Image analysis using modulated light sources

Feng Xiao[a*†], Jeffrey M. DiCarlo[b], Peter B. Catrysse[b], Brian A. Wandell [a]
[a]Dept. of Psychology, Stanford University, CA 94305, USA
[b]Dept. of Electrical Engineering, Stanford University, CA 94305, USA

ABSTRACT

With the development of high-speed CMOS imagers, it is possible to acquire and process multiple images within the imager, prior to output. We refer to an imaging architecture that acquires a collection of images and produces a single result as *multiple capture single image* (MCSI). In this paper we describe some applications of the MCSI architecture using a monochrome sensor and modulating light sources. By using active (modulating) light sources, it is possible to measure object information in a manner that is independent of the passive (ambient) illuminant. To study this architecture, we have implemented a test system using a monochrome CMOS sensor and several arrays of color LEDs whose temporal modulation can be precisely controlled. First, we report on experimental measurements that evaluate how well the active and passive illuminant can be separated as a function of experimental variables, including passive illuminant intensity, temporal sampling rate and modulation amplitude. Second, we describe two applications of this technique: (a) creating a color image from a monochrome sensor, and (b) measuring the spatial distribution of the passive illuminant.

Keywords: Active imaging, digital camera, illuminant estimation, color imaging

1. INTRODUCTION

In biological systems, perception is often linked to action [1]. For example, the human perceptual system appears to contain sensory pathways that are coupled to specific motor goals [2]. The active aspects of human behavior include changes in head and eye position, or adjustments of the sensitivity of the pupil, lens, and visual sensitivity in response to changes in the ambient illumination. Bats emit signals that help them extend the capabilities of their auditory system.

We call an imaging method that improves image understanding through the use of sensor motion or the emission of specific signals an *active imaging method* (AIM). In engineering systems, many methods can serve as active aids to sensory performance. It is possible, for example, to introduce a variety of different signals (e.g., sonar and radar) into the environment to improve sensory measurement. Active imaging methods that add light to a scene can also assist object recognition or detection [3, 4]. An interesting example is the use of spatially modulated illumination to reduce false matches in stereo imaging for depth recovery [5, 6]. Recently, Ando's group [7, 8] has described experiments with a specialized monochrome CMOS sensor that uses time-domain correlation coupled with monochrome illuminant modulation. In this active imaging method, the sensor performs multiple captures and then calculates a single image that represents the scene as if it were illuminated only by the known modulating source.

We have decided to explore the application of active imaging methods with modulating light sources to color imaging. Various color imaging applications can benefit from acquiring image data under a known illumination source, and this can be achieved through such active imaging methods. Here, we describe an experimental apparatus we have built to investigate these applications. We also describe two examples that illustrate the applicability of the method. In one application we use several modulating light sources to derive a color image from a monochrome sensor. In another application we use modulating light sources to estimate the spatial distribution of the ambient illumination.

2. ACTIVE IMAGING THEORY

In this section we develop notation to characterize active imaging method applied to color imaging. The analyses are developed for the case of a temporally modulated illumination source that is incident on a small, uniform surface patch. We show how AIM can be used to derive scene properties from the time-domain correlation of multiple images.

---

[*] For a color version of this paper please go to ftp://white.stanford.edu/users/brian/pdc/spiemodulate.pdf

[†] Correspondence: Email: fxiao@stanford.edu; Telephone: 650 725 1255; Fax: 650 723 0993

Consider a small surface patch located at position x. Suppose the total illumination at location x is comprised of two parts: an active illuminant, $A(x,\lambda,t)$, that is under our control, and a passive (ambient) illumination, $P(x,\lambda,t)$, that is naturally present in the environment. Further suppose that the active illuminant may be comprised of the sum of multiple active illuminants $A_i(x,\lambda,t)$. Hence, the total illumination at a location x is given by

$$\begin{aligned} E(x,\lambda,t) &= P(x,\lambda,t) + A(x,\lambda,t) \\ &= P(x,\lambda,t) + \sum_{i=0}^{N} A_i(x,\lambda,t) \\ &= P(x,\lambda,t) + A_0(x,\lambda) + \sum_{i=1}^{N} A_i(x,\lambda,t) \end{aligned} \qquad (1)$$

The A_0 term is the mean level of the sum of the active illuminants; this mean level is constant over time. The A_i terms are modulations of the active illuminant about this mean level. In most cases, the passive illumination does not change over time, so that $P(x,\lambda,t)=P(x,\lambda)$. Consequently, we may simplify the equation to be

$$E(x,\lambda,t) = (P(x,\lambda) + A_0(x,\lambda)) + \sum_{i=1}^{N} A_i(x,\lambda,t) \qquad (2)$$

In most applications it is desirable to arrange the modulation of the active illuminant so that its relative spectral composition is unchanged as the intensity modulates. This results in the further simplification of the modulating term,

$$E(x,\lambda,t) = (P(x,\lambda) + A_0(x,\lambda)) + \sum_{i=1}^{N} A_i(x,\lambda) F_i(t) \qquad (3)$$

We call the temporal functions $F_i(t)$ the *modulation control functions*. The system designer chooses these functions, and it simplifies subsequent calculations when they are designed to form an orthogonal set. Moreover, because we have separated out the mean, all of these functions have a zero mean:

$$\int F_i(t) F_j(t) dt = \delta_{i,j} \quad \& \quad \int F_i(t) dt = 0 \qquad (4)$$

Next consider the response of an ideal camera with spectral sensitivity $R(\lambda)$ measuring light scattered from a surface patch with reflectance $S(x,\lambda)$. The responses of the camera sensors over time will vary according to

$$r(x,t) = \int R(\lambda) E(x,\lambda,t) S(x,\lambda) d\lambda \qquad (5)$$

We can demodulate the temporal signal of the camera measurements, $r(x,t)$, to calculate a signal associated with each of the different modulated light sources. Namely, to calculate the i^{th} demodulated image we calculate the inner product of the camera response and the i^{th} modulation control function.

$$\begin{aligned} r_i(x) &= \int r(x,t) F_i(t) dt \\ &= \int R(\lambda) S(x,\lambda) \int E(x,\lambda,t) F_i(t) dt \\ &= \int R(\lambda) S(x,\lambda) A_i(x,\lambda) d\lambda \end{aligned} \qquad (6)$$

Hence, the i^{th}-demodulated image depends only on light emitted by one of the active illuminants; the image is independent of the passive illumination and other active illuminants.

3. EXPERIMENTS

3.1. System setup

The basic theory of using AIM to modulate light sources is straightforward. But in practice, many implementation details and noise may limit its accuracy. Thus, we have developed an experimental apparatus to explore the advantages and limitations of AIM for color imaging. Several key elements of the laboratory are shown in Figure 1. The digital imager is a custom-designed 1/3'' CMOS monochrome digital pixel sensor [9] that is controlled by an FPGA (Actel). The scene is imaged through a Computar TV zoom lens with 8-16 mm focal length. The modulated light sources are provided by three LED arrays. Each array is composed of 5x7 identical LEDs whose intensity is controlled using a National Instruments AT-AO-10 DAQ board. With this board, we can create any waveform that we want for the LED arrays. A LCD projector (Sony, Model VPL-X1000) provides a passive illumination whose intensity and spatial distribution can be programmed. A Macbeth chart and other simple targets are used in the experiments. All of the experimental devices are controlled by a high level interface written in Matlab.

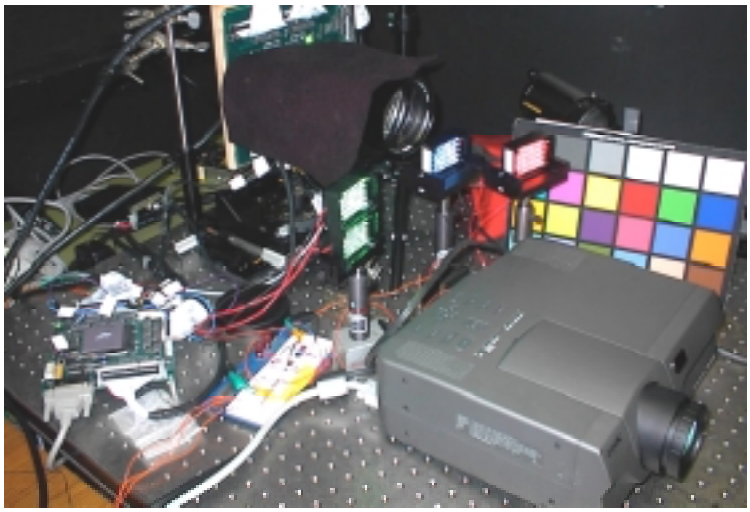

Figure 1: Apparatus used to perform AIM experiments. Images were captured using a digital pixel sensor (DPS) with a large aperture lens. The capture properties of the image sensor (integration time, sampling rate) were controlled by a FPGA using a simple assembler language. LED arrays provided the active illuminants. A calibrated LCD projector provided the passive illumination. Image acquisition and illumination controls were managed from a computer that contained a DAC, frame-grabber, and Matlab software.

The spectral properties of various system elements are shown in Figure 2. These spectral functions were measured using narrowband light sources provided an Oriel monochromator and a PhotoResearch PR-650 spectral radiometer. The spectral responsivity of the CMOS sensor array is shown in panel (a). The spectral power distributions of the three types of LED arrays are shown in panel (b).

3.2. Independence of passive illuminant intensity

In the ideal case, the response to the active illuminant should not be affected by the presence of a steady passive light source (Equation 6). How well does this property hold for a real camera system that includes noise, quantization, and other limitations?

To answer this question, we estimated the intensity of a fixed active illuminant modulation as we varied the passive illuminant intensity. We used the blue LED array as the active illuminant (though the results should be independent of the choice of LED array) and the LCD projector to vary the passive light intensity. We measured light reflected from a white

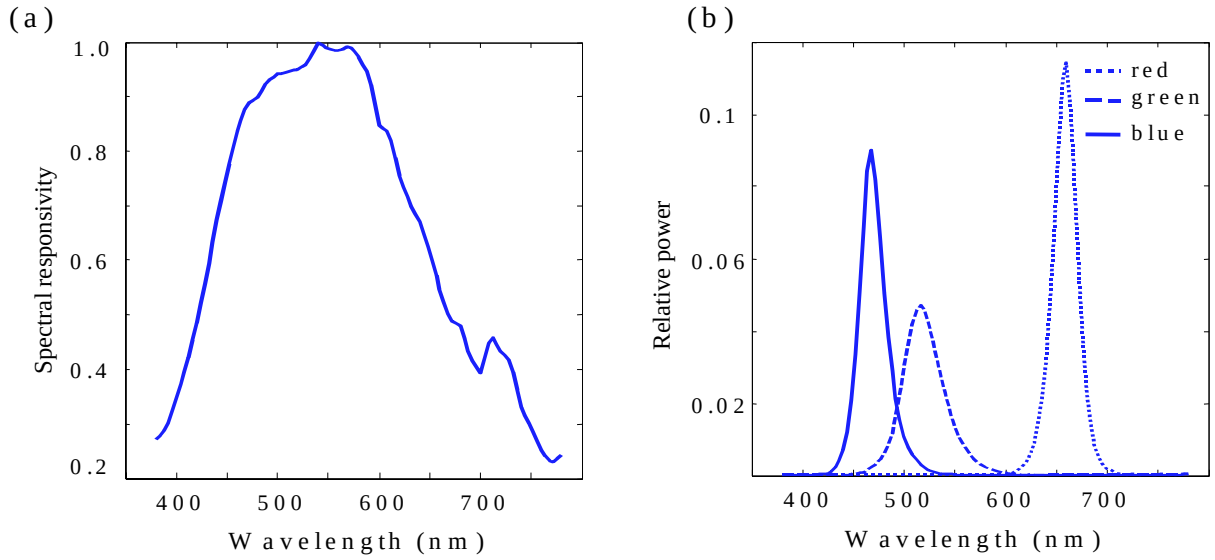

Figure 2: Characterization of the apparatus. (a) The spectral responsivity of the CMOS sensor. (b) The spectral power distributions of the LEDs.

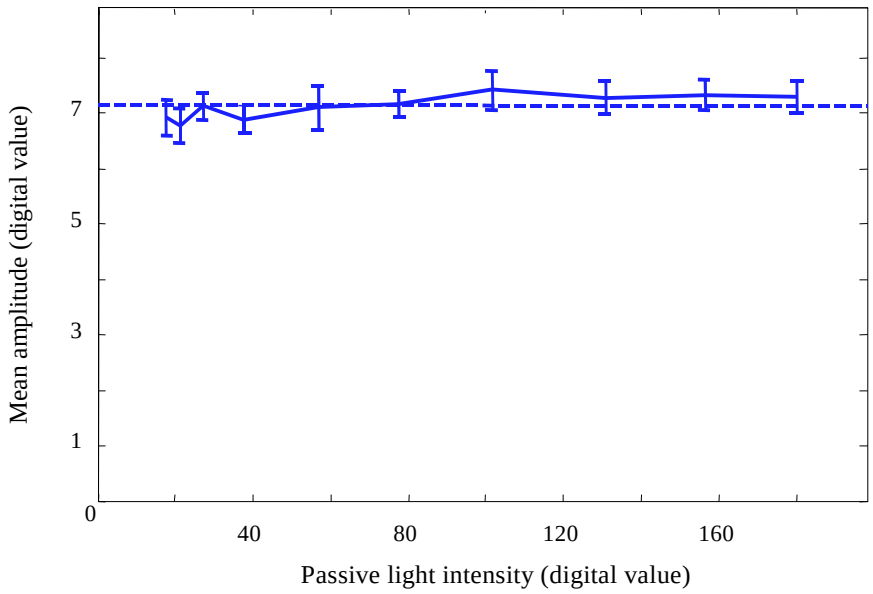

Figure 3: Response to active illuminant as the passive illuminant intensity increases. The blue LED array modulated at 1 Hz with a 14 digital counts peak-to-peak swing. The horizontal axis measures the passive illuminant intensity (digital counts from the linear sensor). The vertical axis measures the amplitude of response to active illuminant. Dash line shows the expected response and solid line shows measured response. Error bar shows the standard error of measurement

patch on the Macbeth chart (9x9 pixels) that was illuminated by both the passive and active illuminants. At each passive intensity level, the blue LED array modulated at 1 Hz in sinusoid waveform; we acquired 8 pictures with an exposure time

(ΔT) of 50 ms and an inter-sample interval (ΔI) of 75 ms. The 8 sample measurements were demodulated to estimate the response to the blue LED array alone. Figure 3 shows the amplitude of the estimated blue LED array modulation at ten different passive illuminant intensity levels. Although the passive light intensity varied from 3 to 25 times the intensity of active illuminant, the estimated value remained relatively constant. The standard error of the amplitude estimate from the 81 pixels in the patch was 0.6 digital counts.

3.3. Effects of sample number

Increasing the exposure time or increasing the number of samples per cycle both improve the accuracy of the estimate. Yet, within a fixed measurement period these two parameters oppose one another: longer exposure times imply fewer samples per cycle. Hence, we measured how the tradeoff between sample number and exposure time influences the accuracy of the estimated active illuminant. The second experiment analyzes this tradeoff for our particular measurement system. In this experiment, we again modulated the blue LED array at 1 Hz, but this time we kept a constant level of passive illumination. In our first measurement, we increased the number of samples by fixing the interval between exposures (ΔI) and decreasing the exposure time (ΔT). Figure 4a shows that increasing the number of samples this way increased the estimation error. Hence, for this imaging system it is more important to have long exposure durations than multiple samples. Figure 4b shows that fixing the exposure time but decreasing the interval between exposures, improves performance, as expected.

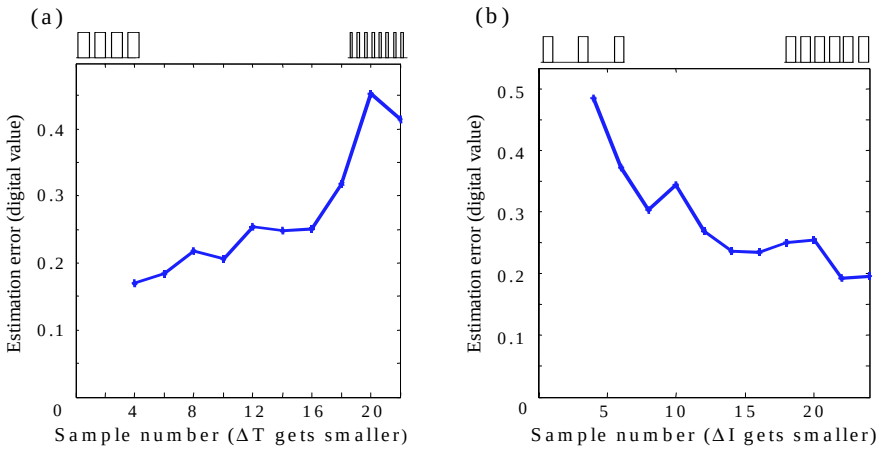

Figure 4: The trade-off between number of samples and exposure duration. The horizontal axis measures the number of samples measured. The vertical axis represents the standard error of the estimated modulation (in digital counts). In panel (a) more samples were obtained by decreasing the exposure duration. In panel (b) more samples were obtained by varying the time interval between the image exposures. For this imaging system, improving the image by extending the exposure duration was more valuable than increasing the number of independent samples of the modulating active illuminant.

3.4. Effect of the intensity of modulated light sources

Finally, we estimated the performance of the system in response to variations of the active illuminant amplitude. Using the blue LED array (1Hz), we varied the illuminant amplitude from a peak-to-peak near zero to a peak-to-peak of 50. Images were acquired from a 9x9 patch of the white Macbeth surface. The amplitude and standard error were estimated using all 81 pixels. The third experiment investigates how the standard error varies as a function of modulation amplitude. The smooth curve in Figure 5 shows the measured error; it is large for modulations less than 5 digital counts and then remains roughly constant for modulations above 20. The dashed curve in Figure 5 shows the expected standard error based on simulations that take into account the quantization error of the imager.

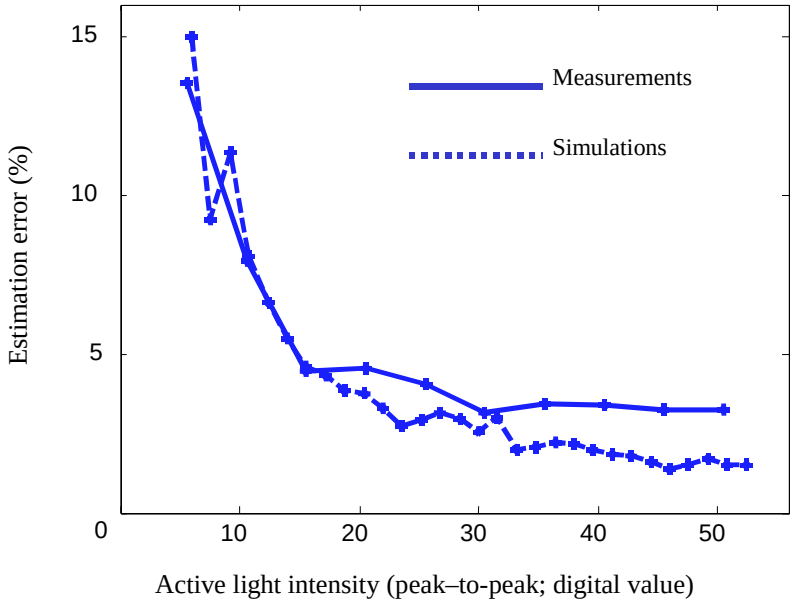

Figure 5: Estimation error for the modulated light source. The horizontal axis measures the blue LED intensity (digital counts from the linear sensor). The vertical axis represents the standard error of the estimated modulation (in percentage). The smooth and dashed curves show experimental and simulated values.

3.5 Selecting the modulation control functions

All modulating control functions are not equivalent. In the absence of noise and system considerations, orthogonal control functions with equal energy will lead to the same system performance; but for real systems other factors may influence the choice of functions. These factors include the need to synchronize the active illuminants and acquisition device, hardware capabilities, scene characteristics, and system noise and quantization characteristics.

For example, some control functions require the modulator and demodulator to be synchronized. If the illuminant controller and camera are not closely coupled, this synchronization may be difficult to achieve. There are some control functions, however, that do not require accurate synchronization. Sinusoidal functions, for example, can be demodulated without precise synchronization between the illuminant and the camera. This is not the case for other control functions, such as square wave functions (Haar basis). For these control functions the emitter and camera should be synchronized to demodulate the signal properly by avoiding acquisitions during the on/off transition.

Scene characteristics are another important factor to consider when selecting control functions. Scene ambient illuminants are usually either steady or flickering at 60 or 120 Hz, so it is important not to pick a bases set that contain energy in the same frequency range as the scene illuminants. By avoiding the frequency range of typical scene illuminants, the demodulator can achieve higher performance because there is less noise in the frequency band that it is detecting and demodulating.

Lastly, the noise properties of the entire system should be considered when selecting modulating control functions. Suppose that the main system noise is due to quantization, and all other sources of noise are negligible. In this case sinusoidal bases will outperform square wave bases, and this will be especially true if many samples can be obtained. The reason for this advantage is that multiple samples from the square wave basis set do not reduce the system quantization error. However when a sinusoidal bases set is used, the detected signal occupies many different levels and the quantization noise is more evenly distributed around zero. This error is added together and averaged away during the demodulation process; the negative error values more evenly match the positive error values and these opposing errors tend to cancel. If other noise sources are the principal limit, then other control functions may be preferred.

4. APPLICATIONS

The ability to measure an image with a controlled active illuminant, even in the presence of a passive illuminant, can be used in a variety of novel image acquisition algorithms. We close this paper by describing briefly two applications.

4.1 Color imaging with a monochrome sensor

From the earliest days of electronic color imaging, engineers have produced and acquired images using a field sequential method. The first television displays used a rotating color wheel to produce a sequence of red, green and blue images. The visual pathways integrate these images to form the impression of a single color image. Similarly, it is possible to acquire a color image using a monochrome sensor by sequentially placing color filters in the light path. Again, the three image sequences can be combined to produce a single color image. Field sequential color has some advantages and many disadvantages compared to capture using an image sensor with a mosaic [10]. One of the most significant disadvantages is that the method is only appropriate for a static scene; object motion between the capture times produces very strong color artifacts.

Using active illuminants, it is possible to modify the field sequential approach to create color images from a monochrome sensor. Suppose an active illuminant comprises three light sources, say red, green and blue sources. If the source control functions are orthogonal, then it is possible to demodulate the acquired time series into three images corresponding to the scene illuminated by the red, green, and blue sources. These three images, acquired simultaneously, can be combined to form a single color image.

Figure 6a shows the time series of an image pixel measuring a white surface while the three illuminants were modulated. The red, green and blue illuminants were modulated at different frequencies and their signatures can be seen in the amplitude spectrum of the time series shown in the panel on the right of Figure 6a. This spectrum has energy at three peaks frequencies

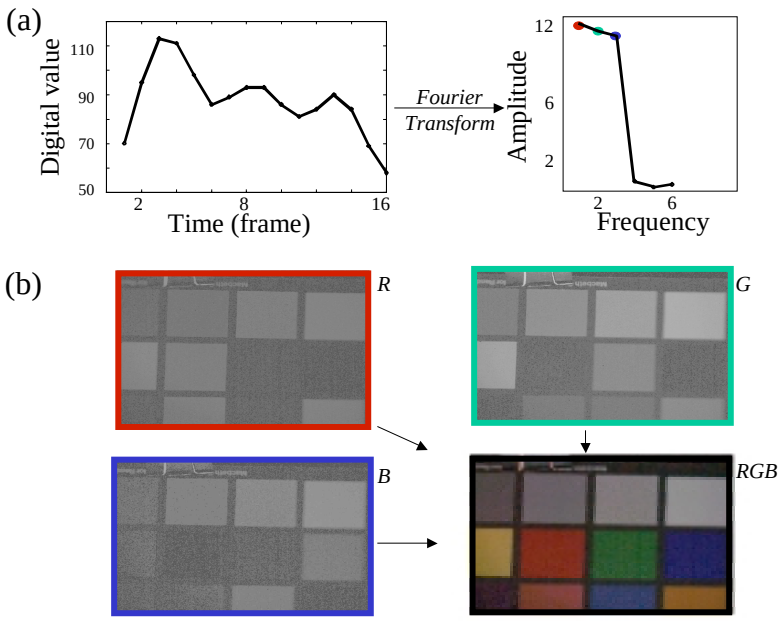

Figure 6. A color image measured with a monochrome image sensor and three active illuminants. Red, green and blue active illuminants were modulated at three different frequencies. (a) The time series at a single pixel measured from a white surface is shown on the left, and the amplitude spectrum of the time series is shown on the right. Three frequencies, corresponding to the modulations of the three active illuminants, contain significant energy. (b) The images obtained by demodulating at the R, G and B illuminant frequencies along with the combined color image from the three modulations (RGB) are shown.

corresponding to the modulations of the red, green and blue light sources. Figure 6b shows the three separate images reconstructed by demodulating the image series at every pixel as well as the final color image assembled from the three color images. It is possible to use this *field simultaneous* method with many different sets of modulation control functions that are suited to specific applications.

4.2 Estimating the spatial distribution of the passive illuminant

In many applications, it is useful to distinguish between image edges caused by surface markings and those caused by illumination discontinuities, such as shadows. In certain circumstances, active illuminants can be used to separate these two kinds of image edges. As an example, suppose that the sensor responds to a narrow wavelength band near λ_0. Further suppose that the active illuminant is spatially constant across the image region. Then, we can express the spatial response of the camera image to the active illuminant, as

$$r_A(x) = \int A(\lambda, x)R(\lambda)S(\lambda, x)d\lambda \approx A(\lambda_0)R(\lambda_0)S(\lambda_0, x) \quad (7)$$

Similarly, the spatial sensor response to the passive illuminant is

$$r_P(x) = \int P(\lambda, x)R(\lambda)S(\lambda, x)d\lambda \approx P(\lambda_0, x)R(\lambda_0)S(\lambda_0, x) \quad (8)$$

The ratio of these two images across space is proportional to the spatial variation of the passive illuminant,

$$\frac{P(\lambda_0, x)}{A(\lambda_0)} \approx \frac{r_P(x)}{r_A(x)} \quad (9)$$

Hence, using the measurements of the active and passive illuminants we can estimate the spatial variation of the passive illuminant. Image edges that are caused by spatial variations in the illumination can be distinguished from those due to a surface marking.

The process is illustrated by the measurements in Figure 7. Panel (a) shows an image of a patterned surface that is illuminated by a passive source. The illumination is uneven, causing a strong shadow edge. The dashed line drawn on top of the figure shows the intensity (digital value) along one of the scan lines. Panel (b) shows an estimate of the spatial structure of the passive illumination. Dividing the active and passive illuminant images as described in Equation 9 derives this estimate. The dotted line superimposed on the figure shows the estimated passive illuminant intensity. Panel (c) shows the measured spatial pattern of the illuminant obtained using a uniform gray card instead of the patterned surface. Panel (d) compares the scan line in the original image with the estimated and measured passive illuminant intensity. In reflective regions of the image, the agreement is good. Systematic errors can be noted in regions when the surface is very dark.

5. DISCUSSION AND CONCLUSIONS

We have implemented and evaluated an active imaging method for color imaging. We found that it is possible to measure images from multiple modulating light sources, and that these images correspond to measuring the scene under one of the active illuminants. The computation of such images is well suited to the MCSI architecture in which a collection of images is captured and processed within a sensor, but only a reduced number of images are produced.

Using our experimental apparatus, we have measured how accurately the modulating and ambient illumination can be separated as a function of experimental variables, including LED intensity, temporal sampling rate and modulation amplitude. Finally, we have demonstrated how this technique can be used to create a color image from a monochrome sensor, and how the technique can be applied to distinguishing image edges caused by surface markings from those caused by illumination edges.

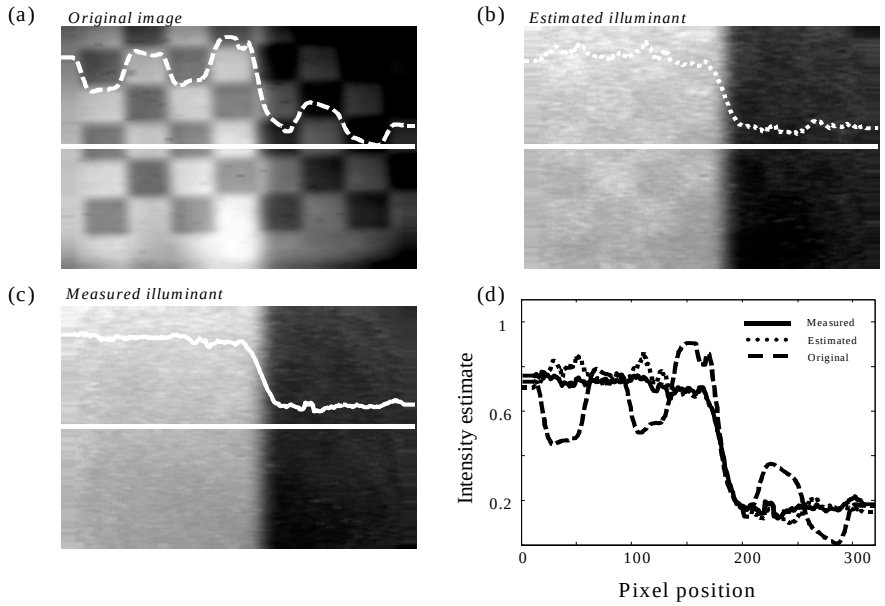

Figure 7: An active imaging method can be used to discriminate surface markings and shadows. See text for details.

ACKNOWLEDGMENTS

This work was supported by the Programmable Digital Camera project at Stanford University (http://smartcamera.stanford.edu/pdc.html). Feng Xiao acknowledges the support from Stanford Graduate Fellowship program. We also thank K. N. Salama for helping with the control of LED arrays.

REFERENCES

1. Ballard, D., *Introduction to Natural Computation*. 1997: MIT Press.
2. Goodale, M.A. and G.K. Humphrey, *The objects of action and perception*. Cognition, 1998. **67**(1-2): p. 181-207.
3. Holmes, R.B., *Applications of lasers to imaging of distant objects.* SPIE Proceedings, 1993. **1871**: p. 306-315.
4. Morimoto, C.H. and M. Flickner. *Real-Time Multiple Face Detection Using Active Illuminant.* in *Fourth IEEE International Conference on Automatic Face and Gesture Recognition.* 2000.
5. McDonald, J.P., R.J. Fryer, and J.P. Siebert, *Stereo scene coding using SLM active illuminant*, in *Proc. 26th Int. Symp. on Automotive Technology and Automation Mechatronics*. 1993: Aachen. p. 169-176.
6. Kang, S., et al. *A Multibaseline Stereo System with Active Illuminant and Real-time Image Acquisition.* in *The Fifth International Conference on Computer Vision*. 1995.
7. Ando, S., K. Nakamura, and T. Sakaguchi. *Ultrafast Correlation Image Sensor: Concept, Design, and Applications,*. in *Proc. IEEE CCD/AIS Workshop*. 1997. Bruges, Belgium: IEEE.
8. Ando, S. and A. Kimachi. *Time-Domain Correlation Image Sensor: First CMOS Realization of Demodulator Pixels Array*. in *Proc. '99 IEEE CCD/AIS Workshop*. 1999. Karuizawa.
9. Yang, D.X.D., et al. *A 640*512 CMOS image sensor with ultra wide dynamic range floating-point pixel-level ADC*. in *1999 IEEE International Solid-State Circuits Conference*. 1999. San Francisco, CA, USA.
10. Catrysse, P., A. El Gamal, and B.A. Wandell. *Comparative analysis of color architectures for image sensors*. in *Proc. of the SPIE*. 1999. San Jose.