Modeling, Calibration, and Rendition of Color Logarithmic CMOS Image Sensors

Dileepan Joseph, *Member, IEEE*, and Steve Collins

***Abstract*—Logarithmic CMOS image sensors encode a high dynamic range scene in a manner that roughly approximates human perception whereas linear sensors with equivalent quantization suffer from saturation or loss of detail. Moreover, the continuous response of logarithmic pixels permit high frame rates and random access, features that are useful in motion detection. This paper describes how to model, calibrate, and render pixel responses from a color logarithmic sensor into a standard color space. The work unifies color theory in conventional linear sensors and fixed pattern noise theory in monochromatic logarithmic sensors. Experiments with a Fuga 15RGB sensor demonstrate calibration and rendition using a Macbeth chart and neutral density filters. Color rendition of the sensor with an empirical model, tested over three decades of dynamic range, competes with conventional digital cameras, tested over 1.5 decades. Photodiode leakage currents complicate modeling and calibration and degrade rendition in dim lighting.**

***Index Terms*—Color rendition, fixed pattern noise, logarithmic pixels.**

I. Introduction

CMOS IMAGE SENSORS are displacing CCD sensors for several reasons, such as the integration of signal processing with image sensing on one die and the economies of scale available in the CMOS industry [1]. Research on CMOS imaging includes the development of sensors to capture high dynamic range scenes, where illuminances and reflectances span several decades, without saturation or loss of perceptible detail [2].

For reasons of high dynamic range and high frame rate, together with random accessibility of pixels, the automotive industry finds logarithmic imaging to be useful in the automation of traffic supervision and the improvement of road safety [3]. Human perception roughly approximates Weber's law, which states that the threshold to sense a difference between the illuminance of a fixation point and its surroundings is a fraction, about 1–10%, of the surrounding illuminance [4]. When illuminances are encoded by a logarithmic sensor, such a law makes the threshold for sensitivity constant, ideal for quantization. Moreover, instead of integrating photo-generated charge over a discrete period of time, as in a linear pixel, a logarithmic pixel converts photon flux continuously into a current, which means that an image is available at virtually any moment. Studies with pulsed lasers have shown a pixel bandwidth of 100 kHz, at normal light levels, that increases with illumination [5]. In practice, the speed of the readout circuit limits the frame rate.

Scenes observed by the human eye span over eight decades of illuminance, ranging from 10^{-3} lux in starlight to 10^{2}–10^{3} lux for indoor lighting, to 10^{5} lux for bright sunlight and to higher levels for specularities or direct viewing of bright sources (such as oncoming headlights or the sun) [2]. At any one time, within this range, the human eye can perceive illuminances spanning five decades [6]. Typical linear CCD and CMOS APS sensors may capture three decades of dynamic range whereas logarithmic CMOS sensors may capture over five decades [7].[1] Linear image sensors can adapt over a high dynamic range by aperture adjustment or global control of integration time but saturated patches of black or white appear when imaging a scene spanning a high dynamic range at once [2].

It would be very difficult, especially at video rates, for a linear image sensor to meet quantization requirements that satisfy Weber's law over a high dynamic range. For example, it would take 23 bits to quantise illuminance on a linear scale with 1% accuracy throughout a five decade range. Even if it were feasible to digitise a scene with 23 bits per pixel (per color channel), most of these bits would be wasted upon display of the image since human perception has less absolute sensitivity to bright illuminances than to dim ones. On the other hand, using fewer bits than required by Weber's law could result in a failure to capture perceptible detail, especially at dim illuminances. A good solution is to encode illuminances on a logarithmic scale so that a fractional threshold becomes a constant threshold, suitable for uniform quantization over a high dynamic range. On a logarithmic scale, capturing five decades of illuminance with 1% accuracy requires only ten bits of quantization.

Applications involving safety require a robust technology. One problem with logarithmic sensors is fixed pattern noise (FPN), a substantial but predictable error that appears in an image due to natural variations of device parameters from pixel to pixel [8]. While FPN correction is necessary to make logarithmic sensors useful, the accurate rendition of scenes on display devices by estimation of real world stimuli from pixel responses is also important. Rendition is more important with color images because the eye is more sensitive to chromatic errors than to intensity errors [4]. Much has been published

Manuscript received May 26, 2002; revised May 20, 2003. This work was supported in part by the Natural Sciences and Engineering Research Council (Canada) and in part by the Engineering and Physical Sciences Research Council (U.K.).

The authors are with the Department of Engineering Science, University of Oxford, Oxford, U.K. (e-mail: dileepan@robots.ox.ac.uk; collins@robots.ox.ac.uk).

Digital Object Identifier 10.1109/TIM.2003.818551

[1]Normally, in the literature, the dynamic range of a linear pixel is taken to be the ratio of the maximum voltage signal to the noise voltage level, whereas the dynamic range of a logarithmic pixel is taken to be the ratio of the maximum to minimum current signal in between which the pixel circuit maintains a logarithmic current-to-voltage relationship.

about color rendition in linear sensors but little has been written on rendition in color logarithmic sensors, the subject of this paper.

Section II unites color theory in linear sensors with monochromatic theory in logarithmic sensors to model color sensation in logarithmic sensors. Section III describes a procedure to calibrate this model and Section IV outlines a method to render the response of a calibrated sensor into a standard color space. Section V demonstrates calibration and rendition with a Fuga 15RGB logarithmic sensor, a color version of the Fuga 15d developed at IMEC [9], and compares color rendition of the sensor to that of conventional digital cameras.

II. Modelling

A color image sensor is made by inserting color filters in the path of light rays before they form an image on a monochromatic sensor [10]. Corresponding to human color vision, three filters are needed, selective to the red (R), green (G), and blue (B) regions of the spectrum. Multisensor imagers use prisms with special coatings to split and filter an image into three images, which are captured by separate sensors and combined to produce a single image. Single-sensor imagers have a pattern of red, green, and blue filters overlaid upon pixels. Though each pixel is selective to one color, its neighbors are selective to the others. By interpolating pixel responses, a red, green, and blue response may be estimated for each pixel at a loss of spatial resolution. As multisensor imagers obey a similar theory, the rest of this paper discusses only single-sensor imagers.

A color filter on a pixel modifies the spectral composition of incident light prior to absorption by the photodiode in the pixel. The photodiode absorbs the filtered light to varying degrees as a function of wavelength λ. Even attenuation in the lens of the camera is wavelength dependent. Equation (1) combines the spectral attenuations of the lens $g_L(\lambda)$, color filter $g_k(\lambda)$, with $k \in \{R, G, B\}$, and photodiode $g_P(\lambda)$ into one function $f_k(\lambda)$ [10]. Equation (2) uses $f_k(\lambda)$ to model the photocurrent I_k induced in a pixel by a spectral irradiance $s(\lambda)$ [1]

$$f_k(\lambda) = g_L(\lambda) g_k(\lambda) g_P(\lambda) \tag{1}$$

$$I_k = \int_0^\infty f_k(\lambda) s(\lambda)\, d\lambda. \tag{2}$$

A color image sensor need not estimate $s(\lambda)$ at each pixel to recreate the sensation of color implied by $s(\lambda)$ on a display device (i.e., a monitor or printer) [10]. In response to a spectral irradiance $s(\lambda)$, human perception of color may be described fundamentally by three numbers X, Y, and Z [11]. These numbers are inner products, over the visible spectrum, of $s(\lambda)$ and three basis functions $\bar{x}(\lambda), \bar{y}(\lambda)$, and $\bar{z}(\lambda)$, which were standardised by the Commission Internationale de l'Eclairage (CIE) in 1931. Normally, $f_R(\lambda), f_G(\lambda)$, and $f_B(\lambda)$ in (1) are designed to approximate linear combinations of $\bar{x}(\lambda), \bar{y}(\lambda)$, and $\bar{z}(\lambda)$ [10]. Therefore, I_R, I_G, and I_B in (2) may be modeled by linear functions of X, Y, and Z, as in (3), where $\mathbf{x}$ is a vector of X, Y, and Z values and $\mathbf{d}_k$ is a vector array of coefficients, called a *mask*, that relates the photocurrent I_k linearly to $\mathbf{x}$

$$I_k = \mathbf{d}_k \cdot \mathbf{x}. \tag{3}$$

Because the circuits for a color logarithmic sensor are identical to those of a monochromatic logarithmic sensor, the same equations relate the output of the sensor to the photocurrent in a pixel. By following the analysis of Joseph and Collins for monochromatic sensors [8], the digital response y of a color logarithmic pixel to a photocurrent I_k, given in (3), may be modeled by (4), where a, b, c, and ϵ are called the *offset*, *gain*, *bias*, and *error*, respectively. The offset depends on threshold voltages of the circuit, the gain depends on the subthreshold slope, the bias depends on the photodiode leakage current and the error depends on random noise

$$y = a + b\, \ln(c + I_k) + \epsilon. \tag{4}$$

III. Calibration

The model in (4) gives the response of a logarithmic pixel to irradiance focused upon it from a point in a scene. To recreate the scene on a standard display, an image must be rendered from pixel responses. Rendering accuracy depends on calibration of the parameters that relate the response of each pixel to real world stimuli. The calibration divides into two parts, one dealing with FPN and the other with color.

A. Fixed Pattern Noise

FPN arises in a logarithmic image sensor, resulting in nonuniform images of uniform surfaces, when a, b, c, or a combination thereof vary from pixel to pixel. This distortion is predictable and largely correctable. Joseph and Collins identify three types of FPN of interest [8]. The first involves offset variation, where only the offset varies with the pixel coordinate j in an array of N pixels, where $1 \le j \le N$. The second involves offset and gain variation and the third involves offset, gain and bias variation. The design and fabrication of a sensor may favor one of these types so all three are considered here.

To correct FPN, the varying parameters are estimated using images of uniform irradiance, preferably white in color, taken with M different intensities [8]. Indexing these images by i, where $1 \le i \le M$, the estimated response $\hat{y}_{ij}$ of the sensor (as opposed to the actual response y_{ij}, which includes an unpredictable error component ϵ_{ij}), is given in (5)–(7) for single, double and triple variation respectively. With $k \in \{R, G, B\}, k_j$ represents the pattern of overlaid color filters

$$\hat{y}_{ij} = a_j + b \ln(c + I_{ik_j}) \tag{5}$$

$$\hat{y}_{ij} = a_j + b_j \ln(c + I_{ik_j}) \tag{6}$$

$$\hat{y}_{ij} = a_j + b_j \ln(c_j + I_{ik_j}). \tag{7}$$

When pixels are partitioned by color, the calibration of FPN in a color sensor becomes the calibration of FPN in three monochromatic sensors. Following Joseph and Collins [8], parameters in (5)–(7) may be estimated by minimizing three sum square errors (SSEs) in (8) between the actual responses y_{ij} and the estimated responses $\hat{y}_{ij}$ for colors $k' = R, G, B$

$$\mathrm{SSE}_{k'} = \sum_{i=1}^{M} \sum_{j \mid k_j = k'} (y_{ij} - \hat{y}_{ij})^2. \tag{8}$$

Parameters estimated by minimizing the SSEs are not unique [8]. Note that (5) and (6) are invariant under transforms (9)–(11), with b replaced by b_j in the case of (6). Similarly, (7) is invariant under transforms (9) and (11), with b and c replaced by b_j and c_j respectively, but (10) does not apply

$$(a_j, b, c, I_{ik}) \equiv (a_j - b\ln\gamma, b, \gamma c, \gamma I_{ik}) \tag{9}$$

$$(a_j, b, c, I_{ik}) \equiv \left(a_j, \frac{b}{\gamma}, 0, (c + I_{ik})^{\gamma}\right) \tag{10}$$

$$(a_j, b, c, I_{ik}) \equiv (a_j, b, c - \gamma, I_{ik} + \gamma). \tag{11}$$

In each partition, estimation of parameters in (5)–(7) is limited by (9)–(11), [8]. Only those parameters that vary from pixel to pixel are determinate from images of uniform (but unknown) irradiance. For single variation in (5), the means of the offsets $\bar{a}_k$, the gain b and the bias c are indeterminate but the deviation of the offsets from the means, denoted $\hat{a}_j$ in (12), is determinate

$$\hat{a}_j \approx a_j - \bar{a}_{k_j}. \tag{12}$$

Similarly, for double variation in (6), the means of the offsets $\bar{a}_k$, the means of the gains $\bar{b}_k$ and the bias c are indeterminate. The estimated offsets and gains, denoted $\hat{a}_j$ and $\hat{b}_j$ in (13) and (14), are linear functions of the actual parameters

$$\hat{a}_j \approx a_j - b_j \frac{\bar{a}_{k_j}}{\bar{b}_{k_j}} \tag{13}$$

$$\hat{b}_j \approx \frac{b_j}{\bar{b}_{k_j}}. \tag{14}$$

For triple variation in (7), the means of the offsets $\bar{a}_k$ and the means of the biases $\bar{c}_k$ are indeterminate. The means of the gains $\bar{b}_k$ are determinate because transform (10) does not apply. The estimated offsets, gains and biases, denoted $\hat{a}_j, \hat{b}_j$, and $\hat{c}_j$ in (15)–(17), are linear functions of the actual parameters

$$\hat{a}_j \approx a_j - b_j \frac{\bar{a}_{k_j}}{\bar{b}_{k_j}} \tag{15}$$

$$\hat{b}_j \approx b_j \tag{16}$$

$$\hat{c}_j \approx e^{\frac{\bar{a}_{k_j}}{\bar{b}_{k_j}}} (c_j - \bar{c}_{k_j}). \tag{17}$$

Equation (18) estimates the standard deviation of the residual error $\hat{\sigma}_\epsilon$ for FPN calibration. The square of this measure equals the total SSE in (8) over the degrees of freedom (DOF), which is the number of constraints (i.e., MN) minus the number of parameters fitted to the same constraints (i.e., $3M + N, 3M + 2N$, and $3M + 3N$ for single, double and triple variation)

$$\hat{\sigma}_\epsilon^2 = \frac{1}{\mathrm{DOF}} \sum_{i=1}^{M} \sum_{j=1}^{N} (y_{ij} - \hat{y}_{ij})^2. \tag{18}$$

B. Color

Once the offset, gain, and bias parameters that vary from pixel to pixel are estimated, the non-varying mask parameters in (3) need estimation to render an image taken by a color logarithmic sensor into a standard color space. Color calibration of conventional linear sensors is done by imaging a color chart with patches of known color and using these ideal values and corresponding image data to estimate parameters of the color model [10]. The approach given below deals with a calibration using M images of the color chart, indexed by i, taken with different illuminant intensities to cover a wide dynamic range.

The images are segmented to locate and separate the pixels that correspond to each color patch. So that FPN does not corrupt segmentation, the images are first corrected using the results of FPN calibration as with monochromatic sensors [8]. After segmentation, the original response of each pixel in each color patch of each image may be estimated by (19)–(21) for single, double and triple variation, where $\mathbf{x}_{ij}$ is the ideal color vector of the jth pixel with the ith illuminant intensity

$$\hat{y}_{ij} = \hat{a}_j + \bar{a}_{k_j} + b\ln(c + \mathbf{d}_{k_j} \cdot \mathbf{x}_{ij}) \tag{19}$$

$$\hat{y}_{ij} = \hat{a}_j + \hat{b}_j \bar{a}_{k_j} + \hat{b}_j \bar{b}_{k_j} \ln(c + \mathbf{d}_{k_j} \cdot \mathbf{x}_{ij}) \tag{20}$$

$$\hat{y}_{ij} = \hat{a}_j + \hat{b}_j \frac{\bar{a}_{k_j}}{\bar{b}_{k_j}} + \hat{b}_j \ln\left(e^{-\frac{\bar{a}_{k_j}}{\bar{b}_{k_j}}} \hat{c}_j + \bar{c}_{k_j} + \mathbf{d}_{k_j} \cdot \mathbf{x}_{ij}\right). \tag{21}$$

Several unknowns remain from FPN calibration in (19)–(21). These parameters may be reduced, as in (22)–(24), by combining some unknowns together with substitutions (25) and (26) (with $\bar{b}_k$ instead of b, and $\bar{c}_k$ instead of c, as appropriate)

$$\hat{y}_{ij} = \hat{a}_j + b\ln\left(c'_{k_j} + \mathbf{d}'_{k_j} \cdot \mathbf{x}_{ij}\right) \tag{22}$$

$$\hat{y}_{ij} = \hat{a}_j + \hat{b}_j \bar{b}_{k_j} \ln\left(c'_{k_j} + \mathbf{d}'_{k_j} \cdot \mathbf{x}_{ij}\right) \tag{23}$$

$$\hat{y}_{ij} = \hat{a}_j + \hat{b}_j \ln\left(\hat{c}_j + c'_{k_j} + \mathbf{d}'_{k_j} \cdot \mathbf{x}_{ij}\right) \tag{24}$$

$$c'_k = e^{\frac{\bar{a}_k}{b}} c \tag{25}$$

$$\mathbf{d}'_k = e^{\frac{\bar{a}_k}{b}} \mathbf{d}_k. \tag{26}$$

The remaining parameters are estimated by minimizing the SSE in (27) between the actual response y_{ij} and estimated response $\hat{y}_{ij}$ for pixels of the color chart, given by the set $\mathcal{C}$. Minimizing the SSE for any of the models in (22)–(24) requires optimization as no analytic solution exists for all the unknowns. However, at the minimum of the SSE, b and $\bar{b}_k$ in (22) and (23) are given by (28) and (29), where l_{ij} is given in (30)

$$\mathrm{SSE} = \sum_{i=1}^{M} \sum_{j \in \mathcal{C}} (y_{ij} - \hat{y}_{ij})^2 \tag{27}$$

$$b = \frac{\sum_{i=1}^{M} \sum_{j \in \mathcal{C}} (y_{ij} - \hat{a}_j) l_{ij}}{\sum_{i=1}^{M} \sum_{j \in \mathcal{C}} l_{ij}^2} \tag{28}$$

$$\bar{b}_{k'} = \frac{\sum_{i=1}^{M} \sum_{j \in \mathcal{C} \mid k_j = k'} (y_{ij} - \hat{a}_j) \hat{b}_j l_{ij}}{\sum_{i=1}^{M} \sum_{j \in \mathcal{C} \mid k_j = k'} \hat{b}_j^2 l_{ij}^2} \tag{29}$$

$$l_{ij} = \ln\left(c'_{k_j} + \mathbf{d}'_{k_j} \cdot \mathbf{x}_{ij}\right). \tag{30}$$

Thus, only 12 variables that make up c'_k and $\mathbf{d}'_k$ need optimization for models (22)–(24). A suitable optimization algorithm is the conjugate gradients method [12]. Care must be taken to ensure that chosen parameters of c'_k and $\mathbf{d}'_k$, in either of (22)–(24), keep the argument of the logarithm positive. This can be done by making the SSE in (27) otherwise return a large value (∞ in Matlab) and ensuring that the line minimization used by the conjugate gradients method can cope with such extremes.

Equation (31) estimates the standard deviation of the residual error $\hat{\sigma}_\epsilon$ in color calibration. The square of this measure equals the SSE in (27) divided by the DOF, which is M times the cardinality of $\mathcal{C}$ minus the number of fitted parameters (i.e., 13, 15, and 12 for single, double and triple variation)

$$\hat{\sigma}_\epsilon^2 = \frac{1}{\text{DOF}} \sum_{i=1}^{M} \sum_{j \in \mathcal{C}} (y_{ij} - \hat{y}_{ij})^2. \tag{31}$$

IV. Rendition

The purpose of a color image sensor is to provide an image of a scene that is similar to the real scene when displayed. Thus, pixel responses must be rendered into a well-defined color space, such as CIE XYZ [11]. Denoting the offset, gain and bias parameters estimated by FPN calibration as $\hat{a}_j, \hat{b}_j$, and $\hat{c}_j$ and those estimated by color calibration as $\hat{b}$ (or $\hat{b}_k$) and $\hat{c}_k$, (32)–(34) estimate the response $\hat{y}_j$ of a logarithmic pixel for single, double and triple variation to a stimulus $\mathbf{x}_j$, with I'_j given in (35)

$$\hat{y}_j = \hat{a}_j + \hat{b} \ln(\hat{c}_{k_j} + I'_j) \tag{32}$$

$$\hat{y}_j = \hat{a}_j + \hat{b}_j \hat{b}_{k_j} \ln(\hat{c}_{k_j} + I'_j) \tag{33}$$

$$\hat{y}_j = \hat{a}_j + \hat{b}_j \ln(\hat{c}_j + \hat{c}_{k_j} + I'_j) \tag{34}$$

$$I'_j = \mathbf{d}'_{k_j} \cdot \mathbf{x}_j. \tag{35}$$

Rendering the response y_j into CIE XYZ space involves estimating the stimulus $\mathbf{x}_j$. First, I'_j is estimated by minimizing the SSE in (36) between the actual response y_j and estimated response $\hat{y}_j$ of the sensor. Such minimization amounts to inversion of (32)–(34), giving estimates $\hat{I}_j$ in (37)–(39)

$$\text{SSE} = \sum_{j=1}^{N} (y_j - \hat{y}_j)^2 \tag{36}$$

$$\hat{I}_j = e^{\frac{y_j - \hat{a}_j}{\hat{b}}} - \hat{c}_{k_j} \tag{37}$$

$$\hat{I}_j = e^{\frac{y_j - \hat{a}_j}{\hat{b}_j \hat{b}_{k_j}}} - \hat{c}_{k_j} \tag{38}$$

$$\hat{I}_j = e^{\frac{y_j - \hat{a}_j}{\hat{b}_j}} - \hat{c}_j - \hat{c}_{k_j}. \tag{39}$$

Next, $\hat{I}_j$ is interpolated to estimate red, green, and blue responses, denoted $\hat{I}_{jk}$, at each pixel j. Linear interpolation over a small neighborhood suffices as the responses of a pixel and its neighbors are highly correlated. Due to (35), the interpolated estimates depend linearly on the stimulus $\mathbf{x}_j$. Inverting this dependence in (40), using the estimated mask $\hat{\mathbf{d}}_k$ in matrix form, gives the desired estimate, denoted $\hat{\mathbf{x}}_j$, of the stimulus

$$\hat{\mathbf{x}}_j = \begin{pmatrix} \hat{d}_{1R} & \hat{d}_{2R} & \hat{d}_{3R} \\ \hat{d}_{1G} & \hat{d}_{2G} & \hat{d}_{3G} \\ \hat{d}_{1B} & \hat{d}_{2B} & \hat{d}_{3B} \end{pmatrix}^{-1} \begin{pmatrix} \hat{I}_{jR} \\ \hat{I}_{jG} \\ \hat{I}_{jB} \end{pmatrix}. \tag{40}$$

Images in CIE XYZ space may be rendered into other spaces. For example, $\hat{\mathbf{x}}_j$ in (40) may be rendered into CIE Lab space [11] or IEC sRGB space [13]. Euclidean distances calculated in Lab space correlate with perceptual differences. For computer hardware and software, however, the sRGB space of the International Electrotechnical Commission (IEC) has been accepted internationally as the default standard. Equation (41) estimates the standard deviation of the perceptual error $\hat{\sigma}_E$ between the ideal Lab vectors $\mathbf{z}_{ij}$ and the estimated Lab vectors $\hat{\mathbf{z}}_{ij}$ for the M calibration images of the color chart. The DOF in (41) is the same as in (31)

$$\hat{\sigma}_E^2 = \frac{1}{\text{DOF}} \sum_{i=1}^{M} \sum_{j \in \mathcal{C}} \|\mathbf{z}_{ij} - \hat{\mathbf{z}}_{ij}\|^2. \tag{41}$$

V. Experiments

Experiments were done with a Fuga 15RGB logarithmic image sensor, which had a 512×512 pixel array (i.e., $N = 512^2$). Rather than vary the intensity of the fluorescent illuminant, neutral density filters with nominal optical densities of 0.5, 1.0, 1.5, and 2.0 were used to simulate two decades of intensity variation. Effective illuminances were measured with a light meter for each filter and for the case of no filter.

A. Calibration

A sheet of white paper made a uniform scene for FPN calibration. Five images were taken (i.e., $M = 5$) using the neutral density filters to span two decades of illuminance. Following Section III-A, the single, double, and triple variation models were calibrated. The standard deviation of the residual error, given in (18), was 5.1, 2.2, and 0.59 LSB for these models. Thus, triple variation represents FPN well for the Fuga 15RGB.

Next, five images were taken of a Macbeth chart, created by McCamy *et al.* [14], which had 24 painted patches covering a wide gamut of colors. Using the neutral density filters to span two decades of illuminance, the images covered a dynamic range of 3.5 decades as the patches spanned 1.5 decades of reflectance. Following Section III-B, color calibration was performed. On average, there were 3 839 pixels in each of the 24 patches in each of the five images. The standard deviation of the residual error, given in (31), was 6.1, 3.9, and 9.4 LSB for single, double, and triple variation respectively.

Fig. 1 plots the standard deviation of the residual error, marked by circles, versus illuminance for color calibration. That triple variation performs much worse than single or double variation is surprising considering the residual error for FPN calibration is much better for triple variation. Investigation of the color chart data reveals that, as with the white paper data, triple variation models FPN better than single or double variation. However, the dependence in (4) of the digital response y on the photocurrent I_k is unsuitable for estimating color. A comparison of ideal colors with estimated colors suggests a model, given in (42), using the function in (43)

$$y = a + b \ln(c + f(I_k)) + \epsilon \tag{42}$$

$$f(I) = (\alpha + I)^\beta. \tag{43}$$

Assuming α and β in (43) are constant from pixel to pixel, replacing the theoretical model of (4) with the empirical model of (42) does not change the results of FPN calibration. The

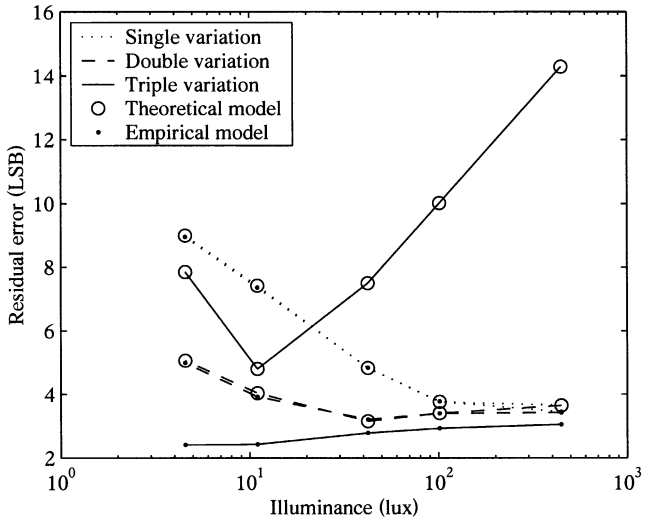

Fig. 1. Residual error of color calibration versus illuminant intensity for the single, double, and triple variation (theoretical and empirical) models.

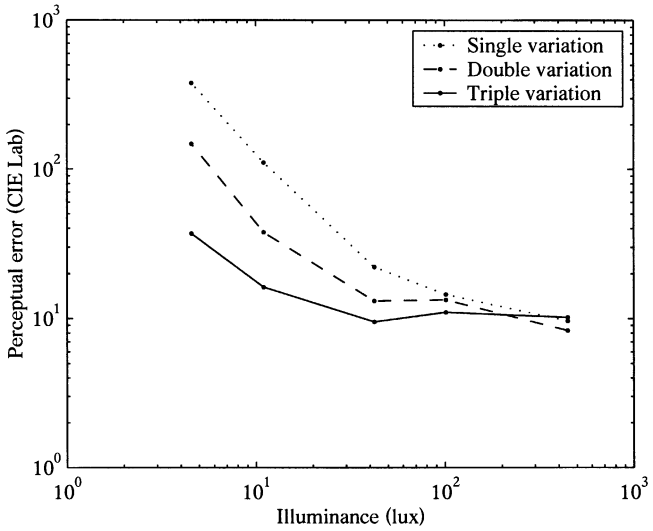

Fig. 2. Perceptual error of rendering a Macbeth chart versus illuminant intensity for the single, double, and triple variation (empirical) models.

unknowns I_{ik} in Section III-A are replaced by the unknowns $f(I_{ik})$ with no change to offset, gain and bias estimates. However, color calibration in Section III-B must estimate α and β by including them in the nonlinear optimization. As they modify the partial derivatives of the SSE in (27), these parameters affect the estimation of other parameters. Furthermore, the DOF in (31) and (41) must account for estimation of α and β.

Repeating color calibration with the empirical model results in a standard deviation of the residual error equal to 6.1, 3.8, and 2.7 LSB for single, double, and triple variation. Fig. 1 plots the standard deviation of the residual error, marked by dots, versus illuminance. Color calibration with the empirical model improves over the theoretical model substantially for triple variation but negligibly for single and double variation. The latter may be unable to discriminate I_k in (4) from $f(I_k)$ in (42) due to a higher residual error from FPN calibration.

The empirical model for triple variation shows a residual error for color calibration that is nearly flat across 3.5 decades of dynamic range (each point in Fig. 1 comprises 1.5 decades). However, with the theoretical or empirical model, the residual error increases with decreasing illuminance for single and double variation. This dependence suggests that bias variation, not considered by single and double variation, degrades color calibration mainly in dim lighting. For triple variation, the slight increase in error with increasing illuminance may be because the neutral density filters, used in taking the dimmer four images, did not really have flat spectral responses and thus modified the color of transmitted light in addition to the intensity.

B. Rendition

After FPN and color calibration, images taken by the Fuga 15RGB may be rendered into a standard color space such as CIE Lab or IEC sRGB, following Section IV for the theoretical model. For the empirical model, the rendering must include an inversion of (43). Using the empirical model, the standard deviation of the perceptual error, given in (41), between the ideal color chart and rendered images of the chart was 180, 69, and 20 for single, double, and triple variation respectively. Fig. 2 plots the standard deviation of the perceptual error versus illuminance. The figure shows how close the colors of the ideal chart match those of the rendered chart from the perspective of a standard observer (as defined by the CIE).

To put the performance of color rendition with the Fuga 15RGB in perspective, the perceptual error between an image of the ideal Macbeth chart and images of the chart rendered by conventional digital cameras was calculated from an article by McNamee [15]. The images were scanned with an HP Scanjet 5300C and converted from sRGB to Lab space. Table I lists the standard deviation of the perceptual error between pixels of the ideal chart and corresponding pixels of each camera's image.

Comparing Table I to Fig. 2 for single and double variation, color rendition is better with conventional cameras than with the Fuga 15RGB. For triple variation, color rendition of the Fuga 15RGB is comparable to conventional cameras except in dim lighting. Excluding the dimmest image, taken with five lux of illuminance, the standard deviation of the perceptual error is 12 with the Fuga 15RGB for triple variation. This result is comparable to the overall standard deviation in Table I, which equals 15. As the Macbeth chart spans 1.5 decades of reflectance and the Fuga 15RGB images, excluding the dimmest, span 1.5 decades of illuminance, color rendition of the logarithmic sensor, tested over three decades of dynamic range, competes with color rendition of conventional cameras, tested over 1.5 decades (McNamee used only one illuminance [15]).

The perceptual error in Fig. 2 increases with decreasing illuminance even for triple variation, which has a residual error in Fig. 1 that decreases with decreasing illuminance. In dim lighting, the bias c (and α) dominates the logarithm in (42) and makes the stimulus $\mathbf{x}$ in (3) difficult to estimate. In other words, photodiode leakage currents reduce the sensitivity of pixels to small photocurrents so that the random noise ϵ in (42) has a greater effect on the response than the stimulus. Decreasing the leakage current, increasing the photocurrent, or reducing the random noise should lessen this degradation. Decreasing the

TABLE I
PERCEPTUAL ERROR OF CONVENTIONAL DIGITAL CAMERAS BETWEEN IDEAL AND ACTUAL IMAGES OF A MACBETH CHART

Digital camera	Error (CIE Lab)
Kodak DCS 265	13
Nikon Coolpix 950	12
Olympus Camedia C-2000 Zoom	16
Canon Powershot Pro 70	17
Ricoh RDC 4200	13
Agfa ePhoto CL 50	15
Fuji MX 2700	15

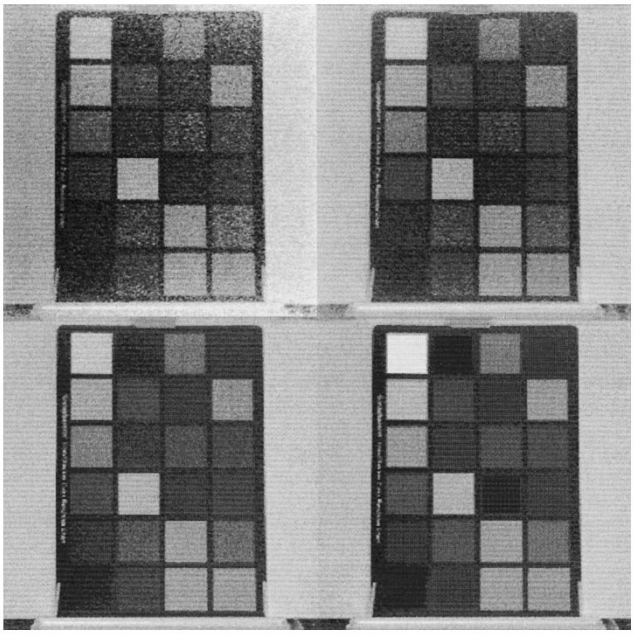

Fig. 3. Macbeth chart with an illuminance of 11 lux (incident on the scene) for single (top-left), double (top-right), and triple (bottom-left) variation rendition of a Fuga 15RGB image into IEC sRGB space. The bottom-right image uses ideal colors for the chart patches with the background of the triple variation image (please contact the author for a color version of the figure).

leakage current would also reduce the significance of bias variation, thereby improving the performance of double variation relative to that of triple variation.

Fig. 3 shows Fuga 15RGB images of the Macbeth chart, taken with an illuminance of 11 lux and rendered into sRGB space for single, double, and triple variation. The figure also shows an image of the chart with ideal values for the color patches. Two mechanisms lead to a deviation of the rendered images from the ideal. The first is a pixel-to-pixel variation that causes uniform textures to appear noisy, especially visible in the single and double variation results of the montage. The second is a color deviation that causes patches to look different from the ideal, discernible in the triple variation result where some patches have too much or too little brightness, even if the observer's eye could filter out the noise.

One reason for a color deviation may be that the mechanism of color sensation is not fully understood in the Fuga 15RGB sensor, evident by the use of an empirical model. Another reason is that, in dim lighting, the dominance of the leakage current over the photocurrent may lead to a biased estimate of the stimulus. Indeed, color matching is much better at higher illuminances, as shown in Fig. 2 (sRGB images are omitted for brevity). The fluorescent illuminant may be another reason as McCamy *et al.* recommended CIE Standard Illuminant C for use with the Macbeth chart [14]. Fluorescent illuminants have spectral irradiance functions with sharp peaks at certain wavelengths that frustrate color rendition and matching [11]. Furthermore, the neutral density filters were not perfectly neutral.

VI. CONCLUSION

Logarithmic CMOS image sensors have a capability to capture scenes bearing a high dynamic range of illuminance and reflectance in a manner that roughly approximates human perception [2]. Permitting high frame rates, they are an attractive technology for motion tracking in outdoor environments [3], [5]. However, research on color logarithmic sensors has been limited by a lack of theory and results on modeling, calibration and rendition of sensor responses in terms of a standard color space. This paper begins to address these problems.

A model for the response of a color logarithmic sensor to spectral irradiance was constructed by combining the color model of conventional linear sensors [10] with the monochromatic model of logarithmic sensors [8]. Thus, the digital response y of a logarithmic pixel to a color stimulus $\mathbf{x}$, given in CIE XYZ space [11], is modeled by $y = a + b\ln(c + \mathbf{d}_k \cdot \mathbf{x}) + \epsilon$ where $a, b, c, \mathbf{d}_k$, and ϵ are called the *offset, gain, bias, mask,* and *error*, respectively, with k identifying if the pixel is selective to the red, green, or blue regions of the spectrum.

Pixel-to-pixel variation of the offset, gain, bias or a combination thereof leads to FPN, which distorts an image in a repeatable and predictable way, most visible with uniform surfaces [8]. Calibration of the image sensor involves estimation of the model parameters. First, the varying parameters are estimated by partitioning pixels by color sensitivity and applying the method of monochromatic FPN calibration to each partition [8]. Second, the mask and other non-varying parameters remaining from FPN calibration are estimated using images of a reference color chart. Calibrated models may be used to render an image taken with the sensor into CIE XYZ space and then into other useful spaces, such as CIE Lab [11] and IEC sRGB [13].

Using neutral density filters to simulate varying illuminance, experiments were performed with a Fuga 15RGB sensor. A pixel-to-pixel variation of offset, gain and bias modeled FPN well, with a residual error of 0.59 LSB for FPN calibration of white paper. Color calibration of a Macbeth chart [14] showed that the theoretical model did not match the sensor response. An empirical model $y = a + b\ln(c + (\alpha + \mathbf{d}_k \cdot \mathbf{x})^{\beta}) + \epsilon$ worked well, with a residual error of 2.7 LSB for color calibration. The perceptual error with this model was 12, in Lab space, over three decades of dynamic range, comparable to conventional digital cameras over 1.5 decades. The perceptual error increased quickly below five lux of illuminance, possibly because leakage currents reduced the sensitivity of pixels.

Instead of focusing on analogue or digital methods to compensate for offset variation, research in logarithmic sensors should aim to minimize bias variation so that offset or offset and gain variation suffices to model FPN, and to minimize bias

magnitude, so that color rendition in dim lighting improves. As the mask depends on spectral responses of photodiodes and overlaid filters and does not seem to vary across pixels, it may be estimated once for a process (a common practice with conventional linear cameras [10]) rather than for every sensor. Furthermore, the effect of temperature on pixel responses needs examination.

REFERENCES

[1] T. Lulé, S. Benthien, H. Keller, F. Mütze, P. Rieve, K. Seibel, M. Sommer, and M. Böhm, "Sensitivity of CMOS based imagers and scaling perspectives," *IEEE Trans. Electron Devices*, vol. 47, pp. 2110–22, Nov. 2000.

[2] O. Yadid-Pecht, "Wide-dynamic-range sensors," *Opt. Eng.*, vol. 38, no. 10, pp. 1650–60, Oct. 1999.

[3] B. Hoefflinger, H.-G. Graf, U. Seger, and A. Siggelkow, "Imager for robust high-speed vision," in *Proc. Dedicated Conf. Robotics, Motion, and Machine Vision in the Automotive Industries*, Sept. 1995, pp. 289–93.

[4] R. M. Boynton, *Human Color Vision*. San Diego, CA: University of California, 1979.

[5] M. Tabet, N. Tu, and R. Hornsey, "Modeling and characterization of logarithmic complementary metal-oxide-semiconductor active pixel sensors," *J. Vac. Sci. Technol. A*, vol. 18, no. 3, pp. 1006–9, May–June 2000.

[6] G. W. Larson, H. Rushmeier, and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Trans. Visual. Comput. Graphics*, vol. 3, no. 4, pp. 291–306, Oct.–Dec. 1997.

[7] "HDRC VGA Imager and Camera Data and Features," Institute for Microelectronics Stuttgart, Tech. Rep., Sept. 2000. IMS Chips.

[8] D. Joseph and S. Collins, "Modeling, calibration, and correction of nonlinear illumination-dependent fixed pattern noise in logarithmic CMOS image sensors," *IEEE Trans. Instrum. Meas.*, vol. 51, pp. 996–1001, Oct. 2002.

[9] B. Dierickx, D. Scheffer, G. Meynants, W. Ogiers, and J. Vlummens, "Random addressable active pixel image sensors," in *Proc. SPIE*, vol. 2950, Oct. 1996, pp. 2–7.

[10] A. C. Luther, *Video Camera Technology*. Boston, MA: Artech House, 1998.

[11] A. K. Roy Choudhury, *Modern Concepts of Colour and Appearance*. Enfield, NH: Science Publishers, 2000.

[12] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.

[13] *Default RGB colour space—sRGB*, International Electrotechnical Commision, Oct. 1999 Document 61966.

[14] C. S. McCamy, H. Marcus, and J. G. Davidson, "A color-rendition chart," *J. Appl. Photo. Eng.*, vol. 2, no. 3, pp. 95–9, Summer 1976.

[15] M. McNamee, "A snapshot in time," *Digital Photo.*, vol. 13, pp. 32–8, July 1999.

[16] D. Joseph, "Modelling and calibration of logarithmic CMOS image sensors," Ph.D. thesis, Univ. of Oxford, Sept. 30, 2002.

Dileepan Joseph (M'96) received the B.Sc. degree in computer engineering from the University of Manitoba, Winnipeg, MB, Canada, in 1997 and the D.Phil. degree in engineering science from the University of Oxford, Oxford, U.K., in 2002. His doctoral research concerned logarithmic CMOS image sensors.

Presently, he works on vortex shedding flowmeters for the Invensys University Technology Centre for Advanced Instrumentation, Oxford, U.K.

Steve Collins received the B.Sc. degree in theoretical physics from the University of York, York, U.K., in 1982 and the Ph.D. degree from the University of Warwick, Warwick, U.K., in 1986.

From 1985 until 1997, he was with the Defence Research Agency on various topics including the origins of $1/f$ noise in MOSFETs and analog information processing. Since 1997, he has been with the University of Oxford, Oxford, U.K., where he has continued his interest in smart imaging sensors and nonvolatile analogue memories.