of the pillars are randomly initialized, and the transmitted light is randomly distributed at the detector. After training, the transmitted light is directed to detector 2, which corresponds to the input handwritten digit. Here, the intensity distribution of the transmitted light after each layer is also normalized by its maximum for clarity.

more layers are added to the system. Figure 5 shows the light field propagation in a five-layer neuromorphic metasurface before and after training. It can be seen that at the beginning of the training, light is directed to a random distribution. After training, light is focused on the right classification spot.

Unlike our previous work demonstrated in Ref. [3], here we did not use nonlinear activation. In this simple recognition task, nonlinear activation does not significantly enhance performance, but nonlinear activation is crucial for more complex tasks such as face recognition. Nonlinear materials such as a layer of saturable absorber can be easily fabricated into multilayer metasurfaces. In Ref. [3], we solve the nonlinear Maxwell's equation to account for the nonlinear activation. To make the computation more manageable, here we did not apply nonlinear activation for these multiscale metasurfaces. Further work is needed to significantly speed up the electromagnetic modeling of nonlinear materials to be used for metasurfaces.

Funding. National Science Foundation (1641006); Defense Advanced Research Projects Agency.

Acknowledgment. The authors thank Mikhail Kats and Lei Ying for their help in processing.

REFERENCES

1. Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics* **11**, 441–446 (2017).
2. X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science* **361**, 1004–1008 (2018).
3. E. Khoram, A. Chen, D. Liu, L. Ying, Q. Wang, M. Yuan, and Z. Yu, "Nanophotonic media for artificial neural inference," *Photon. Res.* **7**, 823–827 (2019).
4. N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, "Light propagation with phase discontinuities: generalized laws of reflection and refraction," *Science* **334**, 333–337 (2011).
5. N. Yu and F. Capasso, "Flat optics with designer metasurfaces," *Nat. Mater.* **13**, 139–150 (2014).
6. C. C. Yann, *MNIST handwritten digit database*, Yann LeCun, Corinna Cortes and Chris Burges (2015).
7. R. Pestourie, C. Pérez-Arancibia, Z. Lin, W. Shin, F. Capasso, and S. G. Johnson, "Inverse design of large-area metasurfaces," *Opt. Express* **26**, 33732–33747 (2018).
8. F. Aieta, P. Genevet, M. A. Kats, N. Yu, R. Blanchard, Z. Gaburro, and F. Capasso, "Aberration-free ultrathin flat lenses and axicons at telecom wavelengths based on plasmonic metasurfaces," *Nano Lett.* **12**, 4932–4936 (2012).
9. A. Arbabi, E. Arbabi, Y. Horie, S. M. Kamali, and A. Faraon, "Planar metasurface retroreflector," *Nat. Photonics* **11**, 415–420 (2017).
10. M. Khorasaninejad, W. T. Chen, A. Y. Zhu, J. Oh, R. C. Devlin, C. Roques-Carmes, I. Mishra, and F. Capasso, "Visible wavelength planar metalenses based on titanium dioxide," *IEEE J. Sel. Top. Quantum Electron.* **23**, 43–58 (2017).
11. F. Aieta, M. A. Kats, P. Genevet, and F. Capasso, "Multiwavelength achromatic metasurfaces by dispersive phase compensation," *Science* **347**, 1342–1345 (2015).
12. M. Khorasaninejad, F. Aieta, P. Kanhaiya, M. A. Kats, P. Genevet, D. Rousso, and F. Capasso, "Achromatic metasurface lens at telecommunication wavelengths," *Nano Lett.* **15**, 5358–5362 (2015).
13. M. Khorasaninejad, Z. Shi, A. Y. Zhu, W. T. Chen, V. Sanjeev, A. Zaidi, and F. Capasso, "Achromatic metalens over 60 nm bandwidth in the visible and metalens with reverse chromatic dispersion," *Nano Lett.* **17**, 1819–1824 (2017).

14. E. Arbabi, A. Arbabi, S. M. Kamali, Y. Horie, and A. Faraon, "Controlling the sign of chromatic dispersion in diffractive optics with dielectric metasurfaces," *Optica* **4**, 625–632 (2017).
15. V.-C. Su, C. H. Chu, G. Sun, and D. P. Tsai, "Advances in optical metasurfaces: fabrication and applications [Invited]," *Opt. Express* **26**, 13148–13182 (2018).
16. B. Groever, C. Roques-Carmes, S. J. Byrnes, and F. Capasso, "Substrate aberration and correction for meta-lens imaging: an analytical approach," *Appl. Opt.* **57**, 2973–2980 (2018).
17. A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," *Nat. Photonics* **9**, 374–377 (2015).
18. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2017).
19. A. Taflov and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed., Artech House Antennas and Propagation Library (Artech House, 2005).