Integrated color pixels in 0.18-μm complementary metal oxide semiconductor technology

Peter B. Catrysse and Brian A. Wandell

Department of Electrical Engineering, Stanford University, Stanford, California 94305

Received January 14, 2003; revised manuscript received June 11, 2003; accepted July 23, 2003

Following the trend of increased integration in complementary metal oxide semiconductor (CMOS) image sensors, we have explored the potential of implementing light filters by using patterned metal layers placed on top of each pixel's photodetector. To demonstrate wavelength selectivity, we designed and prototyped integrated color pixels in a standard 0.18-μm CMOS technology. Transmittance of several one-dimensional (1D) and two-dimensional (2D) patterned metal layers was measured under various illumination conditions and found to exhibit wavelength selectivity in the visible range. We performed (a) wave optics simulations to predict the spectral responsivity of an uncovered reference pixel and (b) numerical electromagnetic simulations with a 2D finite-difference time-domain method to predict transmittances through 1D patterned metal layers. We found good agreement in both cases. Finally, we used simulations to predict the transmittance for more elaborate designs. © 2003 Optical Society of America

OCIS codes: 040.6040, 050.2770, 110.2970, 130.3120, 130.5990, 230.3990.

1. INTRODUCTION

Digital cameras have emerged as the dominant image-capture technology. The core component of a digital camera, its image sensor, records an optical image as a spatially sampled array of pixels. The charge accumulated in these pixels is an approximately linear representation of the local image irradiance. To create color images, color filters are inserted in the optical path before each pixel; these filters control the likelihood that various wavelengths will be transmitted to each pixel and its photodetector. Color filters are commonly superimposed on the image sensor during independent steps in the manufacturing process.[1,2]

An important trend in digital camera design is the development of complementary metal oxide semiconductor (CMOS) image sensors. These sensors are being scaled with CMOS technology to make possible an increasing level of integration of capture and processing to reduce system power and cost.[3] Toward completing design integration of color image sensors in CMOS technology, we have explored the possibility of introducing wavelength selectivity with only standard processing steps, i.e., without inserting a color filter. Specifically, we have implemented filters that consist of patterned metal layers placed within each pixel to control the transmission of light through the pixel to its photodetector. We refer to such pixel design as an integrated color pixel (ICP). ICPs may be useful for digital camera applications but may be even more suited to multispectral imaging and a variety of other image-sensing applications.

We demonstrated the first ICPs featuring visible wavelength selectivity by using subvisible-wavelength patterned metal layers in a standard 0.18-μm CMOS technology.[4] Here we describe the comprehensive optical characterization of CMOS ICPs, including transmittance measurements under various illumination conditions. We also present what is to our knowledge the first model to characterize (a) the optical path from the pixel surface to the photodetector in a CMOS image sensor and (b) the effect of placing patterned metal layers in this path. We compare the measured and the predicted transmittances, and then we show, by simulation, examples of additional transmittances that can be achieved by using advanced CMOS technologies and multiple patterned metal layers simultaneously.

2. BACKGROUND

A. Color Filter Arrays

Today the main method of implementing color image sensors for digital cameras is the color filter array[5] (CFA). The CFA consists of a color filter pattern, comprising a regular array of filters with several different wavelength transmittances, that is deposited onto the monochrome pixels to separate out the wavelengths of incident light. In one common form, this architecture uses a spatial pattern comprising three red–green–blue (RGB) interleaved color mosaics.[6,7] An alternative pattern, often used for video capture, comprises an array of four cyan–magenta–yellow–green (CMYG) color filters.[8]

The CFA color architecture has been quite successful for CCD imagers, but it does have several performance drawbacks when applied to CMOS image sensors. First, the CFA is applied to the pixel surface, whereas the photodetector is located in the substrate [Fig. 1(a)]. In 0.18-μm CMOS technology, the distance between filter and detector can be as much as 10 μm for a typical 6-μm pixel with a 3-μm photodetector. This spacing makes the pixel a less efficient light collector than it could be.[9] It also permits light intended for, say, the red photodetector to refract or scatter and arrive, instead, at a neighboring blue photodetector. This phenomenon is called color cross talk.[10,11] Further, the relatively large separation imposes stringent design constraints on micro-

 © 2003 Optical Society of America

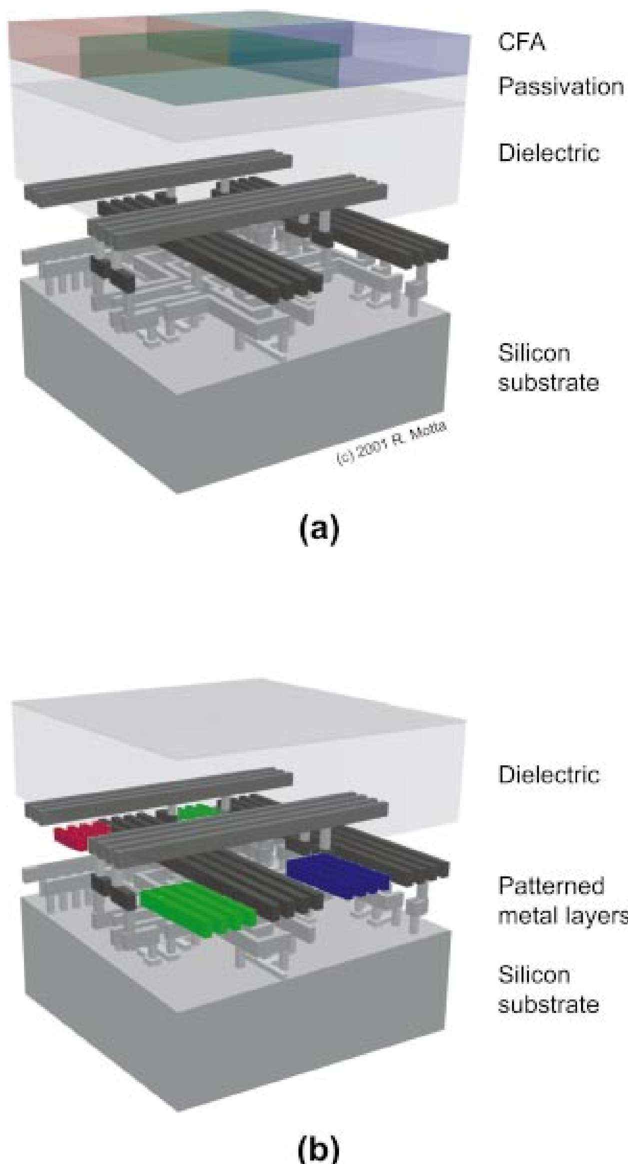


Fig. 1. Geometry of four pixels in a CMOS image sensor. (a) A conventional arrangement including a RGB color-filter array placed on the sensor surface. (b) The integrated color pixel arrangement that contains patterned metal layers (shown in color) within the pixel tunnels. Used by permission of the copyright holder, R. Motta.

lenses; these constraints are usually necessary to increase the sensitivity of CMOS image sensor pixels.[12]

There are several designs that aim to alleviate the problems with the CFA architecture. For example, single-chip three-colors-in-one-pixel CMOS image-sensor designs have emerged. The thin-film-on-application-specific integrated circuit (ASIC) (TFA) approach deposits thin-film detectors with adjustable spectral sensitivity on top of an ASIC.[13] In a second approach, a triple-well structure is used to exploit silicon's differences in the absorption length of light of different wavelengths for color separation.[14] Although promising, neither of these recent designs can be easily implemented in a standard CMOS process without making significant modifications in the process.

B. Wire Grids

The use of wire grids for filtering electromagnetic radiation goes back to the early experiments of Heinrich Hertz in the late 1800s.[15] In the field of microwave engineering, regular metal grids have been used for a long time as filters, transmitting some wavelengths and blocking others.[16] More recently, as a result of continued advances in the manufacturing technology of integrated-circuits, metal grids with periodicity commensurate with the wavelength of near-infrared light have been demonstrated,[17] and some of their properties have been studied numerically. For example, Jensen and Nordin describe a numerical electromagnetic study of the transmission characteristics of finite-aperture wire-grid polarizers.[18]

Integrating the light filters, with use of patterned metal layers, inside image sensor pixels offers several potential advances. Placing the light filters close to the photodetector reduces color cross talk, and pixel vignetting and increases the efficiency of microlenses because of the smaller distance from pixel surface to pixel photodetector. By controlling wavelength responsivity in the pixel design process, ICPs eliminate the need for additional CFA manufacturing steps. The patterns used in the design of an ICP use already available metal layers and are deposited on top of the photodetector in a CMOS image-sensor pixel.

Why have wire grids not been used for color filtering in the visible light regime? To control light in the visible wavelength range requires periodicities smaller than the visible range. This has been impossible until very recent 0.18-μm CMOS technologies were used to make image sensors. Even with the wire sizes possible in 0.18-μm CMOS technology, which is considered state of the art for image sensors, we are on the edge of controlling visible wavelengths. Nevertheless, as CMOS technologies keep scaling and smaller feature sizes become possible, CMOS image sensors are bound to become viable in these more advanced technologies.

3. IMPLEMENTATION: INTEGRATED COLOR PIXEL TEST STRUCTURES

We designed and implemented ICP test structures in a standard 0.18-μm CMOS technology to demonstrate the visible wavelength selectivity of patterned metal layers and to investigate their use for color imaging [Fig. 1(b)]. Periodic patterns were created in the metal layers already present in the process. These patterns comprised an array of elements whose gap widths ranged from 270 to 540 nm. The refractive index of the surrounding dielectric is 1.46. In air the visible wavelength regime is roughly from 400 to 750 nm; in this medium the effective visible wavelength regime is roughly from 270 to 510 nm.

We implemented test structures comprising 12-μm ICPs with n-diffusion/p-substrate ($n+/p\,\mathrm{sub}$) photodiodes. An ICP circuit schematic is shown in Fig. 2(a). A standard three-transistor active-pixel-sensor (APS) circuit, directly connected to a chip output via a two-stage follower amplifier, was used. Transistor gate leakage was reduced, and voltage swing was increased, by using

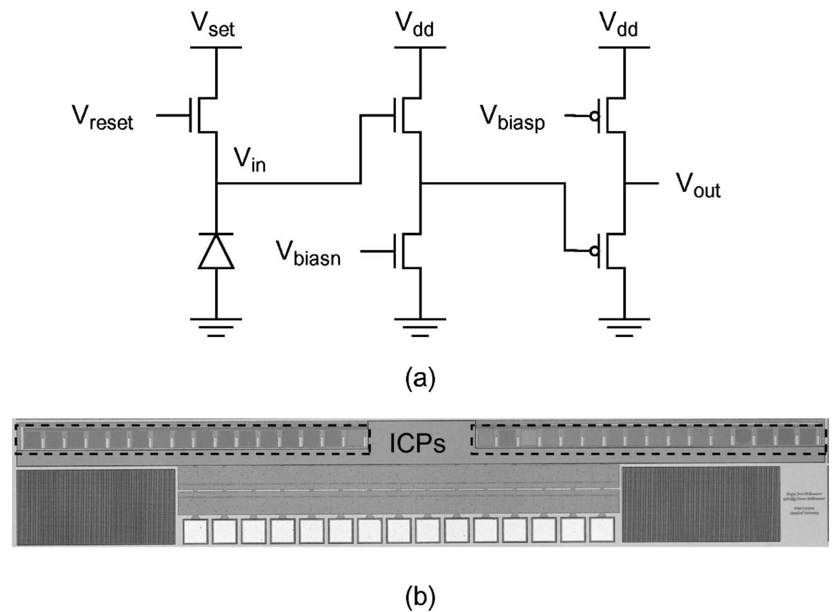


Fig. 2. ICP implementation in 0.18-μm CMOS technology. (a) The APS circuit, (b) a photomicrograph of the ICP test structure chip.

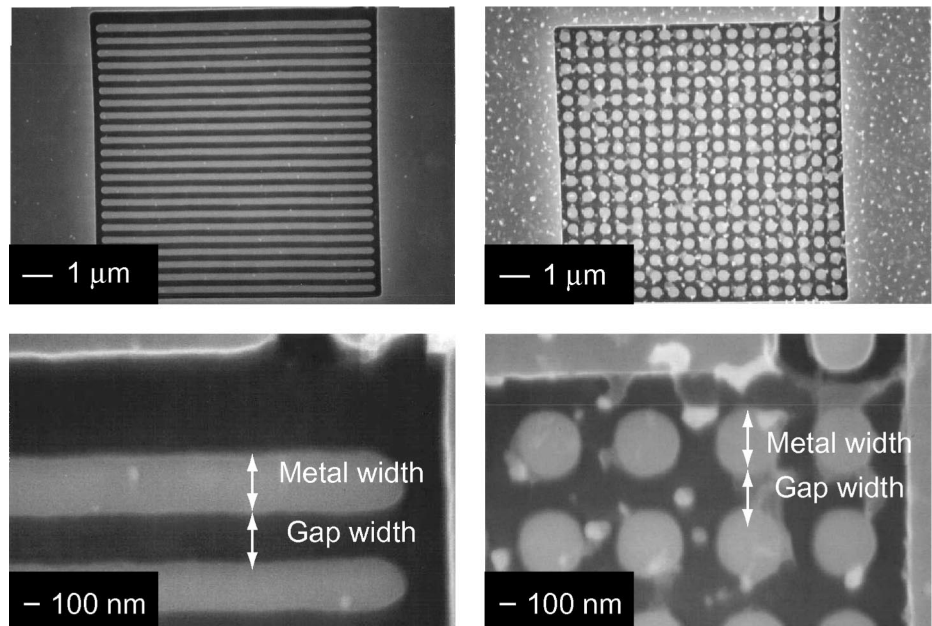


Fig. 3. Scanning electron micrographs of the ICP patterned metal layers in 0.18-μm CMOS technology. The 1D (left) and 2D (right) patterns are shown at two different spatial resolutions. The white spots, seen particularly on the right, are residual traces of material that was not removed completely during the deprocessing necessary to make these images.

3.3-V thick-oxide transistors. The circuit operation of the ICP is otherwise standard.[19–21] A photomicrograph of the chip with the ICP test structures is shown in Fig. 2(b).

The electrical properties of the ICPs are similar to those of standard APS pixels.[19] The transfer function relating the measured voltage, V_{out}, to the voltage at the photodiode node, V_{in}, is sigmoidal over its entire range, but approximately linear,

$$V_{\mathrm{out}} = gV_{\mathrm{in}}, \tag{1}$$

for V_{out} values ranging between 0.9 and 2.1 V with a slope $g = 0.7$. The dark current is approximately 25 mV/s.

The ICP patterned metal layers were deposited on top of APS photodetectors (Fig. 3). The test structures included one-dimensional (1D) and two-dimensional (2D) patterns with a periodicity (metal width + gap width) ranging from 540 to 810 nm and a gap width ranging from 270 to 540 nm. The scanning electron micrographs of the metal patterns shown in Fig. 3 confirm that the actual patterns were close to the design specifications.[4] The principal difference is that the edges of the pattern elements are slightly rounded. Because of the 2D limitation

of the current generation of finite-difference time-domain simulation software,[22] we are able to simulate only the 1D patterned metal layers. Hence we focus our discussion on these arrays and the associated simulation predictions.

4. EXPERIMENTAL METHODS

ICP characterization includes optical and electrical measurements.[21,23,24] To measure the ICP characteristics accurately, one has to take special care to minimize the environmental interferences such as electromagnetic interference and light source and temperature fluctuations. We housed the setup in an air-conditioned dark room on an optical tabletop. All identifiable sources of stray light were blocked, and experiments were performed remotely by use of a networked computer interface. Temperature was kept close to 25 °C. Absolute irradiance $E(\lambda)$ of the light source in W/(m^2 nm) was measured with a spectroradiometer and a diffuse white standard. Calibration measurements were repeated during each experimental series.

A. Electrical Setup

The electrical setup consists of a device-under-test board, a power supply, an analog output board, and a data acquisition board. All electrical equipments are computer controlled.

The device-under-test board holds the ICP test chip, some analog bias circuits, and a digital control signal interface. The board is mounted on an XYZ-translation stage, which in turn is placed on a rotation stage. The board and chip are aligned with the monochromator output port by use of an optical rail.

A highly regulated dc power supply is used to provide a stable $V_{\text{dd}} = 3.3\,\text{V}$ with less than 2 mV of ripple (peak to peak) for the entire chip. The analog output board provides the two bias voltages $V_{\text{bias}n}$ and $V_{\text{bias}p}$ for the two-stage follower that converts the ICP response to the output, the reset voltage level V_{set} for the ICP photodiode, and the reset pulse V_{reset}, which determines the integration time. All these analog output waveforms are generated with a 12-bit resolution at a 1-MHz rate. The analog output board also provides the digital control signal for the 5-bit multiplexer, which is used to select one of the 32 pixels by connecting the multiplexer output node to the global output pad of the chip. The data acquisition board samples and digitizes the analog voltage signal at the ICP output at a 1-kHz rate, using a 12-bit analog-to-digital converter. The analog/digital range from 0 to 5 V matches the ICP output signal range. The output range spans 0.9 to 2.0 V with quantization steps of 1.22 mV. The 1-kHz sampling rate is sufficiently fine to capture integration times ranging from 75 ms to 2 s.

B. Optical Setup

The optical setup comprises a quartz–tungsten halogen light source, a fiber light guide, a monochromator, an integrating sphere, a collimator–polarizer, and a spectroradiometer.

A dc-regulated light source with a 150-W quartz–tungsten halogen lamp provides stable optical power output from 400 to 1100 nm with less than 0.1% short-term ripple and noise (intra-experiment stability) and less than 1% long-term drift (inter-experiment stability). A matched fused-silica fiber light guide delivers the light to the input port of a monochromator with an approximate efficiency of 50% for the entire wavelength range.

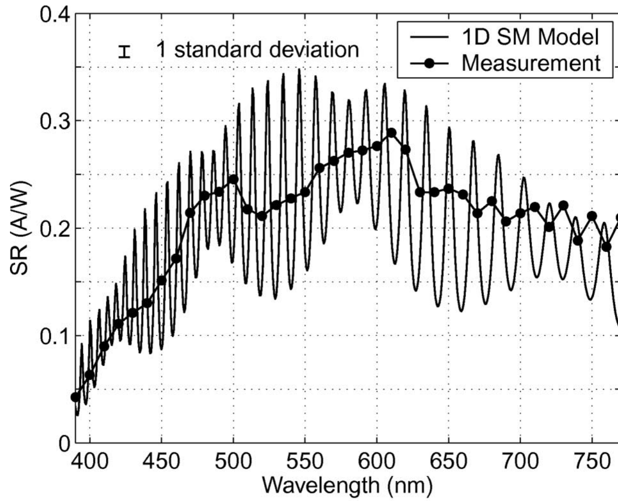
An 1/8-m motorized monochromator, based on an out-of-plane Ebert–Fastie design, separates the broadband white-light input into narrowband spectral components by using a 1200-l/mm ruled grating with a 350-nm blaze wavelength. Its peak efficiency is 80% at 400 nm, and its primary wavelength region is from 200 to 1600 nm (efficiency $> 20\%$). To control the spectral resolution of the light at the output port, we use fixed input and output slits with different widths. The slit width controls the spectral resolution (the narrower the slit, the smaller the wavelength band of the light source selected at the output port) and simultaneously the amount of light transmitted. We used slits with resolution at 500 nm ranging from 1 to 10 nm. Most of the measurements were done with a fixed slit that provided an 8-nm FWHM passband. This slit passes enough light that the photocurrent exceeds the dark current by at least a factor of 2 across measurement wavelengths. The passband also limits the wavelength-sampling interval to exceed 8 nm, which is adequate for characterizing the expected transmittance of the patterned metal layers.

The output of the monochromator can be connected to an integrating sphere to provide a uniform, randomly polarized illumination. This is useful for emulating imaging conditions. We used a 2-in. integrating sphere specifically designed for uniform monochromator illumination. It has excellent diffusing properties provided by a very stable barium-sulfate-based white coating, which is >95% reflective from 400 to 700 nm. Proper baffling prevents "first-strike" radiation from entering the instrument. Alternatively, the output of the monochromator can be connected to a collimator–polarizer to provide a collimated, linearly polarized illumination. This is most useful for performing analysis measurements, i.e., measurements with simple illumination conditions that can be either analytically or numerically simulated.

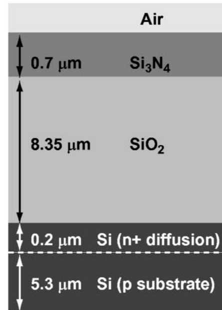
The spectroradiometer is used to make absolute irradiance measurements of the light incident on the ICP. The spectroradiometer measures the spectral radiance from 380 to 780 nm with a spectral bandwidth of 8 nm and an accuracy of 2 nm. To allow irradiance measurements, we aim the spectroradiometer at a diffuse white standard. The white standard has reflectance values ranging from 95% to 99% for the visible ($\pm 1\%$) and near-infrared ($\pm 4\%$) spectrum and is highly Lambertian over its effective spectral range. This enables us to use a simple conversion formula to calculate the incident irradiance $E(\lambda)$, in W/(m^2 nm), at the location of the patch from the measured radiance $L(\lambda)$, in W/(m^2 nm sr), re-emitted by the patch,

$$E(\lambda) = \pi L(\lambda). \tag{2}$$

The RS-232 serial port of a networked computer controls both the monochromator and the spectroradiometer.



(a)



(b)

Fig. 4. Spectral responsivity of an uncovered pixel (circles connected by solid lines) measured with unpolarized narrowband illumination (10-nm sampling intervals, FWHM of 8 nm). (a) The solid curve is the predicted spectral responsivity, based on the scattering-matrix (SM) calculations derived in Subsection 6A. The error bar shows a typical 1 standard deviation for the measurements. (b) The geometry of the pixel assumed in calculating the 1D scattering-matrix model is shown.

C. Transmittance Measurement

If a photodiode is exposed to a narrowband spectral irradiance $E(\lambda)$, the collected photo charge Q can be expressed as

$$Q = A_{\mathrm{eff}} T_{\mathrm{int}} \int SR(\lambda) E(\lambda) \mathrm{d}\lambda \approx A_{\mathrm{eff}} T_{\mathrm{int}} SR(\lambda) E(\lambda) \Delta\lambda, \tag{3}$$

where A_{eff} is the effective photodetector area of a pixel, T_{int} is the integration time, and $SR(\lambda)$ is the spectral responsivity in A/W.[25] The collected charge Q is proportional to the voltage V_{in} at the photodiode node [see Fig. 2(a)],

$$V_{\mathrm{in}} = \frac{Q}{C_{\mathrm{int}}} \approx \frac{A_{\mathrm{eff}} T_{\mathrm{int}}}{C_{\mathrm{int}}} SR(\lambda) E(\lambda) \Delta\lambda. \tag{4}$$

Hence, within the linear regime of the transfer function Eq. (1), the spectral responsivity $SR(\lambda)$ is proportional to the measured voltage V_{out} at the multiplexer output node divided by the input narrowband spectral irradiance $E(\lambda)$ at the input wavelength,

$$SR(\lambda) \approx \frac{C_{\mathrm{int}}}{A_{\mathrm{eff}} T_{\mathrm{int}}} \frac{V_{\mathrm{out}}}{g} \frac{1}{E(\lambda)\Delta\lambda}. \tag{5}$$

The spectral responsivity $SR(\lambda)$ of the uncovered pixel, measured with a collimated source, is shown by the circles connected by solid lines in Fig. 4. The spectral responsivity is typical for the silicon substrate in a CMOS process.[26] It starts out low, near 400 nm, owing to reflection and absorption in the stratified structure preceding the pn-junction photodiode. It peaks near 600 nm as a result of the location of the pn junction 200 nm below the silicon surface. It falls off at higher wavelengths, and finally silicon becomes transparent for wavelengths larger than 1120 nm, which corresponds to its bandgap. We will compare these measurements with theory in more detail in Section 6. The predicted rapid oscillations arise because of Fabry–Perot resonance behavior within the pixel structure. These oscillations are frequently found when one is measuring at a fine enough wavelength resolution (<10 nm) using collimated sources. More typically, however, $SR(\lambda)$ is measured by using wide-angle, unpolarized illumination from an integrating sphere. Such light masks most of the Fabry–Perot resonance behavior. The spectral transmittance through the patterned metal layers is measured indirectly by comparing the spectral responsivity of the ICP with that of an uncovered reference pixel,

$$\mathcal{T}_{\mathrm{meas}}(\lambda) = \frac{SR_{\mathrm{ICP}}(\lambda)}{SR_{\mathrm{Ref}}(\lambda)}. \tag{6}$$

5. EXPERIMENTAL RESULTS

Figure 5 shows the spectral transmittance of a 1D patterned metal layer consisting of 270-nm-wide wires and gap widths ranging from 270 to 540 nm, where sizes represent the design specifications. Transmittance is shown for collimated illumination polarized with the electric field parallel to the wires [TE; Fig. 5(a)] and with the electric field perpendicular to the wires [TM; Fig. 5(b)].

Although we refer to the measurement conditions as TE and TM, the 1D patterned metal layers are parallel wires whose edges do not quite connect to the metal shield (see Fig. 3, left side). These edges form two vertical gaps perpendicular to the main polarization direction. Consequently, the measurements are TE or TM weighted rather than pure TE or TM measurements.

Both TE and TM transmittance behavior is determined by both pattern geometry and material properties. Specifically, we have observed two TE transmittance regions: a constant but attenuated transmission region (left) and a cutoff region (right). We summarize the behavior in these regimes with a three-parameter model consisting of a constant transmittance level α, a cutoff wavelength $\lambda_{\mathrm{cutoff}}$, and a cutoff rate β:

$$\mathcal{T} = \begin{cases} \alpha & \lambda \leqslant \lambda_{\mathrm{cutoff}} \\ \alpha - \beta(\lambda - \lambda_{\mathrm{cutoff}}) & \lambda > \lambda_{\mathrm{cutoff}} \end{cases}. \tag{7}$$

The parameters in this empirical model can be related to the geometrical properties of the grating, as we describe later.

The empirical model for the TE measurements shows that for the three pixels in Fig. 5(a), the $\lambda_{\mathrm{cutoff}}$ value increases with gap width (677, 750, and 875 nm). The transmittance level α increases with increasing gap–metal ratio (0.41, 0.47, and 0.50). The small discrepancies between α and the gap–metal ratio are probably due to a variety of minor factors, such as incomplete periods at the photodetector margin and absorption of the wire material (400-nm aluminum core with a 25-nm coating of titanium nitride on both top and bottom surfaces). The cutoff rate β is roughly the same at 0.001 nm^{-1}. Its accurate calculation is impeded by the rapid oscillations, which coincide with location of the cutoff wavelength.

The TM transmittances in Fig. 5(b) differ in a few ways from the TE transmittances. Mainly, the transmittance in the lower wavelength regime deviates significantly from the constant level expected by the gap–metal ratio. The TM transmittance measurements do decline at longer wavelengths, similar to the falloff in the TE case, but there are some differences. The most notable difference is that the falloff is steeper in TM than in TE measurements.

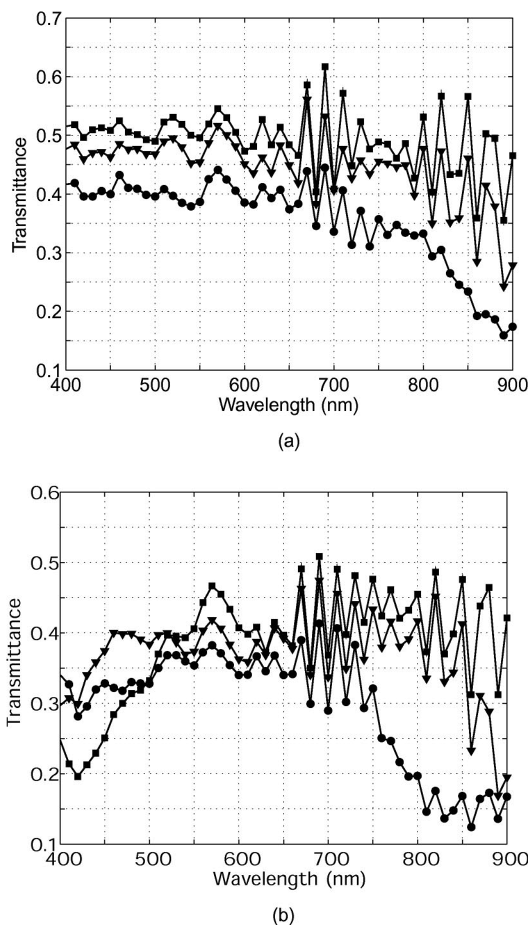


Fig. 5. Measured transmittance of 1D ICPs with gap widths designed to be 270 (circles), 360 (triangles), and 450 (squares) nm. Measurements are shown for collimated, polarized illumination: (a) electric field parallel with wires (TE) and (b) electric field perpendicular to wires (TM).

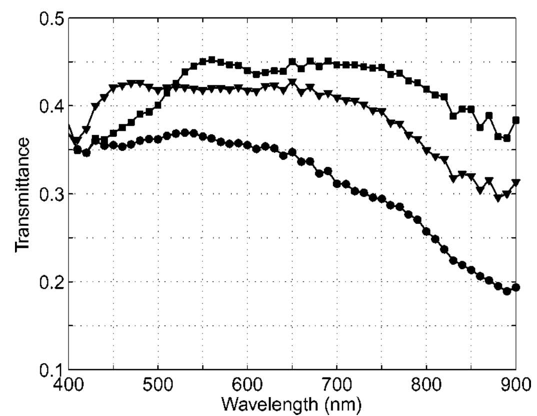


Fig. 6. Measured transmittance of 1D ICPs with gap widths designed to be 270 (circles), 360 (triangles), and 450 (squares) nm. Measurements are shown for integrating-sphere illumination.

The transmittance measurements with wide-angle, unpolarized illumination (Fig. 6) can be explained on the basis of the linearity of the transmitted light; i.e., the wide-angle, unpolarized results are a weighted sum of the single-angle collimated illumination at each polarization. Even under these more natural imaging conditions, changing the gap width changes the spectral transmittance of the patterned metal layers.

The pixel spectral responsivity obtained by using plane waves (collimated, polarized illumination) has rapid oscillations [see Fig. 4(a)]. These oscillations are also plainly visible in the transmittance measurements of the patterned metal layers in regions near 600–700 nm [Figs. 5(a) and 5(b)]. The oscillations are greatly reduced when measuring is with incident waves spanning many angles, say, when an integrating sphere is used. As we describe in Section 6, the position and spacing of these oscillations can be understood as caused by Fabry–Perot reflections between the silicon nitride (Si_3N_4) and silicon dioxide (SiO_2) layers in the pixel. The introduction of the patterned metal layer further subdivides the pixel cavity and introduces additional reflections. Hence some of these oscillations are visible in the transmittance measurements.

6. THEORY

The key properties of the experimental data include (a) the spectral responsivity of the uncovered reference pixel and (b) the transmittance of the patterned metal layers in the ICP. In this section we show that both of these properties can be predicted. The reference pixel responsivity can be predicted with an analytical model based on a scattering-matrix approach and a model of the photodiode. The measured TE and TM transmittance of the patterned metal layers can be predicted by using a numerical electromagnetic field simulation based on the finite-difference time-domain (FDTD) method. We describe the numerical methods in some detail at the end of this section.

Although a CMOS image-sensor pixel is a three-dimensional (3D) structure, in this paper we simulate only a finite 2D cross section in a plane along the direction of light propagation and perpendicular to the long

axis of the wires. We treat the third dimension, along the axis of the wires, as extending infinitely. This simulation accounts for the finite thickness of the metal wires and gaps as well as the periodicity. This approximation is inexact at the edges of the wires, but it is acceptable because the wires are almost 50 times longer than they are wide. For the reference pixel, which has no patterned metal layer, we can further simplify the modeling because there is no alternation between metal and gap, and thus the refractive index is constant in this dimension. This approximation is inexact at the edges of the pixel, but it is acceptable for the case of a 12-μm reference pixel because the contribution to the photodiode signal from the perimeter is modest compared with the signal from the main area.

A. Spectral Responsivity Model for a CMOS Pixel

The spectral responsivity of the uncovered reference pixel [Fig. 4(a)] is dominated by two mechanisms. There is a mean spectral responsivity, which is determined mainly by the absorption properties of the silicon photodiode, and there are substantial oscillations superimposed on this mean, which are determined by the stratified dielectric tunnel leading to the photodiode. To examine the hypothesis that these oscillations are produced by Fabry–Perot resonance within the stratified dielectric tunnel leading to the photodiode, we created a cross-sectional model for the dielectric tunnel and the pn-junction photodiode in the silicon substrate [Fig. 4(b)]. The model consists of a four-layer stratified dielectric structure comprised of air, Si_3N_4, SiO_2, and silicon with known thickness and complex refractive index. For a plane wave at normal incidence, we calculate the transmittance through the stratified structure and the absorption of the transmitted light into the silicon substrate by using a scattering-matrix approach (Appendix A.1).[9] We model the conversion of the absorbed photon flux into a photocurrent by using an abrupt pn-junction photodiode model (Appendix A.2).

The predicted spectral responsivity exhibits both the mean spectral absorption behavior of the silicon substrate and the Fabry–Perot behavior of the stratified dielectric structure (see Fig. 4). The Fabry–Perot reflections can be separated into two fundamental periodicities that can be related to the thickness of the Si_3N_4 and SiO_2 layers. The oscillations in the short-wavelength regime are too rapid to be detected by our measurement apparatus. Beginning at 650 nm, however, small oscillations at the proper spacing are present in the measurements.

B. Spectral Transmittance of the Patterned Metal Layer

We have used the 2D FDTD method to calculate the transmittance of the patterned metal layers in the ICP; the 1D scattering-matrix approach is insufficient to handle the spatial structure of the patterns where the finite dimensions are a significant factor. The transmittance can be calculated without computing the entire pixel responsivity. Instead, we can calculate the transmittance of the metal pattern by measuring the ratio of the power just past the silicon interface for both the uncovered reference pixel and the ICP. This approach eliminates the need for an FDTD dispersive model of the wavelength-dependent silicon absorption, which appears not to exist.

The main effect of the transmittance for TE polarization can be seen in Fig. 7, which plots the transmittance through only that portion of the tunnel containing the patterned metal. In this simulation the metal pattern is modeled as falling within an aluminum aperture within the SiO_2 layer. The dielectric and the metal at the edge of the simulation domain are terminated by use of uniaxial phase-matched layers (UPMLs). The smooth curve is the theoretical transmittance based on the FDTD calculation; the connected circles are the measured transmittance. The plane wave transmittance is constant but attenuated up to 750 nm, and then transmittance decreases.

A more complete simulation, including the stratified dielectric tunnel, is compared with the same measurements in Fig. 8(a). The agreement between the measurements and the simulation is improved by including the oscillations caused by Fabry–Perot resonances. The slight increase in transmittance in the very long wavelengths (beyond 850 nm) occurs because the wires are not connected to the surrounding metal. In the measurements, some light from the TM polarization passes through the gap between the wires and the surrounding metal and is included in this estimated transmittance. Figure 8(b) shows the agreement between measurements and simulation for the TM case.

Figure 8 compares the TE and TM transmittances with a prediction from a weighted sum of the theoretical TE and TM transmittances with use of an aluminum material model. A mixture is expected because the measurements include a contribution from the main horizontal gaps between the wires and the surrounding vertical gaps. The data are reasonably well fitted by use of a relative (horizontal, vertical) proportion of (0.7, 0.3). The contribution from the vertical gaps is larger than might be expected by simply calculating the surface area of these two types of gaps.

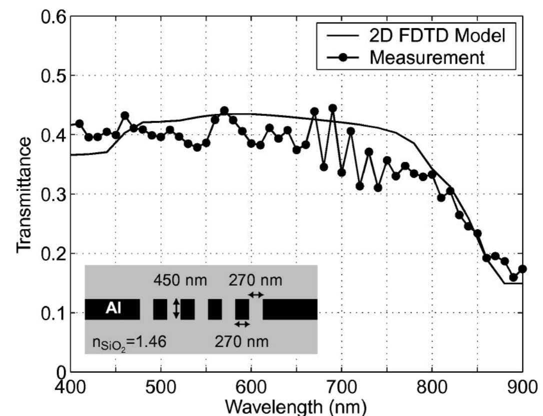
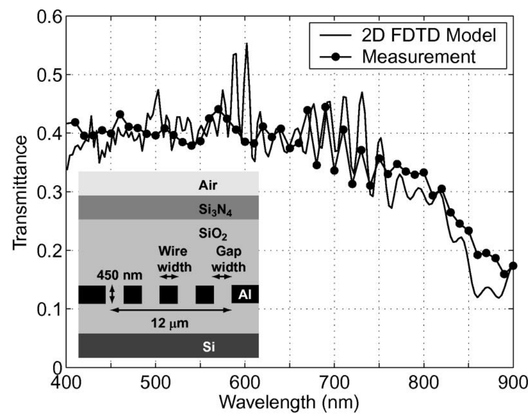


Fig. 7. ICP transmittance measurements (circles connected by solid lines) are compared with 2D FDTD simulation assuming TE polarization (solid curve). For this simulation, the domain includes only the patterned metal layer embedded in a homogeneous SiO_2 medium. The inset indicates the geometry and dimensions of the FDTD simulation domain.



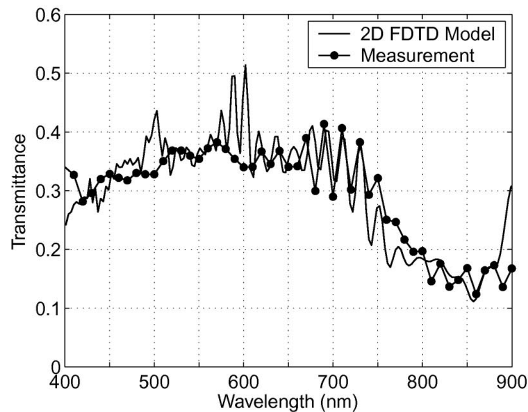


Fig. 8. (a) ICP transmittance measurements (circles connected by solid lines) are compared with 2D FDTD simulation assuming TE polarization (solid curve). The wire width and gap width of the wire grid are the same as those in Fig. 7. The inset shows the FDTD simulation domain, which includes the patterned metal layer embedded in the stratified dielectric tunnel. (b) ICP transmittance measurements (circles connected by solid lines) are compared with 2D FDTD simulation assuming TM polarization (solid curve).

C. Simulations of New Designs

Controlling wavelength transmittance by varying the patterned metal layers creates opportunities to design a wide variety of filters. In this section we illustrate some new designs that involve technology scaling and stacking of the patterned metal layers. We performed these simulations assuming that the patterned metal layers are perfect electrical conductors so that we might visualize the salient features of the transmittance.

The implementation in 0.18-μm CMOS technology was limited to one metal layer and was restricted by current design rules. As technology scales, it will be possible to implement the patterned metal layers at finer scales. The consequence of using increasingly fine patterns is illustrated in Fig. 9, where we compare CMOS technologies with 0.18-, 0.15-, and 0.13-μm feature sizes. In the TE-polarization condition, it is possible to create low-pass filters with a range of cutoff frequencies that span the visible. Hence using 0.13-μm CMOS technology makes it possible to create a set of RGB cutoff filters for visible imaging.

An extension of the current implementation is to use multiple patterned metal layers to form the light filters. A huge number of possibilities exist for these kinds of designs. For example, it is possible to adjust the separations between the metal lines to match the spacing between the layers, creating a Fabry–Perot cavity between the layers. Pairs of metal layers can be used to create rather narrow band transmittances for TE polarization. The simulations in Fig. 10 show the transmittance for two layer designs that use layers Metal 1 and Metal 2 or Metal 1 and Metal 3 and assume a 0.13-μm CMOS technology. The periodicities between the layers are an integer multiple of the periodicity within each layer, resulting

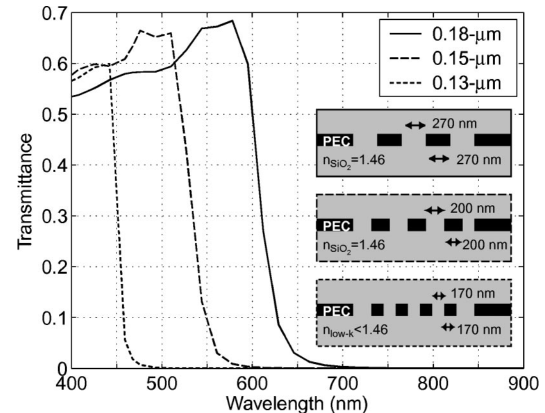
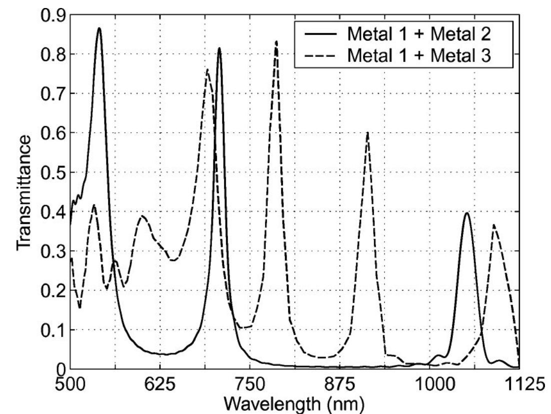


Fig. 9. Simulated transmittance of 1D ICPs assuming CMOS technologies with decreasing feature sizes (0.18-μm, solid curve; 0.15-μm, dashed curve; 0.13-μm, dotted curve). Simulations assume the TE polarization. The simulated patterned metal layer geometries are shown in the inset.



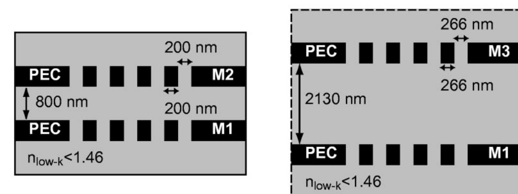


Fig. 10. (a) Simulated transmittance of 1D ICPs for stacked patterned metal layers in 0.13-μm CMOS technology (solid curve for layers Metal 1 and Metal 2, dashed curve for layers Metal 1 and Metal 3). Simulations assume TE polarization. (b) The simulated patterned metal-layer geometries.

in a powerful Fabry–Perot effect. This design produces narrow high-transmittance regions in a regime that is effectively zero transmittance in the one metal layer design (0.13-μm CMOS technology in Fig. 9). The Metal 1/Metal 2 simulation has a design with an integer multiple of 2 between the layer separations and wire periodicity, and the Metal 1/Metal 3 simulation has an integer multiple of 4. The consequence of this change is to produce additional high-transmittance peaks.

At optical wavelengths, perfect electrical conductor (PEC) is an unrealistic model for the patterned metal layers. The degradation in performance, however, is not severe. For example, we simulated the structure described in Fig. 7 as a perfect electrical conductor and as aluminum; the transmittance maintained its features but dropped by 20%.

D. Finite-Difference Time-Domain Method

The FDTD method is a numerical method for solving electromagnetic problems.[27,28] It has been applied extensively to simulating optical devices.[18,29–31] More specifically, it has been used to model pixels in the established CCD technology.[32] To the best of our knowledge, this is the first time the method has been applied to predicting the optical transmission in CMOS image-sensor pixels with their deeper structures and multiple metal layers.

Simulations in this paper were based on a commercially available 2D FDTD implementation.[22] The patterned metal layer, in the homogeneous medium or in the stratified dielectric medium, is specified in the x–z plane with electromagnetic field propagation taking place along the z axis (the y direction is assumed to be infinite). This assumption removes all the $\partial/\partial y$ derivatives from Maxwell's equations and allows a convenient split into two independent sets of equations for the TE and TM polarizations.

The sampling grid in space for the FDTD simulation domain is set to 10 nm in x and z, which corresponds to ten steps per wavelength in the medium with the highest refractive index for the smallest wavelength (380 nm). The sampling step in time is selected to ensure numerical stability of the FDTD algorithm and is automatically determined by the Courant limit,

$$\Delta t = \frac{1}{c[1/(\Delta x)^2 + 1/(\Delta z)^2]^{1/2}} = 1.7 \times 10^{-7} \text{ s}. \quad (8)$$

To reduce numerical dispersion, we use a total-field FDTD implementation as opposed to a scattered-field approach.[33] The location of the incident field divides the computational domain into scattered-field and total-field regions. UPML absorbing boundaries, which are based on a physical Maxwell formulation of an equivalent absorbing uniaxial anisotropic medium, truncate the simulation domain.[34] Minimization of the numerical reflectance at the UPMLs requires spatial scaling of the conductivity profile from zero (at the interface of the PMLs) to a maximum value at the end of the computational window, following

$$\sigma(x) = \sigma_{\max}(x/L)^m, \qquad \kappa(x) = 1 + (\kappa_{\max} - 1)(x/L)^m. \quad (9)$$

With choices of a UPML thickness of $L = 10$, a maximum tensor value of $\kappa_{\max} = 5$, and a polynomial power of $m = 3.5$, the absorbing boundaries have a theoretical reflection coefficient of 10^{-12}.[34]

We start an FDTD simulation by introducing a normally incident plane wave at the scattered-field–total-field interface $z = z_{\text{inc}}$,

$$E_y^{\text{inc}}(x, z_{\text{inc}}) = T_{\text{pulse}}(t)F(x, z_{\text{inc}})\sin(\omega t + \theta_i), \quad (10)$$

with pulsed time profile,

$$T_{\text{pulse}}(t) = \exp\left[-\frac{1}{2}\left(\frac{t - t_{\text{off}}}{t_0}\right)^2\right], \quad (11)$$

and transverse Gaussian space profile $F(x, z_{\text{inc}})$ chosen to avoid edge effects at the UPML. A pulsed excitation enables us to obtain the entire spectral behavior of the structure in a single FDTD simulation. The center frequency ω is chosen at 550 nm to cover the visible and the near-infrared wavelength range; the phase shift θ_i is the phase difference between points in the incidence plane, and this offset can be adjusted to define the direction of the incident field. The input plane $z = z_{\text{inc}}$ is chosen a few 100 nm above the metal layer. The Gaussian space profile $F(x, z_{\text{inc}})$ has a FWHM of 6 μm (for an aperture of 12 μm). The pulsed time profile is chosen to have significant optical power in the 380-to-1100-nm wavelength regime, i.e., the time offset $t_{\text{off}} = 10^{-14}$ s and the pulse width $t_0 = 2 \times 10^{-15}$ s. The time stepping continues until the desired late-time pulse response is observed at the field points of interest. We empirically selected 10^4 steps after which the pulse has either (a) left the domain on first pass (for the homogeneous-medium case) or (b) all multiple reflections have decayed (for the stratified-medium case).

The FDTD simulation data for determining transmittance are collected at an observation line located in the total-field region, i.e., below $z = z_{\text{inc}}$ and the metal layer. The fields propagated by the FDTD algorithm are time-domain fields. To get the full complex (amplitude and phase) wave information at the observation line below the metal layer, we calculate the complex electric and magnetic fields by taking a discrete Fourier transform of the full time trace. These complex field values are combined to compute the real power transmitted through the observation line. Since we are interested in the total transmitted power, we directly integrate the Poynting vector resulting from the electric ($\mathcal{E}$) and magnetic ($\mathcal{H}$) fields on the output plane (observation line) located a few 100 nm below the metal layer. We get expressions for the total power transmitted by the clear aperture (reference pixel)

$$P_{\text{without}} = \frac{1}{2}\text{Re}\left[\int_{\text{without}}(\mathcal{E} \times \mathcal{H}^*) \cdot \hat{z}\,\mathrm{d}x\,\mathrm{d}y\right] \quad (12)$$

and the aperture with wires (ICP)

$$P_{\text{with}} = \frac{1}{2}\text{Re}\left[\int_{\text{with}}(\mathcal{E} \times \mathcal{H}^*) \cdot \hat{z}\,\mathrm{d}x\,\mathrm{d}y\right]. \quad (13)$$

Any power contained in the evanescent field will be imaginary and will therefore not contribute to the Poynting vector. The transmittance of the patterned metal layer

in an ICP is defined as the ratio of the transmitted power with, P_{with}, to the transmitted power without, P_{without}, wires,

$$\mathcal{T}_{\text{FDTD}} = P_{\text{with}}/P_{\text{without}}. \tag{14}$$

We perform this calculation both for the patterned metal layer within a homogeneous dielectric medium (Fig. 7 inset) and for the more realistic case where the patterned metal layer is embedded in a stratified medium consisting of a stack of dielectrics [Fig. 8(a) inset].

The materials used in the FDTD simulations are either (a) homogeneous dielectric materials with constant refractive index, e.g., air ($n = 1$), Si_3N_4 ($n = 2.08$), SiO_2 ($n = 1.46$), and silicon ($n = 3.5$), or (b) dispersive metals, e.g., aluminum. At first we modeled the aluminum wires assuming a PEC model. The PEC model, for which $\mathcal{E}$ or $\mathcal{H}$ fields are zero inside the wires and for which the tangential and perpendicular components for the $\mathcal{E}$ and $\mathcal{H}$ fields at its boundary are zero as well, proved to be an unrealistic assumption and yielded too-large transmittance values. This is in agreement with the well-known observation that real metals do not act as PECs in the visible wavelength range. A more realistic three-parameter Drude model for aluminum proved to be more appropriate,

$$n = \sqrt{\epsilon_\infty - \frac{{\omega_p}^2}{\omega^2 + i\omega_\tau\omega}}, \tag{15}$$

where the relative permittivity at infinite frequency is $\epsilon_\infty = 1$, the plasma frequency is $\omega_p = 1.8 \times 10^{16}$ rad/s, and the collision frequency is $\omega_\tau = 1.1 \times 10^{15}$ rad/s. The values for these parameters were derived for the visible wavelength regime on the basis of published tabulated values for the complex refractive index of aluminum and following the procedure of Ordal.[35] This frequency-based dispersive material model was implemented in the time domain by using the polarization-equation approach with a single resonance.[22]

E. Integrated Color Pixels and Concepts from Waveguide and Grating Theory

The geometrical structure of the patterned metal layers is similar to that of conventional wire-grid transmission gratings. Considerable analytical electromagnetic-grating theory exists,[36] and in this section, we analyze how to apply this theory to ICPs. On the basis of these observations, we suggest how it may be possible to derive approximate closed-form formulas to predict the transmittance of the patterned metal layers.

Electromagnetic-grating theory divides the wavelength spectrum into several regimes with respect to transmission mechanism. For TE the boundaries between these wavelength regimes are set by the gap width w and the refractive index of the surrounding medium n. Above $2nw$ there is no TE transmission. Between nw and $2nw$ there is a single propagating waveguide mode. For the gap widths of the 1D patterned metal layers studied here, visible wavelengths fall between nw and $2nw$; above $2nw$ is the infrared.

The empirical three-parameter-model (Section 5) is inspired by this theory. In the region from nw to $2nw$, TE transmission is attenuated but approximately constant. As we pass from this region to wavelengths above $2nw$, TE transmission falls off. Given a value of $n = 1.46$ for the surrounding medium, and measured gap widths 236 ± 14 nm, the predicted TE falloff is 689 ± 40 nm and defines the boundary of the zero-transmittance regime. The estimated value with use of the three-parameter model, 677 nm [Fig. 8(a)], is surprisingly close to this value despite the facts that (a) the formulas for the boundaries between the transmission regimes are valid only for the PEC and (b) the measurements are a mixture of TE and TM signals. A similar agreement is observed for all pixels shown in Fig. 5(a), with predicted TE falloff at 689 ± 41, 724 ± 58, and 940 ± 82 nm, and the three-parameter model estimates yield 677, 750, and 875 nm.

According to electromagnetic-grating theory, the drop in the TM transmittance spectra differs from the cutoff observed in the TE transmittances. The drop in the TM transmittance is associated with a Rayleigh–Wood anomaly[37] rather than with a waveguide-mode cutoff. Rayleigh–Wood anomalies are a pure diffraction effect associated with a redistribution of energy and phase among several diffraction orders of a grating. They occur at wavelengths for which a transmission order becomes grazing and emerges parallel to the patterned metal layer.[38] For normal incidence, the anomaly wavelengths correspond to an integer number of periods of the pattern. For patterned metal layers with a duty cycle close to 50%, as is the case in the ICPs, the waveguide cutoff wavelength in the TE case and the Rayleigh–Wood anomaly wavelength in the TM case coincide [Fig. 8(b)]. In the TM case, the Rayleigh–Wood regime is followed by a transmittance regime with Fabry–Perot resonances occurring when the layer thickness equals a multiple of half the waveguide-mode wavelength.[39,40]

To the extent that waveguide modes and Rayleigh–Wood anomalies are the only important physical mechanisms governing the transmission spectra, it should be possible to develop a simple analytical theory. This suggestion can be confirmed only by further experiments with a range of materials and geometric configurations.

7. CONCLUSIONS

We implemented light filters that use patterned metal layers placed on top of each pixel's photodetector. These experiments extend the integration of optical functionality in CMOS image sensors. We designed and prototyped integrated color pixels in a standard 0.18-μm CMOS technology based on an active pixel sensor circuit. To demonstrate the wavelength selectivity of the light filters' design, we measured the transmittance through the patterned metal layers under various illumination conditions and found that they exhibit wavelength selectivity. We then performed (a) wave optics simulations to predict the spectral response of an uncovered reference pixel and (b) numerical electromagnetic simulations using a 2D finite-difference time-domain method to predict transmittances through 1D patterned metal layers. We found good agreement for both the spectral response and the transmittance. Finally, we used simulations to predict the transmittance for more elaborate designs, that is, fil-

ters using more than one metal layer and filters in more advanced CMOS technologies.

APPENDIX A

1. Scattering-Matrix Formalism

Consider a stratified medium along the z axis that consists of a 1D stack of m parallel layers, with infinite dimensions in x and y, sandwiched between two semi-infinite ambient (0) and substrate $(m + 1)$ media. Let all media be linear homogeneous and isotropic, let the complex index of refraction and the thickness of the jth layer be n_j and d_j, respectively, and let n_0 and n_{m+1} represent the complex indices of refraction of the ambient and the substrate media. A monochromatic plane wave incident on the parallel layers, originating from the ambient, generates a resultant transmitted plane wave in the substrate. We are now interested in determining the amplitude of the resultant wave. The intuitive method of adding multiple reflected and transmitted waves quickly becomes impractical even for a few dielectric layers. A more elegant approach can be taken that is based on the fact that the equations that govern the propagation of light are linear and that continuity of the tangential components of the light fields across an interface between two isotropic media can be regarded as a 2×2 linear-matrix transformation. This method, referred to as the scattering-matrix approach, was pioneered by Abelès.[41,42] The present development is due to Hayfield and White.[43] The method here is essentially equivalent to the scattering matrix method in microwave theory.[44]

The total field inside the jth layer, which is excited by the incident plane wave, consists of two plane waves: a forward-traveling and a backward-traveling plane wave with complex amplitudes $\mathcal{E}^+$ and $\mathcal{E}^-$, respectively. The total field in a plane z, parallel to the boundary, can be described by a 2×1 column vector,

$$\mathcal{E}(z) = \begin{pmatrix} \mathcal{E}^+(z) \\ \mathcal{E}^-(z) \end{pmatrix}. \tag{A1}$$

The fields at two different planes $\mathcal{E}(z')$ and $\mathcal{E}(z'')$ are related by a 2×2 matrix transformation,

$$\begin{pmatrix} \mathcal{E}^+(z') \\ \mathcal{E}^-(z') \end{pmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{pmatrix} \mathcal{E}^+(z'') \\ \mathcal{E}^-(z'') \end{pmatrix}. \tag{A2}$$

The 2×2 matrix defined by the planes immediately adjacent to the 01 and $m(m + 1)$ interfaces is called the scattering matrix $\mathbf{S}$ of the stratified structure. The scattering matrix can be expressed as a product of interface and layer matrices $\mathbf{I}$ and $\mathbf{L}$,

$$\mathbf{S} = \mathbf{I}_{01}\mathbf{L}_1 \ldots \mathbf{I}_{(j-1)j}\mathbf{L}_j \ldots \mathbf{L}_m \mathbf{I}_{m(m+1)}. \tag{A3}$$

The matrix $\mathbf{I}$ of an interface between two media relates the fields on both of its sides when Fresnel's reflection and transmission coefficients $t_{(j-1)j}$ and $r_{(j-1)j}$ are used for the interface:

$$\mathbf{I}_{(j-1)j} = \frac{1}{t_{(j-1)j}} \begin{bmatrix} 1 & r_{(j-1)j} \\ r_{(j-1)j} & 1 \end{bmatrix}. \tag{A4}$$

The formulas to compute the Fresnel coefficients can be found in many optics textbooks[45]; they depend on the complex indices of refraction of the materials used in the CMOS process. The matrix $\mathbf{L}$ describes the propagation through a layer with index of refraction n_j and thickness d_j

$$\mathbf{L}_j = \begin{bmatrix} \exp(j\beta_j) & 0 \\ 0 & \exp(-j\beta_j) \end{bmatrix}, \tag{A5}$$

where the phase shift, i.e., the layer phase thickness, is given by $\beta_j = (2\pi/\lambda) d_j n_j \cos\theta_j$. The overall reflection, R, and transmission, T, coefficients of the stratified structure are

$$R = S_{21}/S_{11}, \qquad T = 1/S_{11}. \tag{A6}$$

2. CMOS Photodiode Model

The photodetector converts incident spectral irradiance $E_0(\lambda)$, expressed in W/(m^2 nm), into photocurrent that is proportional to the radiant power. The conversion is done in two steps: (a) incident photons generate electron-hole (e-h) pairs in the silicon substrate, and (b) some of the generated carriers are converted into photocurrent. Since photocurrent is very small (10s to 100s of femtoamperes), it is typically integrated into charge and then converted to voltage before it is read out. In what follows, we quantify the rate of e-h-pair generation in silicon for photons in the visible wavelengths, we describe how a photodetector converts generated carriers into current, and we derive a simplified equation for the photocurrent and its conversion into charge/voltage by direct integration.

1. Electron-Hole-Pair Generation

An incident photon reaching the silicon substrate needs to have at least the bandgap energy of 1.124 eV (or wavelength below 1120 nm) to excite an electron to jump from the valence band to the conduction band, leaving a hole behind and generating an e-h pair. Photons in the visible wavelength regime have enough energy to generate e-h pairs. However, no photon in the visible can generate more than one e-h pair.

Assume incident irradiance $E_0(\lambda)$ at the surface of the silicon substrate, or if the substrate is covered with a stratified structure (as is the case for a photodetector in a CMOS image-sensor pixel), assume transmitted irradiance $|T|^2 E_0(\lambda)$ at the silicon substrate interface. If $E(\lambda, z)$ is the irradiance at depth z with respect to the silicon surface, then the number of photons absorbed per second between z and $z + \mathrm{d}z$ is given by

$$\frac{hc}{\lambda} \frac{\mathrm{d}E(\lambda, z)}{\mathrm{d}z} = \frac{\mathrm{d}F(\lambda, z)}{\mathrm{d}z} = -\alpha(\lambda) F(\lambda, z), \tag{A7}$$

where $F(\lambda, z)$ is the photon flux, expressed in photons/(s m^2 nm), at depth z and $\alpha(\lambda)$ is the wavelength-dependent absorption coefficient derived from silicon's complex index of refraction.[46,47] Solving this equation, we obtain that

$$F(\lambda, z) = F_0(\lambda) \exp[-\alpha(\lambda) z]. \tag{A8}$$

Thus the rate of e-h pairs generated at z is

$$G(\lambda, z) = \frac{\mathrm{d}}{\mathrm{d}z}[F_0(\lambda) - F(\lambda, z)] = \alpha(\lambda)F_0(\lambda)\exp[-\alpha(\lambda)z]. \tag{A9}$$

2. Current Generation

There are several types of photodetectors in CMOS technology. The most commonly used is the photodiode, which is a reverse biased pn junction. In our chip we used an $n+/p$sub photodiode. Surprisingly, the junction depths of CMOS photodiodes are quite consistent with the absorption depths for visible light, although as CMOS technology scales beyond 0.35 μm, junctions are becoming too shallow for complete absorption. Assuming the depletion approximation, we can calculate the photocurrent i_{ph} as the sum of three components: (a) a current due to generation in the depletion region $i_{\mathrm{ph}}^{\mathrm{sc}}$, where almost all carriers generated are swept away by the strong electric field; (b) a current due to the holes generated in the n-type quasi-neutral region i_{ph}^{p}, where some diffuse to the space charge region and get collected; and (c) a current due to the electrons generated in the p-type region i_{ph}^{n}. We derive an explicit expression for i_{ph} under the following (standard) simplifying assumptions: (a) abrupt pn junction, (b) depletion approximation, (c) low-level injection, i.e., flux-induced carrier densities are much smaller than the majority carrier densities, (d) short base-region approximation, which is quite acceptable for advanced CMOS processes where junction depths are much smaller than diffusion lengths. Details of the derivation can be found elsewhere.[48] We only show the result here. Assuming that all generated carriers in the space charge region are collected, the current density (in A/m^2) due to generation in the space charge region is

$$j_{\mathrm{ph}}^{\mathrm{sc}}(\lambda) = qF_0(\lambda)\{\exp[-\alpha(\lambda)z_1] - \exp[-\alpha(\lambda)z_2]\}. \tag{A10}$$

The diffusion current density due to the generation in the n-type quasi-neutral region is

$$j_{\mathrm{ph}}^{p}(\lambda) = -\frac{qF_0(\lambda)}{\alpha(\lambda)z_1}\{[\alpha(\lambda)z_1 + 1]\exp[-\alpha(\lambda)z_1] - 1\}, \tag{A11}$$

and similarly for the p-type quasi-neutral region:

$$j_{\mathrm{ph}}^{n}(\lambda) = \frac{qF_0(\lambda)}{\alpha(\lambda)(z_3 - z_2)}\{[\alpha(\lambda)(z_3 - z_2) - 1] \times \exp[-\alpha(\lambda)z_2] + \exp[-\alpha(\lambda)z_3]\}. \tag{A12}$$

We then need only reasonable numbers for z_1, z_2, and z_3 and the photodiode area in order to determine the photocurrent density, which is the sum of the three aforementioned components:

$$j_{\mathrm{ph}}(\lambda) = \frac{qF_0(\lambda)}{\alpha(\lambda)}\left\{\frac{1 - \exp[-\alpha(\lambda)z_1]}{z_1} - \frac{\exp[-\alpha(\lambda)z_2] - \exp[-\alpha(\lambda)z_3]}{z_3 - z_2}\right\} \tag{A13}$$

or

$$j_{\mathrm{ph}}(\lambda) = \frac{q}{\alpha(\lambda)^2}\left[\frac{G(\lambda, 0) - G(\lambda, z_1)}{z_1} - \frac{G(\lambda, z_2) - G(\lambda, z_3)}{z_3 - z_2}\right]. \tag{A14}$$

3. Spectral Response Model

For the spectral response, we put the photocurrent together with the light transmitted into the silicon substrate,

$$SR(\lambda) = \frac{j_{\mathrm{ph}}(\lambda)}{E(\lambda)} = \frac{qhc}{\lambda}\frac{|T(\lambda)|^2}{\alpha(\lambda)}\left\{\frac{1 - \exp[-\alpha(\lambda)z_1]}{z_1} - \frac{\exp[-\alpha(\lambda)z_2] - \exp[-\alpha(\lambda)z_3]}{z_3 - z_2}\right\}, \tag{A15}$$

or, equivalently, we can get the (external) quantum efficiency

$$QE(\lambda) = \frac{|T(\lambda)|^2}{\alpha(\lambda)}\left\{\frac{1 - \exp[-\alpha(\lambda)z_1]}{z_1} - \frac{\exp[-\alpha(\lambda)z_2] - \exp[-\alpha(\lambda)z_3]}{z_3 - z_2}\right\}. \tag{A16}$$

The external quantum efficiency is then related to the internal quantum efficiency by the transmittance,

$$QE(\lambda) = |T(\lambda)|^2 QE_{\mathrm{internal}}(\lambda). \tag{A17}$$

The transmittance $|T(\lambda)|^2$ is determined in part by the 1D layered dielectric stack (i.e., dielectric materials and layer thickness) and in part by the 2D structure within each layer (i.e., the geometry of nearby metal wires obstructing light and the lateral geometry of the photodiode). The former can be dealt with by using a scattering-matrix approach, as we do here, and is sufficient provided that the diode is large (e.g., 12 μm) and subject to normal incident-plane-wave illumination (see Fig. 4). The latter is the reason for pixel vignetting and can be dealt with by a phase-space approach, which we discuss elsewhere, and is more prominent when pixel size shrinks to a few micrometers and is subject to oblique incident wide-angle illumination.[9]

ACKNOWLEDGMENTS

This work is supported by the Programmable Digital Camera (PDC) project and by Philips Semiconductors. P. B. Catrysse is "Aspirant" with the Fund for Scientific Research—Flanders (Belgium). We thank B. Fowler and M. Godfrey for their valuable help on the ICP implemen-

tation; M. Peeters, J. Zhang, and S. Fan for assistance with the FDTD simulations; A. El Gamal, S. Fan, J. W. Goodman, D. A. B. Miller, and I. Veretennicoff for insightful comments. We also thank R. Motta for the original 3D pixel drawings and J. DiCarlo for assistance with the automated experimental setup.

Author contact information: Peter B. Catrysse (c/o Wandell Lab), Jordan Hall, Building 420, Room 490, Stanford University, Stanford, California 94305-2130. Phone, 650-725-1255; fax, 650-322-5379; e-mail: pcatryss @stanford.edu.

REFERENCES

1. K. A. Parulski, "Color filters and processing alternatives for one-chip cameras," IEEE Trans. Electron Devices **ED-32**, 1361–1389 (1985).
2. A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, Solid-State Science and Technology Library (Kluwer Academic, Boston, Mass., 1995), p. 388.
3. A. El Gamal, D. Yang, and B. Fowler, "Pixel level processing—why, what, and how?" in *Sensors, Cameras, and Applications for Digital Photography*, N. Sampat and T. Yeh, eds. (SPIE Press, Bellingham, Wash., 1999), pp. 2–13.
4. P. B. Catrysse, B. A. Wandell, and A. El Gamal, "An integrated color pixel in 0.18 μm CMOS technology," in *2001 International Electron Devices Meeting—Technical Digest* (Institute of Electrical and Electronics Engineers, New York, 2001), pp. 559–562.
5. J. Adams, K. Parulski, and K. Spaulding, "Color processing in digital cameras," IEEE Micro **18**, 20–30 (1998).
6. P. Dillon, D. Lewis, and F. Kaspar, "Color imaging using a single CCD array," IEEE Trans. Electron Devices **ED-25**, 102–107 (1978).
7. B. E. Bayer, "Color imaging array," U.S. Patent 3,971,065 (July 20, 1976).
8. H. Nabeyama, "All-solid-state color camera with single-chip MOS imager," IEEE Trans. Consumer Electron. **CE-27**, 40–45 (1981).
9. P. B. Catrysse and B. A. Wandell, "Optical efficiency of image sensor pixels," J. Opt. Soc. Am. A **19**, 1610–1620 (2002).
10. A. J. Blanksby and M. J. Loinaz, "Performance analysis of a color CMOS photogate image sensor," IEEE Trans. Electron Devices **47**, 55–64 (2000).
11. W. Li, P. Ogunbona, S. Yu, and I. Kharitonenko, "Modelling of color cross-talk in CMOS image sensors," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* (Institute of Electrical and Electronics Engineers, New York, 2002), pp. 3576–3579.
12. P. B. Catrysse, X. Liu, and A. El Gamal, "Quantum efficiency reduction due to pixel vignetting in CMOS image sensors," in *Sensors and Camera Systems for Scientific, Industrial and Digital Photography Applications*, M. M. Blouke, N. Sampat, G. M. Williams, Jr., and T. Yeh, eds. (SPIE Press, Bellingham, Wash., 2000), pp. 420–430.
13. M. Böhm and H. Stiebig, "Trichromatic sensor," U.S. Patent 5,923,049 (July 13, 1999).
14. R. B. Merrill, "Color separation in an active pixel cell imaging array using a triple-well strucuture," U.S. Patent 5,965,875 (October 12, 1999).
15. H. Hertz, "Ueber Strahlen electrischer Kraft," Annal. Phys. Chem. **36**, 769–783 (1889).
16. P. J. Bliek, L. C. Botten, R. Deleuil, R. C. McPhedran, and D. Maystre, "Inductive grids in the region of diffraction anomalies: theory, experiments and applications," IEEE Trans. Microwave Theory Tech. **28**, 1119–1125 (1980).
17. H. Tamada, T. Doumuki, T. Yamaguchi, and S. Matsumoto, "Al wire-grid polarizer using the *s*-polarization resonance effect at the 0.8-μm-wavelength band," Opt. Lett. **22**, 419–421 (1997).
18. M. A. Jensen and G. P. Nordin, "Finite-aperture wire grid polarizers," J. Opt. Soc. Am. A **17**, 2191–2198 (2000).
19. E. R. Fossum, "Active pixel sensors: are CCD's dinosaurs?" in *Charge-Coupled Devices and Solid State Optical Sensors III*, M. M. Blouke, ed. (SPIE Press, Bellingham, Wash., 1993), pp. 2–14.
20. B. Fowler, A. El Gamal, D. Yang, and H. Tian, "A method for estimating quantum efficiency for CMOS image sensors," in *Solid State Sensor Arrays: Development and Applications II*, M. M. Blouke, ed. (SPIE Press, Bellingham, Wash., 1998), pp. 178–185.
21. H. Tian, X. Q. Liu, S. H. Lim, S. Kleinfelder, and A. El Gamal, "Active pixel sensors fabricated in a standard 0.18 μm CMOS technology," in *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications II*, M. M. Blouke, J. Canosa, and N. Sampat, eds. (SPIE Press, Bellingham, Wash., 2001), pp. 441–449.
22. OptiFDTD 2.0, Optiwave Corporation, Ottawa, Canada, 2001.
23. D. Yang, H. Min, B. Fowler, A. El Gamal, M. Beiley, and K. Cham, "Test structures for characterization and comparative analysis of CMOS image sensors," in *Advanced Focal Plane Arrays and Electronic Cameras*, T. M. Bernard, ed. (SPIE Press, Bellingham, Wash., 1996), pp. 8–17.
24. E. Roca, F. Frutos, S. Espejo, R. Dominguez-Castro, and A. Rodríguez-Vázquez, "Electrooptical measurement system for the DC characterization of visible detectors for CMOS-compatible vision chips," IEEE Trans. Instrum. Meas. **47**, 499–506 (1998).
25. P. B. Catrysse, A. El Gamal, and B. A. Wandell, "Comparative analysis of color architectures for image sensors," in *Sensors, Cameras, and Applications for Digital Photography*, N. Sampat and T. Yeh, eds. (SPIE Press, Bellingham, Wash., 1999), pp. 26–35.
26. T. Lulé, S. Benthien, H. Keller, F. Mütze, P. Rieve, K. Siebel, M. Sommer, and M. Böhm, "Sensitivity of CMOS based imagers and scaling perspectives," IEEE Trans. Electron Devices **47**, 2110–2122 (2000).
27. K. Shlager and J. Schneider, "A selective survey of the finite-difference time-domain literature," IEEE Trans. Antennas Propag. Mag. April 1995, pp. 39–56.
28. A. Taflove and S. C. Hagness, *Computational Electrodynamics: the Finite-Difference Time-Domain Method*, 2nd ed. (Artech House, Norwood, Mass., 2000).
29. A. Taflove, "Review of the formulation and applications of the finite-difference time-domain method for numerical modeling of electromagnetic wave interactions with arbitrary structures," Wave Motion **10**, 547–582 (1988).
30. E. A. Navarro, B. Gimeno, and J. L. Cruz, "Modelling of periodic structures using finite difference time domain method combined with the Floquet theorem," Electron. Lett. **29**, 446–447 (1993).
31. K. H. Dridi and A. Bjarklev, "Optical electromagnetic vector-field modeling for the accurate analysis of finite diffractive structures of high complexity," Appl. Opt. **38**, 1668–1676 (1999).
32. T. O. Körner and R. Gull, "Combined optical/electric simulation of CCD cell structures by means of the finite-difference time-domain method," IEEE Trans. Electron Devices **47**, 931–938 (2000).
33. R. Holland and J. W. Williams, "Total-field versus scattered-field finite-difference codes: a comparative assessment," IEEE Trans. Nucl. Sci. **NS-30**, 4583–4588 (1983).
34. S. D. Gedney, "An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD Lattices," IEEE Trans. Antennas Propag. **44**, 1630–1639 (1996).
35. M. A. Ordal, "Optical properties of metals Al, Co, Cu, Au, Fe, Pb, Ni, Pd, Pt, Ag, Ti, and W in the infrared and far infrared," Appl. Opt. **22**, 1099–1119 (1983).
36. R. Petit, *Electromagnetic Theory of Gratings*, Vol. 22 of Topics in Current Physics (Springer-Verlag, Berlin, 1980).
37. R. W. Wood, "Anomalous diffraction gratings," Phys. Rev. **48**, 928–936 (1935).
38. A. Hessel and A. A. Oliner, "A new theory of Wood's anomalies on optical gratings," Appl. Opt. **4**, 1275–1297 (1965).
39. P. Lalanne, J. P. Hugonin, S. Astilean, M. Palamaru, and K.

D. Moller, "One-mode model and Airy-like formulae for one-dimensional metallic gratings," J. Opt. A, Pure Appl. Opt. **2**, 48–51 (2000).
40. Q. Cao and P. Lalanne, "Negative role of surface plasmons in the transmission of metallic gratings with very narrow slits," Phys. Rev. Lett. **88**, 57403-1–57403-4 (2002).
41. F. Abelès, "Recherches sur la propagation des ondes électromagnétiques sinusoïdales dans les milieux stratifiés: application aux couches minces," Ann. Phys. **5**, 596–640 (1950).
42. F. Abelès, "Recherches sur la propagation des ondes électromagnétiques sinusoïdales dans les milieux stratifiés: application aux couches minces," Ann. Phys. **5**, 706–782 (1950).
43. P. C. S. Hayfield and G. W. T. White, "An assessment of the suitability of the Drude–Tronstad polarized light method for the study of film growth on polycrystalline metals," in *Ellipsometry in the Measurement of Surfaces and Thin Films*, N. M. Bashara, A. B. Buckman, and A. C. Hall, eds. (National Bureau of Standards, Washington, D.C., 1964), pp. 157–200.
44. D. M. Pozar, *Microwave Engineering*, 2nd ed. (Wiley, New York, 1997).
45. M. Born and E. Wolf, *Principles of Optics*, 6th (corrected) ed. (Pergamon, Oxford, UK, 1980), pp. 38–41.
46. M. V. Klein and T. E. Furtak, *Optics*, 2nd ed., Wiley Series in Pure and Applied Optics (Wiley, New York, 1986), p. 71.
47. J. H. Weaver and H. P. R. Frederikse, eds., *Optical Properties of Metals and Semiconductors*, 74th ed., CRC Handbook of Chemistry and Physics (CRC, Boca Raton, Fla., 2000), pp. 12–109 and 112–131.
48. A. El Gamal, "EE392B: introduction to image sensors and digital cameras" (2001), retrieved 2002, http://www.stanford.edu/class/ee392b.