Bayesian Method for Reconstructing Color Images from Trichromatic Samples

David H. Brainard
Department of Psychology, University of California, Santa Barbara, California

Introduction

We typically specify color images by a triplet of red, green, and blue (RGB) values at each pixel on a rectangular sampling grid. This representation is not, however, available directly from the output of many color CCD cameras. Rather, these cameras provide only single photosensor response at each pixel. Color information comes because the camera as a whole contains multiple classes of photosensors, with each class characterized by a distinct spectral sensitivity. Usually there are three classes, referred to as red (R), green (G), and blue (B). Fig. 1 provides a schematic illustration of this design.

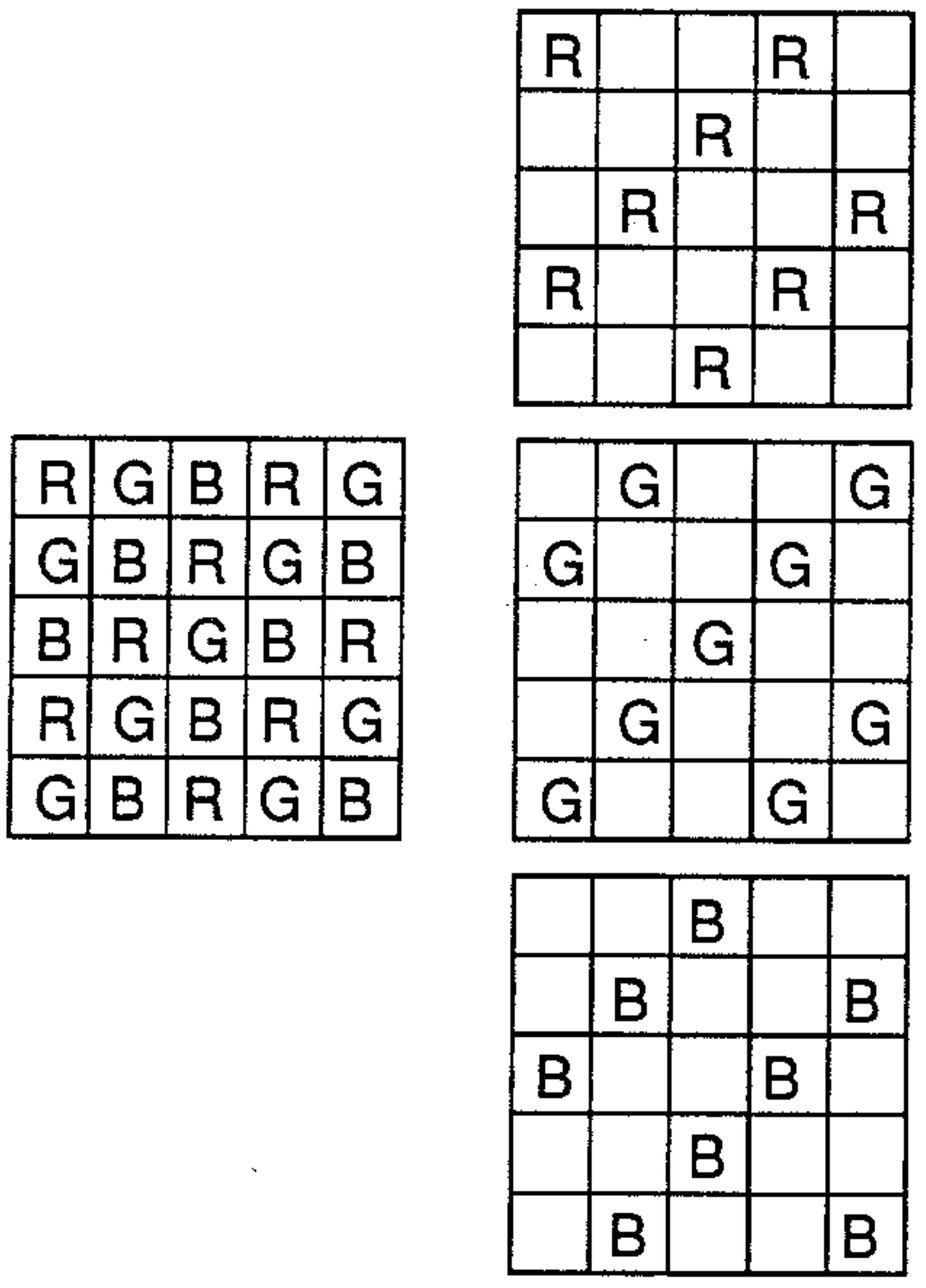

Figure 1. Color CCD camera. Left: The CCD camera contains a rectangular array of sensors. Each sensor has either an R, G, or B spectral sensitivity. Right: We may think of the overall sensor array as three interleaved subarrays. Each subarray corresponds to one of the three sensor classes. Note that although the overall geometry of the array is rectangular, this is not the case for the individual subarrays.

To obtain a full color image, we must reconstruct at each pixel the responses of the two missing sensor classes. I call this polychromatic reconstruction. Polychromatic reconstruction generalizes the classic problem of reconstructing signals from samples.[1]

The fundamental result for regular one-dimensional, monochromatic sampling is that a signal may be reconstructed from samples if it contains no frequency components greater than the Nyquist limit of the sampling array.[1] The Nyquist limit is defined as half the sampling rate. Signal components at frequencies above the Nyquist limit cannot be distinguished from lower frequency components, as illustrated in Fig. 2. When standard methods are used, high frequency components are reconstructed at lower frequencies. This phenomenon is referred to as aliasing.

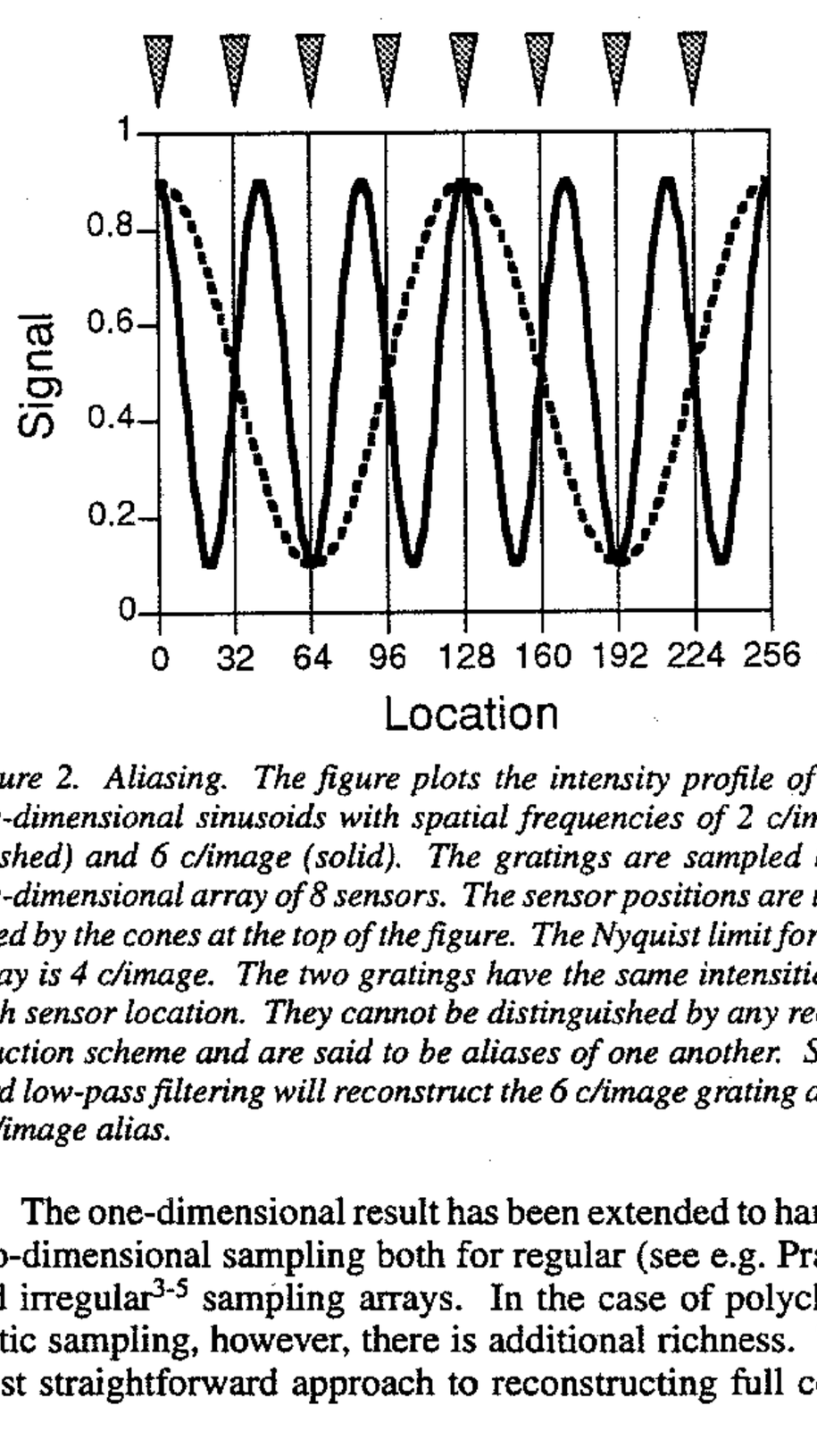

Figure 2. Aliasing. The figure plots the intensity profile of two one-dimensional sinusoids with spatial frequencies of 2 c/image (dashed) and 6 c/image (solid). The gratings are sampled by a one-dimensional array of 8 sensors. The sensor positions are indicated by the cones at the top of the figure. The Nyquist limit for this array is 4 c/image. The two gratings have the same intensities at each sensor location. They cannot be distinguished by any reconstruction scheme and are said to be aliases of one another. Standard low-pass filtering will reconstruct the 6 c/image grating as its 2 c/image alias.

The one-dimensional result has been extended to handle two-dimensional sampling both for regular (see e.g. Pratt[2]) and irregular[3-5] sampling arrays. In the case of polychromatic sampling, however, there is additional richness. The most straightforward approach to reconstructing full color

images from color CCD sensor responses is to treat each color band separately. Standard methods may then be applied to each subarray in turn. The disadvantage of this approach is that aliasing will occur when the image contains spatial frequencies above the Nyquist limits of the individual subarrays, as illustrated in Fig. 3. This sort of chromatic aliasing can produce objectionable artifacts in images recorded with color CCD cameras, particularly near intensity edges (see Fig. 4).

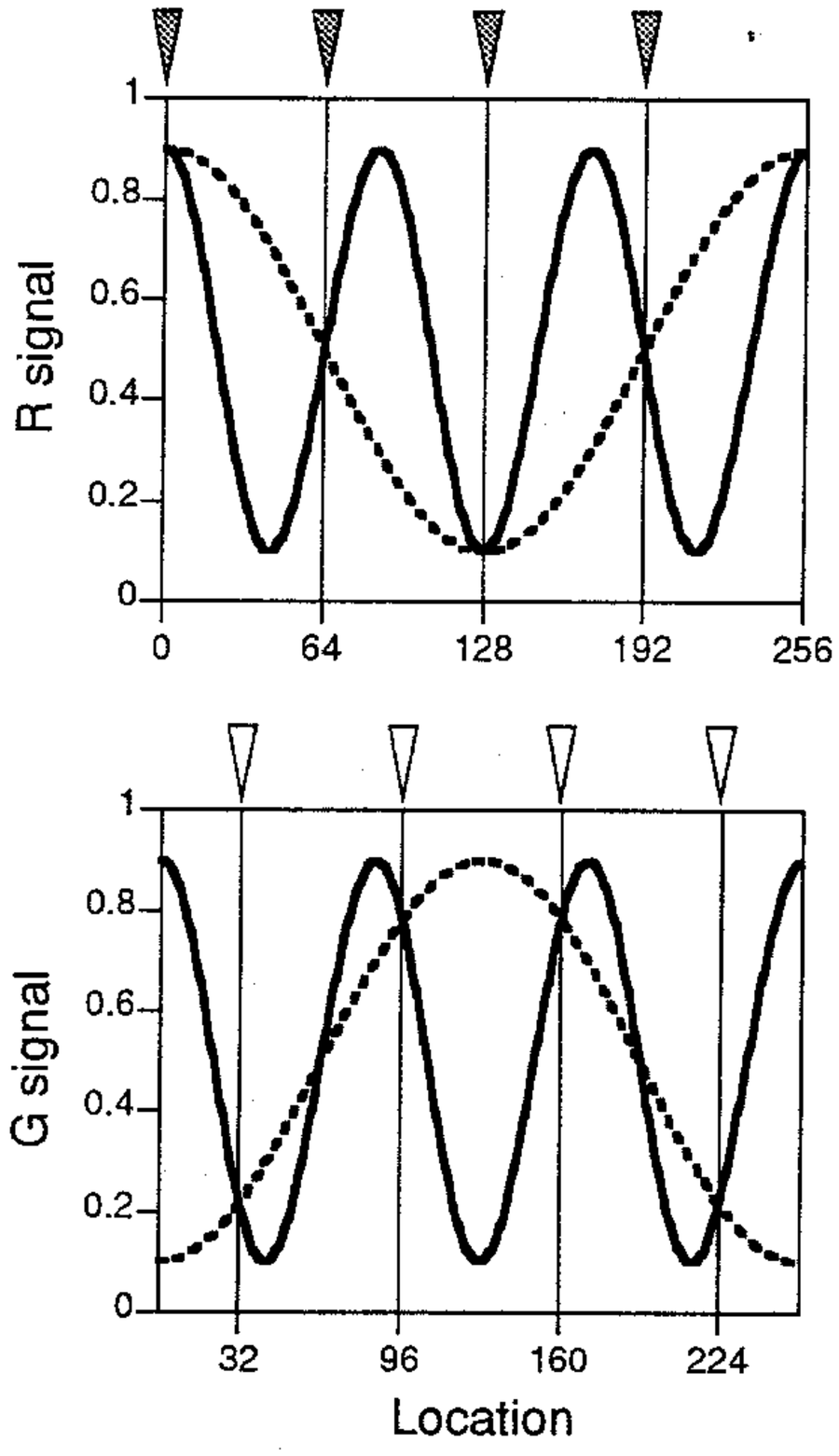

Figure 3. Chromatic aliasing. The upper and lower panels plot the R and G components of two sinusoids. One sinusoid is an intensity grating with spatial frequency 3 c/image (solid). The R and G components for this sinusoid are in pahse. The second sinusoid is a red/green grating with spatial frequency 1 c/image (dashed). The R and G components for this siusoid are out of phase. The sinusoids are sampled by two interleaved subarrays. The R sensor positions are indicated by the shaded cones above the top panel. The G sensor positions are indicated by the open cones above the bottom panel. The two sinusoids produce the same responses in all of the sensors and cannot be distinguished by any reconstruction scheme. Low-pass filtering applied separately to each subarray will reconstruct the 3 c/image intensity grating as its 1 c/image red/green alias. Although the total number of sensors is the same as it was in Fig. 2, distortion from aliasing occurs at lower spatial frequencies.

Brute force may be used to reduce chromatic aliasing. One possibility is to blur the image reaching the camera sensors so that spatial frequencies above the subarray Nyquist limits are heavily attenuated. This has the obvious disadvantage that the resulting image will be blurred. An alternative is to increase the overall sampling density of the camera sensors so that the Nyquist limit for the subarrays is raised. This has the disadvantage that it raises the cost of the camera.

A second approach is to use reconstruction methods that are designed specifically for polychromatic sampling. Such methods cannot restore the information lost by sampling. In many cases, however, they can be designed to reconstruct intensity signals accurately at spatial frequencies above Nyquist limits of the individual subarrays. This will be possible if there is a statistical correlation between the responses of sensors in different color classes, so that, say, the response of an R sensor at a location contains information about how a G sensor would have responded at that location.

My interest in polychromatic sampling began with the observation that the human visual system is effective at reconstructing images from the responses of interleaved subarrays of retinal cones.[6,7] In trying to understand how this is possible, I cast the polychromatic sampling problem in the form of a statistical decision problem: given the responses of interleaved subarrays of sensors and a model of the statistical distribution of images, what image estimate minimizes the expected reconstruction error? This line of thinking led to a method that may be applied to color CCD camera data. In this paper, I present my analysis and show some examples of its performance.

Polychromatic reconstruction techniques have also been developed in the engineering community.[8,9] These share with my work the feature that responses from the entire sensor array are used jointly in the reconstruction process.

Formal Development

Image representation

I represent an image by the discrete function $i(x_i, y_j, \lambda_k)$. The coordinates x_i $(1 \le i \le N_{rows})$ and y_j $(1 \le j \le N_{cols})$ represent $N_{rows}\,N_{cols}$ evenly spaced sample locations on a rectangular grid. This grid should be much denser than the sampling array being studied. The λ_k $(1 \le k \le N_{wls})$ represent the number of color bands in the image. The λ_k may be evenly spaced sample wavelengths throughout the visible spectrum, in which case the function $i(x_i, y_j, \lambda_k)$ represents the spectrum at each image location. To simplify the discussion here, we regard λ_k as an indicator variable for R, G, and B. The use of discrete representations for both spatial position and wavelength has been discussed extensively elsewhere.[2,10]

I use the $N_{rows}\,N_{cols}\,N_{wls}$ dimensional column vector $\mathbf{i}$ to represent the function $i(x_i, y_j, \lambda_k)$. The n^{th} entry of $\mathbf{i}$ is $i(x_i, y_j, \lambda_k)$ with

$$\begin{aligned} i &= 1 + ((n-1) \text{ modulo } N_{rows}) \\ j &= 1 + (\lfloor (n-1)/N_{rows} \rfloor \text{ modulo } N_{cols}) \\ k &= 1 + \lfloor (n-1)/(N_{rows}N_{cols}) \rfloor . \end{aligned} \tag{0}$$

This choice of indexing enumerates all of the values of the discrete function $i(x_i, y_j, \lambda_k)$.

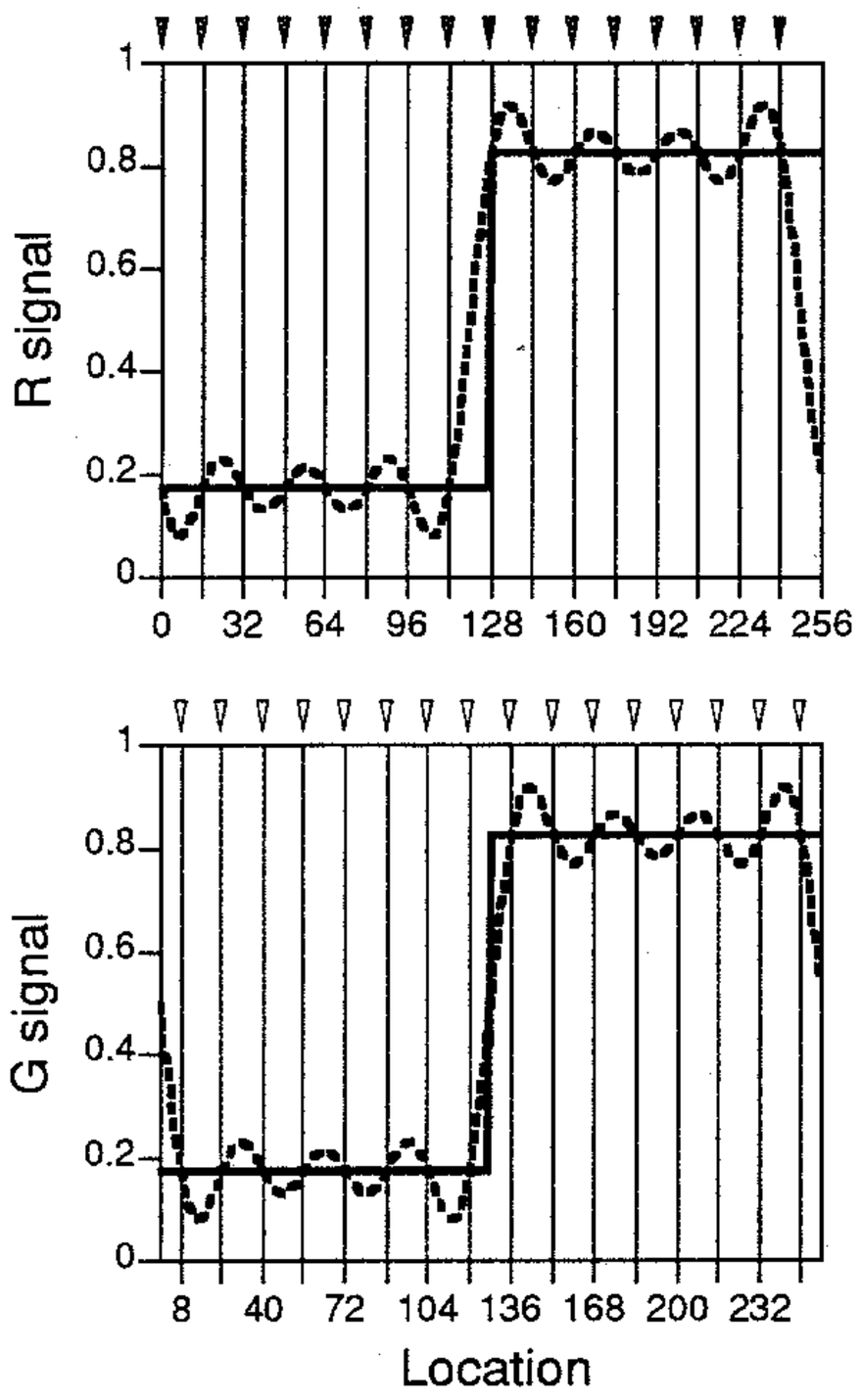

Figure 4. Chromatic aliasing at edges. The upper and lower panels plot the R and G components of two signals. One signal is an intensity step edge (solid). The R and G components of this edge are in phase with one another. The second signal is a low frequency alias of the intensity step edge (dashed). The sinusoids are sampled by two interleaved subarrays. The R sensor positions are indicated by the shaded cones above the top panel. The G sensor positions are indicated by the open cones above the bottom panel. The two signals produce the same responses in all of the sensors. Careful examination of the low frequency alias reveals that its R and G components are not in phase at locations adjacent to the edge. This alias will contain visually salient colored fringes.

Sampling

Consider the relation between the response of a single sensor and the image data. Each sensor is completely characterized by a polychromatic receptive field, which specifies how strongly it responds to light from each color band at each location in the image. The receptive field of a single sensor may be described by a function $s(x_i, y_j, \lambda_k)$. To response r of the sensor is given by

$$r = \sum_{i,j,k} s(x_i, y_j, \lambda_k)\, i(x_i, y_j, \lambda_k) + e \qquad (1)$$

where e is a random variable representing sensor noise. We can rewrite Eq. 1 as

$$\mathbf{r} = \mathbf{s}\,\mathbf{i} + \mathbf{e} \qquad (2)$$

where the $N_{rows}\,N_{cols}\,N_{wls}$ dimensional row vector **s** is the vector representation of $s(x_i, y_j, \lambda_k)$.

Suppose there are N_{sensor} sensors in the overall camera array, enumerated by the index m. The responses of all the sensors can be represented by a single N_{sensor} dimensional column vector **r** whose m^{th} entry is the response of the m^{th} sensor. Let **S** be the N_{sensor} by $N_{rows}\,N_{cols}\,N_{wls}$ matrix whose m^{th} row is the row vector that represents the receptive field of the m^{th} sensor. Then

$$\mathbf{r} = \mathbf{S}\,\mathbf{i} + \mathbf{e} \qquad (3)$$

where **e** is an N_{sensor} dimensional column vector representing sensor noise. In this paper, I assume that **e** is multivariate Normal with mean $\mathbf{u}_e$ and covariance $\mathbf{K}_e$.

Eq. 3 describes the sampling process. The goal of a reconstruction algorithm is to provide an inverse mapping, $\hat{\mathbf{i}} = f(\mathbf{r})$, where $\hat{\mathbf{i}}$ is an estimate of the original image **i**. In this sense, the polychromatic estimation problem is an inverse problem. It is a linear inverse problem because, in the absence of noise, **r** is a linear function of **i**. It is an underdetermined inverse problem because the number of samples is typically much smaller than the number of image parameters to be estimated. In this paper, I adopt a Bayesian approach to the finding an estimator $f(\mathbf{r})$. Other ways to approach underdetermined linear inverse problems have been extensively discussed (see e.g. Pratt[2] or Menke[11]).

Bayesian Approach

The Bayesian approach to inverse problems is very simple (see e.g. Berger[12] or Brainard and Freeman[13]). The problem is defined in terms of two probability distributions and a loss function.

The first probability distribution is the prior $p(\mathbf{i})$. The prior specifies the likelihood of encountering each possible image. It captures what is known about images before the sensor data are examined. Prior information is used in most reconstruction methods. For example, low-pass filtering (i.e. sinc interpolation) is based on the prior assumption that the image contains no frequency components above the Nyquist limit of the array. What differs in the Bayesian framework is that the prior information is expressed explicitly as a probability distribution. I discuss the construction of sensible prior distributions below.

The second probability distribution $p(\mathbf{r}\,|\,\mathbf{i})$ is the likelihood of the sensor data conditional on the image being viewed. The likelihood is really a specification of the imaging model. Here $p(\mathbf{r}\,|\,\mathbf{i})$ is determined by Eq. 3 and we write

$$p(\mathbf{r}\,|\,\mathbf{i}) = N(\mathbf{S}\,\mathbf{i}, \mathbf{K}_e)\ . \qquad (4)$$

Given the prior and the likelihood, we may compute the posterior distribution $p(\mathbf{i}\,|\,\mathbf{r})$. The posterior specifies how likely any given image is, conditional on the observed sensor responses. The posterior is computed using Bayes' rule:

$$p(\mathbf{i}\,|\,\mathbf{r}) = p(\mathbf{r}\,|\,\mathbf{i})\ p(\mathbf{i})\ . \qquad (5)$$

The loss function $L(\hat{\mathbf{i}}, \mathbf{i})$ specifies the cost of estimating $\hat{\mathbf{i}}$ when the underlying image is **i**. Given a loss function and the posterior, we may compute the Bayes risk

$$R(\hat{\mathbf{i}}\,|\,\mathbf{r}) = E_{p(\mathbf{i}\,|\,\mathbf{r})}\{L(\hat{\mathbf{i}}, \mathbf{i})\} \qquad (6)$$

where $E_{p(i|r)}$ denotes the expected value with respect to the posterior. The Bayes risk specifies the expected loss for each choice of $\hat{\mathbf{i}}$, conditional on $\mathbf{r}$. We choose the reconstructed image as the particular $\hat{\mathbf{i}}$ that minimizes the Bayes risk. In this paper, I only consider the squared error loss function $L(\hat{\mathbf{i}}, \mathbf{i}) = |\mathbf{i} - \hat{\mathbf{i}}|^2$. Brainard and Freeman provide some discussion on loss function choice.[13]

Specifying the Prior

To find an explicit Bayesian estimator, it is necessary to specify the prior distribution $p(\mathbf{i})$. I use linear models to construct prior distributions. I assume that each image can be expressed as a weighted sum of N_{basis} basis images. Let the basis images be represented specified by columns of a $N_{rows}N_{cols}N_{wls}$ by N_{basis} basis matrix $\mathbf{B}$. I assume that every image $\mathbf{i}$ may be expressed as a weighted sum of these columns, so that

$$\mathbf{i} = \mathbf{B}\mathbf{w} . \tag{7}$$

Images constrained by Eq. 7 are said to lie within the linear model defined by $\mathbf{B}$. The N_{basis} dimensional vector $\mathbf{w}$ contains the linear model weights.

The linear model constraint by itself does not express a probability distribution. To turn Eq. 7 into a statement about probability, I assume that $\mathbf{w}$ is multivariate Normal with mean $\mathbf{u}_w$ and covariance $\mathbf{K}_w$. This implies that

$$p(\mathbf{i}) = N(\mathbf{B}\mathbf{u}_w, \mathbf{B}\mathbf{K}_w\mathbf{B}^T) \tag{8}$$

The general form of Eq. 8 captures two features of natural images. First, the average power spectrum of natural images falls off fairly rapidly as a function of spatial frequency.[14,15] Second, the signals in different color bands are positively correlated.[15] This latter observation is particular true of the long and middle wavelength sensitive cones that mediate human vision, because the spectral sensitivities of these cones are very similar.[16]

By choosing the basis images to be spatial sinusoids in each color band, it is relatively straightforward to construct $\mathbf{K}_w$ given a) the average image power spectrum in each color band and b) the correlation between signals in the different color bands. The simplest case, and the only one I have implemented, is where the spatial statistics are the same in each color band and the color correlations are independent of spatial location.

The general form of Eq. 8 does not capture all of the structure in natural images. For example, Normal priors cannot express the fact that image intensities must be positive. None-the-less, these priors are more flexible than the assumptions of band-limited signals that have been used in previous reconstruction algorithms. It seems worthwhile to investigate the performance possible when such priors are assumed. Future specification of better prior distributions may be incorporated within the same framework presented here.

Reconstruction

When the prior and the likelihood are both multivariate normal, so is the posterior. Moreover, the mean and covariance of the posterior may be computed in closed form. (See Lee[17] for the derivation in the univariate case. The multivariate generalization is straightforward.) The estimator that minimizes the Bayes risk for quadratic loss is exactly the posterior mean, and using this fact we may derive that

$$\hat{\mathbf{i}} = \mathbf{I}\,\mathbf{r} + \mathbf{i}_0 \tag{9}$$

with

$$\mathbf{I} = \mathbf{B}\mathbf{K}_w(\mathbf{R}\mathbf{B})^T\left((\mathbf{R}\mathbf{B})\mathbf{K}_w(\mathbf{R}\mathbf{B})^T + \mathbf{K}_n\right)^{-1}$$
$$\mathbf{i}_0 = \mathbf{B}\,\mathbf{w}_0 - \mathbf{I}\,(\mathbf{R}\mathbf{B})\,\mathbf{u}_w + \mathbf{I}\,\mathbf{u}_e . \tag{10}$$

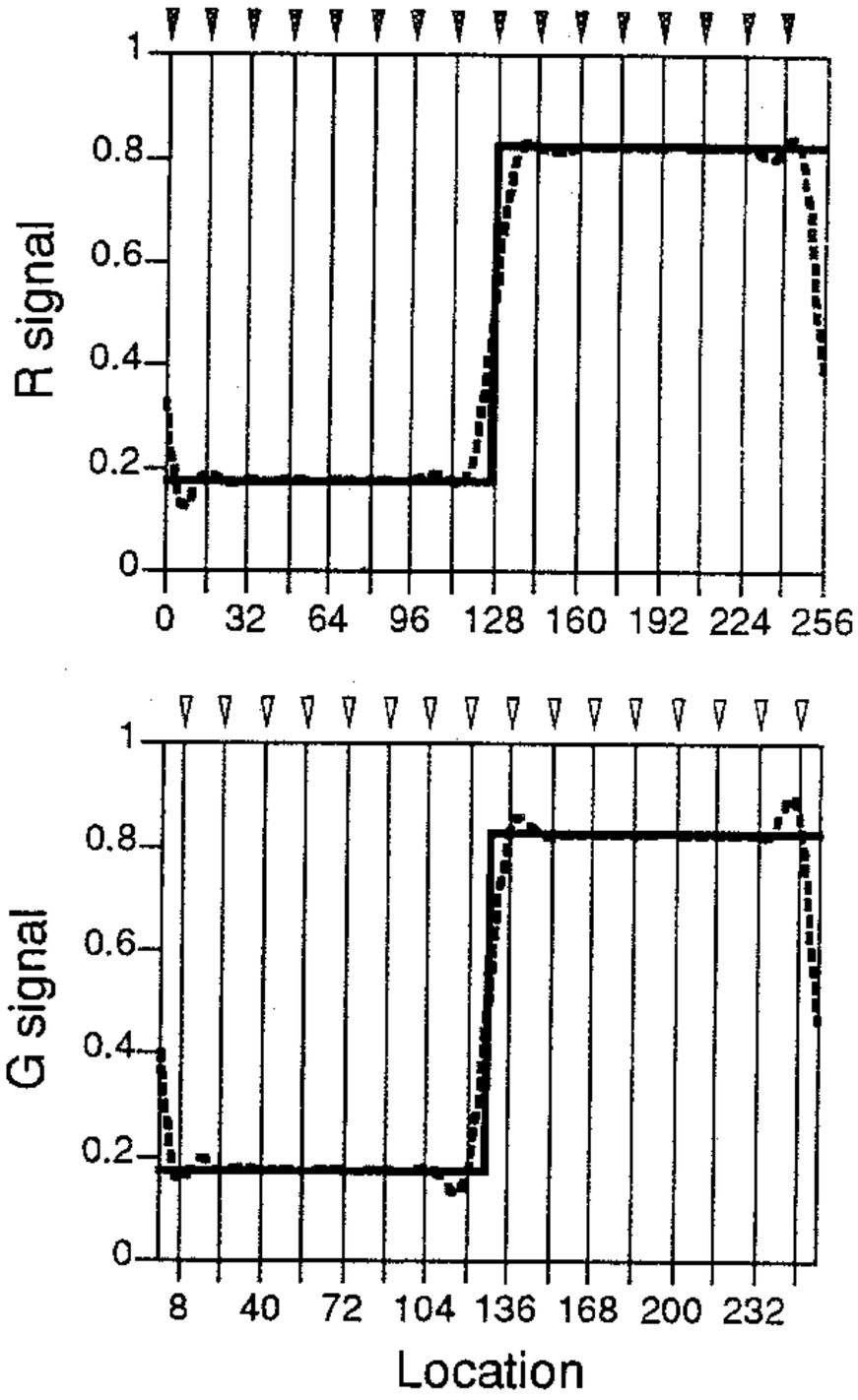

Figure 5. Polychromatic reconstruction of an intensity step edge. The upper and lower panels plot the R and G components of two signals. One signal is an estimate of the step edge, reconstructed using the Bayesian method described in this paper (dashed). The sensor array is the same as in Fig. 4. The reconstruction was made assuming a) that the camera's optics blur the incident image with a blur standard deviation equal to one-half the overall spacing between sensors, b) that Rand G signals are correlated with a correlation coefficient of 0.8, c) that the average amplitude spectrum of the image population is inversely proportional to spatial frequency, and d) that there is additive sensor noise with a standard deviation of one percent of the mean sensor response. The optical blurring was taken into account when simulating the sensor responses, but no noise was added. The reconstruction shows much less chromatic fringing than the low frequency edge alias in Fig. 4.

Eq. 9 describes a linear estimator. The estimator is determined by parameters describing the camera optics and sensor array ($\mathbf{R}$, $\mathbf{u}_e$, $\mathbf{K}_e$) and parameters describing the image population ($\mathbf{B}$, $\mathbf{u}_w$, $\mathbf{K}_w$). An advantage of the Bayesian approach is that it lets us design the reconstruction algorithm in terms of parameters that may (in principle) be measured or controlled.

Example

I have implemented the Bayesian polychromatic reconstruction method both for one and two-dimensional signals. Here I show results for one-dimensional signals, which are easier to plot. Fig. 5 shows a reconstruction of an intensity step edge sampled by the same camera array as in Fig. 4. The reconstruction shows much less chromatic fringing than the edge alias shown in Fig. 4. There are three main reasons for this. First, I have incorporated a small amount of optical blur into the imaging model. This has the general tendency to reduce the type of ringing seen in Fig. 4. Second, the reconstruction algorithm is based on the assumption that the R and G signals are correlated. This favors intensity interpretations of the sensor data. Third, the reconstruction algorithm is based on the assumption that the average image amplitude spectrum is inversely proportional to spatial frequency. This assumption is reasonable for step edges.

The good performance illustrated in Fig. 5 is achieved at a cost. Reconstructions of less likely signals will contain distortion. Such distortion is inevitable for some input signals, since blurring, sampling, and noise lead to information loss. What the Bayesian approach makes possible is improved performance for likely images.

One simple way to examine reconstruction tradeoffs is illustrated by Fig. 6. The two panels plot a measure of how completely intensity and red/green gratings are reconstructed for various choices of R and G sensor correlation. As the assumed correlation increases, the intensity gratings are reconstructed more completely while the red/green gratings are reconstructed less completely. Apparently, the reconstruction algorithm may be adjusted to control this tradeoff.

Fig. 6 does not, of course, provide a complete evaluation of the algorithm's performance. In addition to knowing how completely the input signal is reconstructed, it is also important to know the character of the distortion produced by different reconstruction methods for different input signals. This issue will be discussed in a future paper.

Discussion

Space limitations prevent extensive discussion. I close by noting that the method I have developed is very general. It provides a recipe for designing optimal reconstruction algorithms that handle irregular polychromatic sampling in the presence of optical blur and sensor noise. Because the method tailors reconstruction algorithms to the properties of the camera and the image population, it can be used to compare how well different camera designs will perform. For example, the number of sensor classes, their spectral sensitivities, the relative numbers and placement of sensors in each class, and the amount of optical blur are all design parameters of CCD cameras. In Fig. 6, the correlation between R and G sensors was treated as a reconstruction parameter for a fixed camera. The same principle, however, could be used to compare the performance of different cameras, when signals from each are reconstructed optimally.

Acknowledgments

I thank A. Ahumada, B.Wandell and D. Williams for helpful discussions. Initial results were presented at the 1989 and 1993 ARVO annual meetings.[18,19] Supported by NEI EY 10016.

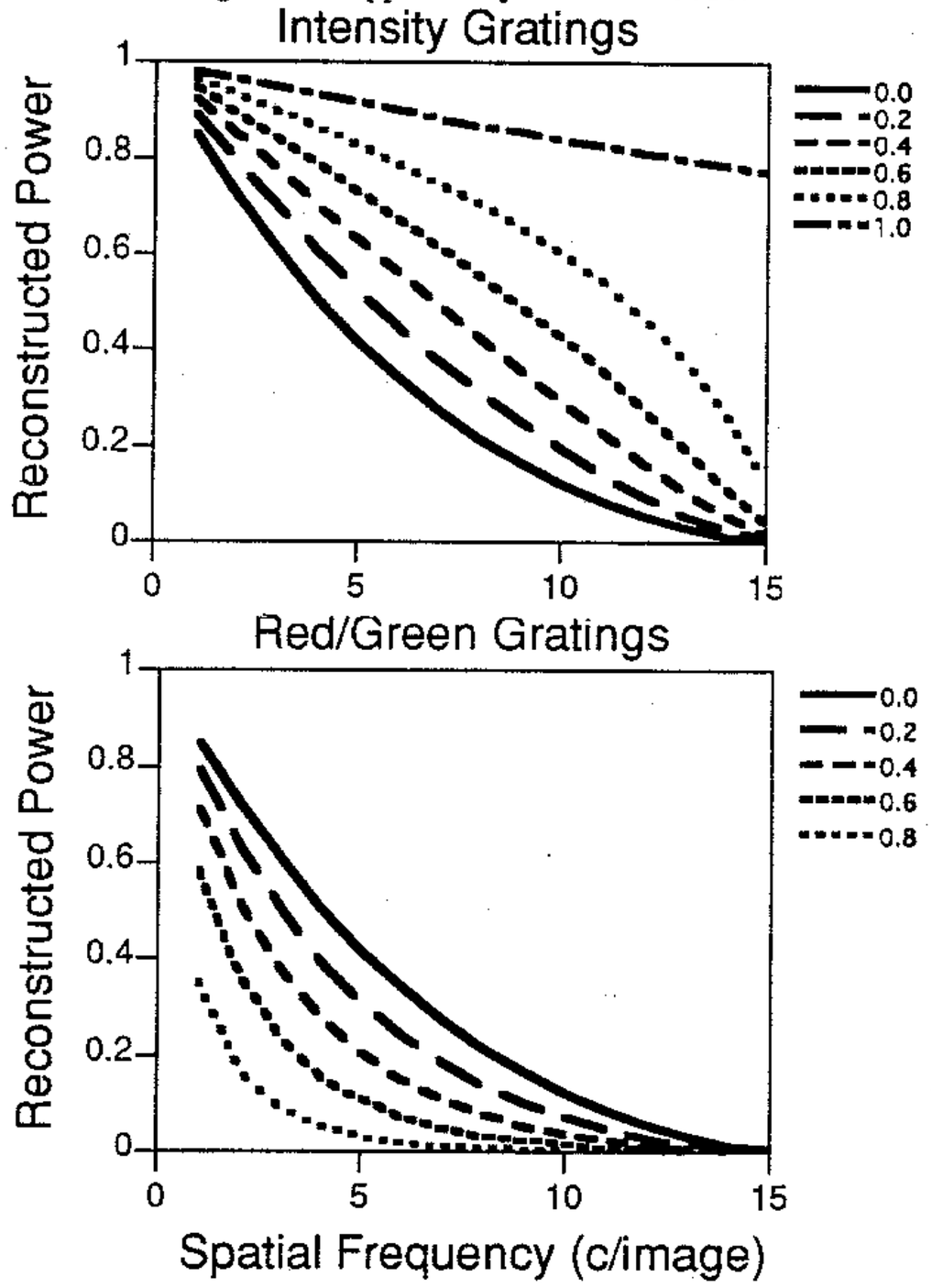

Figure 6. Reconstruction tradeoffs. To generate the figure, I reconstructed sinusoidal gratings of unit power at a series of spatial frequencies. The sinusoids were either intensity gratings (top panel) or red/green gratings (bottom panel). Each line in the plots show the power of the reconstructed signal in the same color direction and at the same spatial frequency as the input. The different lines in each plot are parameterized by the assumed correlation between R and G sensors. I assumed that the camera's optics blur the incident image with a blur standard deviation equal to one-half the overall spacing between sensors and that there is additive sensor noise with a standard deviation of one percent of the mean sensor response. The optical blurring was taken into account when simulating the sensor responses, but no noise was added.

References

1. Bracewell, R., *The Fourier transformation and Its Applications*, (McGraw-Hill: New York, 1978).
2. Pratt, W. K., *Digital Image Processing*, (John Wiley & Sons: New York, 1978).
3. Yen, J. L., "On nonuniform sampling of bandwidth-limited signals," *IRE Transactions on circuit theory* **CT-3**, 251-257 (1956).

4. Chen, D. S. and Allebach, J. P., "Analysis of error in reconstruction of two-dimensional signals from irregularly spaced samples," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-35**, 173-180 (1987).
5. Yellott, J. I., Jr., "The photoreceptor mosaic as an image sampling device." In *Advances in Photoreception*; (National Academy Press: Washington, D.C., 1990).
6. Williams, D. R., Sekiguchi, N., Haake, W., Brainard, D. H. and Packer, O., "The cost of trichromacy for spatial vision." In *From Pigments to Perception*; B. B. Lee and A. Valberg eds., (Plenum Press: New York, 1991).
7. Brainard, D. H. and Williams, D. R., "Spatial reconstruction of signals from short wavelength cones," *Vision Research* **33**, 105-116 (1993).
8. Freeman, W., "Method and apparatus for reconstructing missing color samples," *United States Patent 4663655* (1987).
9. Cok, D., "Signal processing method and apparatus for producing interpolated chrominance values in a sampled color image signal," *United States Patent 4642678* (1987).
10. Brainard, D. H., "Colorimetry." In *OSA Handbook of Optics*; *2cd edition*, M. Bass ed., (Optical Society of America: Washington, D.C., in press).
11. Menke, W. *Geophysical Data Analysis: Discrete Inverse Theory*, (Academic Press: San Diego, 1989).
12. Berger, T. O. *Statistical Decision Theory and Bayesian Analysis*, (Springer-Verlag: New York, 1985).
13. Brainard, D. H. and Freeman, W. T., "Bayesian method for recovering surface and illuminant properties from photosensor responses," *IS&T/SPIE Symposium on Electronic Imaging Science & Technology*, (in press).
14. Field, D. J., "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A* **4**, 2379-2394 (1987).
15. Burton, G. J. and Moorhead, I. R., "Color and spatial structure in natural scenes," *Applied Optics* **26**, 157-170 (1987).
16. Schnapf, J. L., Kraft, T. W. and Baylor, D. A., "Spectral sensitivity of human cone photoreceptors," *Nature* **325**, 439-441 (1987).
17. Lee, P. H. *Bayesian Statistics: an Introduction*, (Oxford University Press: New York, 1989).
18. Brainard, D. H., Wandell, B. A. and Poirson, A. B., "Discrete analysis of spatial and spectral aliasing," *Investigative Ophthalmology and Visual Science, Supplement* **30**, 53 (1989).
19. Brainard, D. H. and Williams, D. R., "Bayes estimator for reconstruction from samples," *Investigative Ophthalmology and Visual Science, Supplement* **34**, 777 (1993).