Reference data set for camera spectral sensitivity estimation

Maryam Mohammadzadeh Darrodi, Graham Finlayson, Teresa Goodman and Michal Mackiewicz

Reference data set for camera spectral sensitivity estimation

Maryam Mohammadzadeh Darrodi,1,* Graham Finlayson,1 Teresa Goodman,2 and Michal Mackiewicz1

1School of Computer Science, University of East Anglia, Norwich, Norfolk NR4 7TJ, UK
2National Physical Laboratory, Hampton Road, Teddington, Middlesex TW11 0LW, UK
*Corresponding author: m.darrodi@uea.ac.uk

Received June 5, 2014; revised November 6, 2014; accepted January 5, 2015; posted January 8, 2015 (Doc. ID 208071); published February 5, 2015

In this article, we describe a spectral sensitivity measurement procedure at the National Physical Laboratory, London, with the aim of obtaining ground truth spectral sensitivity functions for Nikon D5100 and Sigma SD1 Merill cameras. The novelty of our data is that the potential measurement errors are estimated at each wavelength. We determine how well the measured spectral sensitivity functions represent the actual camera sensitivity functions (as a function of wavelength). The second contribution of this paper is to test the performance of various leading sensor estimation techniques implemented from the literature using measured and synthetic data and also evaluate them based on ground truth data for the two cameras. We conclude that the estimation techniques tested are not sufficiently accurate when compared with our measured ground truth data and that there remains significant scope to improve estimation algorithms for spectral estimation. To help in this endeavor, we will make all our data available online for the community. © 2015 Optical Society of America

OCIS codes: (040.1490) Cameras; (120.4140) Monochromators; (120.4820) Optical systems; (150.1488) Calibration.

http://dx.doi.org/10.1364/JOSAA.32.000381

1. INTRODUCTION

The knowledge of camera spectral sensitivities is useful in many applications, for example, illuminant estimation [1], color correction [2,3], and multispectral imaging [4,5]. Camera manufacturers do not provide their data and so it is the user’s task to either measure or estimate their camera’s sensor sensitivities. As we will address in this paper, measuring camera spectral sensitivities is not an easy task. So, it is more practical to estimate—through statistical argument—the spectral sensitivities of a camera given, for example, the RGB response to known spectral stimuli. The need for spectral sensitivity estimation is increased because cameras made by the same manufacturer can have significantly different spectral sensitivities [6].

A standard methodology for measuring the sensor sensitivities is to record the response of the camera to monochromatic light produced by a monochromator. This can be carried out in specialist calibration laboratories with equipment suitable to transmit a mechanically selectable narrow waveband of light chosen from a wide range. Alternatively, one might use a set of narrowband interference filters [7,8] and a known light source to generate narrowband spectral stimuli. While this second approach is notionally easier, interference filters are expensive. Both approaches require significant lab-based measurements of the kind not available to most users.

In this paper, we describe a camera calibration procedure using a monochromator at the National Physical Laboratory (NPL), the UK National Measurement Institute. We discuss the measurement procedure in depth along with the setup of the system. The novelty of this approach arises from the fact that we report the measured spectral sensitivities along with the measurement uncertainties. The stated uncertainty values are a combination of camera and calibration uncertainties and are reported per measured wavelength. These values help us understand how close the measured sensitivity values are to the actual camera sensor sensitivity functions. To the best of our knowledge, this detailed uncertainty evaluation has not been addressed in other camera calibration methodologies reported in the literature. Prior calibration procedures discussed in the literature follow a much simpler setup, with less controlled conditions that may be easier and faster in practice [7,9]. The downside of their approach is that they are not in control of the variations and errors of the system and so cannot report the uncertainty of their recovery. One of the motivations for our detailed uncertainty driven approach is that we wish to provide the community with an accurate measured data set to aid the development of algorithms for spectral sensitivity estimation.

Algorithms for estimating the spectral sensitivities are found through regression by inverting the color formation equation with respect to the sensor sensitivity. The idea here is that simple image formation is linear, in the sense that the RGB recorded by a camera for given spectral stimuli is the integral of this spectrum with the 3 spectral sensitivity functions. Given enough spectral measurements and corresponding camera RGBs, the spectral sensitivities of the camera can be solved for through regression. One of the main difficulties in this approach is that the limited dimensionality of the reflectance spectra and the presence of noise in the measurements leads to an ill-posed problem [8,10]. Using more spectrally high-frequency color samples, such as LED-based emissive targets [11–13] or fluorescent color targets [14] have been shown to lead to some improvements. Equally, multiple LED illuminations [15] could be used to illuminate a target. Though, of course, more light implies a lengthier calibration.
In this paper, we review several regression techniques proposed in the literature. Novelty, we evaluate these methods with respect to the ground truth spectral sensitivities (e.g., using the same RGB data of the Macbeth ColorChecker [16,17] viewed under D65 that was used to validate our spectral measurements). Tested methods are Tikhonov regularization [18,19], Tikhonov based on derivatives [20], quadratic programming [21], and simple linear regression where the spectral sensitivities are assumed to lie in the basis of a linear model of small dimension [6,22].

We show that most of these techniques provide quick estimates of the spectral sensitivities that are not as accurate as the measured ones, though encouragingly good reasonable recovery is possible. Interestingly, for our test data sets, different algorithms sometimes work better than others, suggesting that no estimation technique is preferable. We hope that by making the data available to the community, we aid in the development of spectral estimation algorithms.

In Section 2, we present the background necessary for understanding camera spectral sensitivity estimation by regression. NPL calibration is explained in Section 3. Validation tests on ground truth and estimated sensors are explained in Section 4, followed by comparisons of the estimation techniques using ground truth spectral response functions. We conclude in Section 5.

2. BACKGROUND
Mathematically, the simplest color formation model for the ith sensor and jth pixel response $p_{ij}$ can be written as

$$p_{ij} = \int_\omega oE(\lambda)S_j(\lambda)Q_i(\lambda)d\omega, \quad i = 1, 2, 3, \tag{1}$$

where $E(\lambda)$ is the spectral power distribution of the scene illuminant, $S_j(\lambda)$ is the surface reflectance imaged at pixel $j$, and $Q_i(\lambda)$ is the spectral response of sensor $i$. The visible spectrum $\omega$ runs from 400 to 700 nm (we use the notation nm for nanometers). By sampling spectral quantities at 10 nm intervals, the integral in Eq. (1) can be replaced by a summation over the $m$ intervals ($m = 31$ in the example case) transforming Eq. (1) into

$$p_{ij} = \sum_{\ell=1}^{m} E(\lambda_{\ell})S_j(\lambda_{\ell})Q_i(\lambda_{\ell})\Delta\lambda, \quad i = 1, 2, 3, \tag{2}$$

where the scalar $\Delta\lambda$ denotes the nanometer sampling difference. Let us assume that $\Delta\lambda$ is incorporated in $Q_i(\lambda)$ for ith sensor and the multiplication of the $m \times 1$ vectors $E$ and $S_j$ for a single pixel can be rewritten as follows:

$$q = \text{Diag}(E)S_j. \tag{3}$$

In Eq. (3) and throughout this paper underscoring denotes vector properties, $q$ is an $m \times 1$ vector, and the Diag is an operator that converts any $m \times 1$ vector into an $m \times m$ diagonal matrix whose diagonal elements are the elements of the vector. For a single pixel and one sensor channel, Eq. (2) can be transformed into Eq. (4) which can be written in vector notation as follows:

$$p = z^T q. \tag{4}$$

Assuming we have $n$ known reflectance—i.e., $m \times n$ matrix $S$—then under a known illuminant we can write $C = (\text{diag}(E)S)^T$ where $C$ is the $n \times m$ color stimulus matrix, in which $z^j$ is the jth row. Thus, we can write the $n$ responses, $P$, for each sensor as

$$P = C q. \tag{5}$$

Equation (5) is very powerful as, in a single equation, it teaches that for known color signal spectra that induced known camera responses we might, at least in principle, solve for the spectral sensitivities through regression. Indeed, Eq. (5) is similar to the standard regression equation $Ax = b$. When there are more equations than unknowns ($n \geq m$) and $C$ is full rank, minimizing the sum of squares error between the sensor responses and their estimate, using the standard—closed-form—Moore Penrose inverse [23] illustrated in Eq. (6), estimates a solution for $q$.

$$\min_{q} ||Cq - P||^2 \Rightarrow q = [C^T C]^{-1} C^T P. \tag{6}$$

In Eq. (6), it is assumed that $C^T C$ is invertible ($n$ must be larger than $m$). In the case where there are fewer equations than unknowns ($n < m$), one condition is where the matrix is not invertible and there are infinite numbers of solutions. In this case, we assume that the sensor $q$ is in the space of the reflectances:

$$q = C' X, \tag{7}$$

in which $X$ is derived by applying Eq. (7) to Eq. (5) which substitutes to Eq. (9) as

$$X = [C'C]^{-1} P. \tag{8}$$

$$q = C'[CC']^{-1} P. \tag{9}$$

Of all sensors that solve Eq. (7) exactly, the solution in Eq. (9) is the solution with least norm [24].

A full discussion of the problem of inverting (ill-conditioned) matrices can be found in the literature [25]. Algorithms for spectral sensor estimations, in effect, try to formulate the regression problem so that the system of equations can be solved more stably.

A. Spectral Sensitivity Recovery

Many spectral estimation algorithms begin with Eq. (7) and attempt to provide solutions, which are more stable. We provide only a short summary of some of these estimation methods and we refer the reader to the original references for further explanation.

1. Tikhonov Regularization

Tikhonov [19] regularization adds a penalty term to the normal least-squares error in order to prevent jaggy (high norm) sensor recoveries:

$$\min_{q} ||Cq - P||^2 + r\|q\|^2. \tag{10}$$
where $\gamma > 0$ is the regularization parameter defined by the user. As the value of $\gamma$ increases the norm of the sensor becomes a significant part of the minimization and in effect forces a smoother solution. Tikhonov regularization also has a closed form solution [19]:

$$ q = (C^T C + \gamma I)^{-1} C^T P, $$

where $I$ denotes the $m \times m$ identity matrix. It is useful to plot the fitting error against the norm of the recovered error for different values of $\gamma$. This forms an L-curve [18], which for larger $\gamma$ shows an increase in the fitting errors and for smaller $\gamma$ the Tikhonov regularization resorts to simple least-squares. A good solution is one which makes both the recovery error and the norm of the sensor small which is usually the point on the L-curve closest to the origin (see [18] for more details).

2. Tikhonov Regularization based on Derivatives

Dyas [20] has proposed a variant of Tikhonov regularization which minimizes the norm of the second derivative solution. Minimizing the seminorms explicitly targets a smooth solution. The Dyas minimization is given below:

$$ \min_2 \| C q - P \|^2 + \| T q \|^2 $$

in which $T$ is the discrete 2nd derivative operator in matrix form:

$$ T = \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 \\ 1 \\ 1 \end{bmatrix}. \tag{13} $$

Again, the L-curve criterion is applied for finding the best $\gamma$ [18]. The closed form solution for Eq. (13) with an optimal $\gamma$ is

$$ q = (C^T C + \gamma T^T T)^{-1} C^T P. \tag{14} $$

3. Linear Models of Spectral Sensitivity

The adaption of a linear basis (in which spectral sensitivity must lie) on itself is a powerful constraint. Here, we assume that the spectral sensitivity can be approximated by a linear combination of $h$ number of basis functions as follows [5,10,22]:

$$ Q(\lambda) = \sum_{k=1}^{h} a_k B_k(\lambda). \tag{15} $$

where $B_k(\lambda)$ denotes the basis as a function of wavelength $\lambda$, and $a_k$ is a weighing coefficient. From all the choices of basis functions, a sine function with basis $B_k(\lambda) = \sin(k\lambda \sigma)$ is proven to perform best [22]. Jiang et al. [6] employed linear models that are inferred from the real measured spectral sensitivities. They carried out a principal component analysis (PCA) of a data set of 28 different cameras previously measured by [26].

Substituting Eq. (16) into matrix form we write

$$ q = B a, \tag{16} $$

where $B$ is the $m \times h$ matrix of basis functions, and $a$ is the $h$ element vector of coefficients. Substituting in our regression formulae we search to minimize

$$ \min_2 \| CBq - P \|^2. \tag{17} $$

4. Quadratic Programming

Quadratic programming was used by Finlayson et al. [21] to estimate a spectral sensitivity function subject to linear constraints. In essence, they adopted the standard quadratic error criterion

$$ \min_2 \| C q - P \|^2 \tag{18} $$

but added several linear constraints to simultaneously ensure the recovered spectral sensitivities were all positive, smooth, and unimodal (single peak). The positivity constraint is applied with

$$ q \geq 0. \tag{19} $$

Unimodality at index $z$ is applied with the following set of inequalities:

$$ q_l \leq q_l + 1, \quad l = 1, \ldots, z - 1, $$

$$ q_l \geq q_l + 1, \quad l = z, \ldots, m. \tag{20} $$

Smoothness is implemented by enforcing the recovered sensor to be a linear combination of a small number of low-order Fourier series basis functions. An example of sensor recovery using 5 basis functions is

$$ q = \sigma_1 + \sigma_2 \sin(\chi) + \sigma_3 \cos(\chi) + \sigma_4 \sin(2\chi) + \sigma_5 \cos(2\chi), \tag{21} $$

where $\sigma$ is a vector of basis weights and $\chi = (\lambda - 400)\pi/150$ for $\lambda = 400, \ldots, 700$ nm in 0 nm intervals. Subject to the above constraints, quadratic programming will find the optimal solution that minimizes the error in Eq. (18). Various choices of $\sigma$’s are tested to find the optimal solution with minimum error.

Finally, we remark that for most of the above-mentioned estimation techniques additional constraints can be easily added to enhance the performance of the algorithms. As an example, it is possible to implement Tikhonov-type regularizations subject to linear constraints. However, the objective in this paper is not to advance the state of the art in algorithm design but rather to accurately measure the spectral sensitivities of cameras and then benchmark the performance of existing published algorithms.

B. Number of Patches in a Color Target

Let us return to Eq. (6) where $C^T C$ is not invertible when it is low rank ($n < m$) and—at least from the simplest equation solving perspective—we are required to use the least-norm solution in Eq. (9). One might suppose that the solution to this
problem would be to calibrate the camera with more surface reflectances. However, since reflectances (and so color signals) are well described by low-dimensional linear models \([27-29]\) the effective dimension of \(C^tC\) is low (the matrix has a very high condition number and so its inverse is not stable). The import of this is that very small perturbation to the measured camera data can result in very different recovered sensors:

\[
[C^tC]^{-1}C^tP \neq [C^tC]^{-1}C^t(P + \varepsilon).
\]  

(22)

Of course, having more reflectances may still help. Let us test this idea. In simulation, we use Eq. (5) to numerically calculate the synthetic camera response for the 24 color patches of the Macbeth ColorChecker \((n < m) [16]\), and then for a target comprising the 1995 samples \((n > m) [17]\). In carrying out the numerical integration we used the Nikon camera sensitivities (measured in this paper) and a D65 illuminant. Importantly, each reflectance is represented in a simulated image as a 10 x 10 pixel patch. To each patch, a small amount of random shot noise (less than 1%) has been added per pixel of the synthetic camera response image. We then average over each patch before estimating the sensor spectral sensitivity using Tikhonov regularization. In Fig. 1, we see that, visually at least, the recovery using 24 or 1995 reflectances is similar. Thus, the 1995 values do not result in significantly more accurate recovery, assessed using the Vora value [30]. This is a metric that can be used to measure how well a test spectral sensitivity set can be used to predict the actual responses for a reference sensor and will be discussed later in this document. The Vora values close to 0 indicate two sensor sets ample light identically. Here, the Vora value is calculated between the estimated sensitivities recovered using Tikhonov regularization (for 24 or 1995 reflectances) and the actual Nikon spectral sensitivities. For both reflectance sets, the Vora values are almost the same: 5.3 and 5.5, respectively. For all practical purposes, these two estimated sensor sets are the same. Thus, we conclude that a small number of reflectances provide equally good spectral sensitivity estimation. Moreover, measuring the color signal spectra reflected from a target is time consuming. Thus, in this paper, like previous authors [20,21], we carry out calibration using the 24 patch Macbeth ColorChecker.

3. NATIONAL PHYSICAL LABORATORY SPECTRAL SENSITIVITY MEASUREMENT SYSTEM

In this section, we provide a detailed description of the detector spectral sensitivity calibration facility of the NPL in London. The system consists of a tungsten halogen lamp (operated at a correlated color temperature of approximately 2580 K), a double monochromator, which selects a narrowband of wavelengths from the lamp, and optics, which direct the emergent optical radiation from the monochromator to either a reference silicon photodiode or the detector being calibrated (see Fig. 2). The optics are housed within a blackened enclosure, which prevents ambient light from reaching the detectors, and baffles are used to reduce the effects of any scattered light within the enclosure.

The system is based on the principle of substitution, in which the detector to be calibrated (the test detector) and the silicon reference detector are placed on a linear stage and irradiated sequentially by radiation from the monochromator; in this way, the calibration is transferred from the reference detector to the test detector. A slightly different optical arrangement was used for the calibration of the cameras. As in the usual setup, a concave mirror was used to image the exit slit of the monochromator onto the reference silicon photodiode. However, instead of direct substitution of the camera for the reference detector, a white tile was moved into the beam, in order to provide a uniform radiance field which was then imaged by the camera. The reflectance of the white tile in the 0/45 geometry used (i.e., irradiation at 0°, viewing at 45°) was calibrated at NPL and was allowed for when determining the sensitivity of the camera [31].

The calibration methodology was as follows: first, the monochromator was set to the required wavelength and an image of the white tile was taken with the camera. The shutter was then closed and a dark image of the tile was captured with the camera. The white tile was then moved out of the monochromator beam, by means of the translation slide, and light and dark readings were taken using the reference detector. Finally, the monochromator was set to the next wavelength and the process repeated. Apart from the camera, which was controlled independently using a laptop PC, the rest of the system was automated and controlled by the laboratory computer.

![Fig. 1. Spectral response function of the Nikon camera measured at NPL (solid line) and estimated using Tikhonov’s method (dashed line) using reflectance data of 24 patches of the Macbeth ColorChecker and 1995 samples of the online database.](image1.png)

![Fig. 2. Illustration of the calibration setup of the NPL automated spectral sensitivity measuring system and test camera.](image2.png)
According to Eq. (2), the camera signal \( P \) depends on the spectral radiance of the radiation incident on the camera at the measurement wavelength and the sensitivity of the camera at that wavelength. The radiance is measured using the reference silicon photodiode whose sensitivity at each wavelength has been calibrated against SI units [32], allowing for the reflectance of the tile.

This is denoted by \( L_i \) for the \( i \)th wavelength and (assuming that the diffuser behaves as an ideal cosine reflector) is calculated using

\[
L_i = p_i \delta_i / q_i \pi, \tag{23}
\]

where \( p_i \) is the dark-corrected signal from the reference detector, \( q_i \) is the calibrated irradiance sensitivity of the reference detector, and \( \delta_i \) is the reflectance of the white tile. Thus the absolute spectral radiance sensitivity for a camera sensor (R, G or B) at \( i \)th wavelength, denoted as \( q_i^\lambda \), can be calculated using

\[
q_i^\lambda = p_i / L_i \tag{24}
\]

where \( p_i \) is the dark-corrected camera signal of the sensor, which is derived by subtracting the dark camera signals from light signals and averaging over the central area of the captured images.

Note that Eq. (24) relates to the absolute spectral radiance sensitivity of the camera sensor expressed in units of signal per watt per meter squared per steradian. Absolute spectral sensitivity measurements would require the substitution of the reference detector and the white tile to be made in the same plane at the focus of the mirror, in which position the irradiated area was relatively small and nonuniform. In practice, lower calibration uncertainties could be achieved by working in terms of relative spectral sensitivities, which meant that the white tile could be placed closer to the monochromator as shown in Fig. 2, generating a larger and more uniform measurement area onto which the camera was focused. The camera signal value over all wavelengths was therefore normalized by the maximum signal for each channel before performing the calculation given in Eq. (24), yielding the relative spectral sensitivity function; note this is dimensionless, i.e., has unit \( 1 \).

Two different cameras were calibrated using this system: a Nikon D5100 and a Sigma SD1 Merrill. Measurements were made over the wavelength range of 380–770 nm, at 5 nm intervals for the Nikon camera calibration and at 10 nm intervals for the Sigma camera calibration. (The difference in the wavelength interval used for the two cameras was only due to lab time constraints. In practice, more than 2 \( \times \) 78 initial test measurements were taken for the Nikon camera and 2 \( \times \) 40 measurements for the Sigma. This was to configure the calibration setup and exposure settings of the two cameras.) The camera readings were averaged over a 21 \( \times \) 21 pixel patch in the center of the image, in order to reduce the effects of noise on each pixel. Although averaging over a larger area would have resulted in a further reduction in the noise effects, it would have introduced additional measurement uncertainties, such as image nonuniformities, which arise from slight differences in the individual pixel response across the sensor and small nonuniformities in the radiance across the tile.

The cameras were set to full aperture in each case. One of the challenging parts of the calibration was to set the camera shutter speed (exposure time) to avoid saturation of the spectral sensitivity of the sensors. Ideally, the same exposure time would be used for all wavelengths and for all 3 sensors. However, it was found that the range of radiances used and the differing responsivities of the sensors meant that there was insufficient dynamic range available to achieve this. Therefore, an initial measurement was carried out to determine a camera exposure which performed reasonably well on all three channels, i.e., which minimized the range of wavelengths for which a saturated signal was recorded while ensuring that a measurable signal, above the noise floor, was recorded for regions where each sensor was sensitive. This was found to be a 15 s exposure time. This setting caused some saturation with the Nikon camera in the region of 580–630 nm for the red channel. This was not the case for the Sigma camera. Therefore, it was decided to repeat the measurement on the Nikon camera with a faster shutter speed (5 s) for the saturated wavelength region in order to recover the lost information due to saturation for that channel. This second data set was collected over a slightly wider region (550–680 nm) and the overlapping results in the nonsaturated region (i.e., from 550 to 580 nm and 630 to 680 nm) were used to determine the scaling ratio between the two sets of results. Informally, we call this stitching the two data sets together. Figure 3(a) illustrates the two measured data sets for the Nikon.

The scaling ratio value was determined from the ratio \( Z_\lambda \) in the overlap regions:

\[
Z_\lambda = q_\lambda^r / q_\lambda^{*r}. \tag{25}
\]

where \( q_\lambda^r \) indicates the red channel spectral response measured with 15 s exposure time and \( q_\lambda^{*r} \) with 5 s. Figure 3(b) illustrates this ratio with the dashed ellipsoids indicating the overlapping region of the two data sets. The average value of \( Z_\lambda \) over the region shown with dashed ellipsoids is equal to the reciprocal of the scaling ratio \( r \). Thus,

\[
1/r = Mean(Z_\lambda) \quad 550 \leq \lambda \leq 580 \quad \text{and} \quad 630 \leq \lambda \leq 680. \tag{26}
\]

which is then applied as follows:

\[
q_\lambda = \begin{cases} 
  q_\lambda^r & 550 \leq \lambda \leq 580, \\
  r q_\lambda^{*r} & 585 \leq \lambda \leq 625, \\
  q_\lambda^{*r} & 630 \leq \lambda.
\end{cases} \tag{27}
\]

resulting in the final spectral sensitivity function of the red sensor. The expected value of \( r \) is 3.0 (the ratio of the two shutter speeds) and this was indeed the value obtained in our measurements. However, the value of \( r \) was not assumed equal to 3 in case of shutter timing errors or other possible camera errors. The spectral response function determined for the red channel of the Nikon camera before and after stitching is illustrated in Fig. 3(c). Although this method was time consuming, it was found to be the most reliable way of calibrating the spectral responsivity of the camera sensors using the NPL facility.

The measured spectral sensitivity functions of both cameras are subject to a certain level of uncertainty due to possible calibration errors. One of the motivations for calibration at NPL was to be able to determine these errors and uncertainties for the first time, to the best of our knowledge. This also enabled us to assess how well the functions calculated using the sensor sensitivity estimation algorithms
A. Calibration Uncertainties

A number of factors need to be considered when evaluating the measurement uncertainty associated with the spectral responsivity calibration of the camera sensors. These can be summarized as follows:

- **Uncertainty associated with the reference detector.** This arises from the uncertainty associated with the calibration of the reference detector by comparison with the NPL spectral responsivity scale and additional uncertainties that arise due to possible drift in the detector between its time of calibration and time of use. The combined 2σ uncertainty for these two effects varies from about 0.6% at 380 nm to less than 0.2% for wavelengths from 500 to 800 nm [33].

- **Uncertainty associated with the spectral reflectance of the white tile.** This arises from the calibration of the white tile on the NPL National Reference Reflectometer in the 0/45 geometry used and drift in the reflectance since the time of calibration, and is 0.75% for wavelengths between 400 and 800 nm, increasing to 1.35% for shorter wavelengths [31].

- **Uncertainty associated with the NPL spectral responsivity facility.** This includes uncertainties due to the scaling factor (used in the saturated region for the Nikon’s red sensor), nonuniformity, dark noise, nonlinearity and repeatability (note that by using the raw camera signals, uncertainties due to white balance were eliminated). It is important to appreciate that these effects may have a significant impact on subsequent measurements using the camera, particularly in terms of the absolute ratios between the sensors, and therefore some normalization of the absolute RGB values (e.g., by referencing to a calibrated white tile illuminated with white light source) was required when applying the calibrated camera sensitivity functions for measurements on spectrally broad samples (see Section 4 for further explanation). Below, we explain each of these uncertainties associated with the camera in detail.

- **Uncertainty due to scaling factor.** The uncertainty associated with the scaling factor r from Eq. (25) was determined from the standard deviation of the individual ratios Z of Eq. (25) in the overlap region, expressed as a percentage of the mean value. This uncertainty is around 3% which was only applicable for wavelengths between 580 and 628 nm for the Nikon’s red sensor. This uncertainty for other wavelengths and other sensors is zero.

- **Uncertainty due to nonuniformity effects.** As mentioned earlier, we averaged the pixel response p_j over the calibration area. The uncertainty due to pixel–pixel variations of the pixel response p^j can be calculated from the following for each sensor:

\[
\frac{\sigma}{\sqrt{n}} \cdot \tilde{p}.
\]  

where n is the number of pixels in the calibration area (which is 21 x 21 pixels), \(\tilde{p}\) is average of the dark-corrected signal (over the 21 pixel x 21 pixel area), and \(\sigma\) is the standard deviation of \(p^j\) over the calibration area of the dark-corrected image. The outcome of dividing the standard error of the mean

![Figure 3](image-url)
by the averaged signal is the relative nonuniformity uncertainty for that image. Calculating this value for each of the calibration images captured at each wavelength gave an estimate of this uncertainty over the spectrum. It should be noted that nonuniformities in the radiance of the surface imaged by the camera (i.e., the white tile irradiated by the light coming from the monochromator) were negligibly small over the 21 x 21 pixel area. In regions of the spectrum where the camera sensor sensitivity was low, the apparent nonuniformity was dominated by noise and in these cases only the residual uncertainty, after allowing for noise effects, was included in the uncertainty budget.

- **Uncertainty due to dark noise.** This uncertainty was measured in the same way as the nonuniformity except that the standard error was obtained from the dark-reading images taken at each wavelength \( \sigma_{\text{dark}} \) in Eq. (28) resulting in the relative dark noise uncertainty for that wavelength.

- **Uncertainty associated with sensor nonlinearity.** The impact of any sensor nonlinearity on the calibration results was minimized by ensuring that the input radiance varied by less than 100:1 for all wavelengths over the full visible wavelength range; in other words, the radiance was closely the same at all wavelengths. The uncertainty associated with sensor nonlinearity was evaluated by examining the relationship between the camera signal and the sample reflectance for captured images of a series of achromatic samples from the Macbeth ColorChecker illuminated by the D65 illuminant. We examined this for three different levels of exposure: 1/40, 1/20, and 1/10 s for the Sigma camera and 1/40, 1/25, and 1/15 s for the Nikon. Figure 8 in Appendix A illustrates the results of the linearity test for the Nikon. The points fell closely on a straight line for all three sets of exposures, indicating that the camera response is linear. Similar results were obtained for the Sigma camera. The maximum separation of the points from the ideal straight line in this graph is treated as the bounds of a rectangular uncertainty distribution function, so allowing the standard uncertainty associated with camera nonlinearity to be evaluated using Eq. (28) with \( n = 3 \) when \( R, G, \) and \( B \) are between [0,1] (see Table 1). Note that the same uncertainty was used for all wavelengths and for each camera channel.

- **Uncertainty associated with camera repeatability.** This was evaluated by capturing 10 images of the same scene (Macbeth ColorChecker illuminated by the illuminant D65). For each data set, the maximum standard deviation of the dark-corrected camera signals of the achromatic colors across the 10 images was used in Eq. (28) (with \( n = 10 \)) to result in the relative standard uncertainty associated with camera repeatability. This uncertainty was determined for each of the channels for each camera (see Table 1), and the same value assigned to all wavelengths for that channel.

The root of sum of squares of the camera relative uncertainties listed above is calculated to estimate the total camera uncertainty. Finally, the root of sum of squares of the relative uncertainties associated with the reference detector, white tile, responsivity facility and camera is calculated to give the uncertainty value for the entire calibration process. Figure 9 in Appendix A illustrates the detailed uncertainty values for every wavelength of the Nikon’s red sensor response. Figure 4 illustrates the measured spectral response function of the Nikon camera, with error bars marking the uncertainties that are higher than 0.5% for the high-frequency region of 450–650 nm. Similarly, Fig. 5 illustrates the measured spectral response function of the Sigma camera, with error bars marking the uncertainties that are higher than 1%.

These graphs are a simple illustration of the benefit of the calibration approach at NPL. Full uncertainty data are available online.

Note that the uncertainties in the regions of low responsivity are relatively high. However, ultimately what we require is the uncertainty associated with the spectrally integrated RGB values. This integration has the effect not only of “averaging out” uncertainties due to noise (which are the dominant uncertain-

![Fig. 4. Measured Nikon camera sensitivities from NPL calibration facility. Uncertainty levels over 0.5% are illustrated with black error bars across 450–650 nm.](image)

<table>
<thead>
<tr>
<th>Table 1. Table of Linearity and Repeatability Uncertainty Percentages for the Nikon and Sigma Cameras* b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearity</td>
</tr>
<tr>
<td>——</td>
</tr>
<tr>
<td>Nikon</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Sigma</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*The linearity test was carried out on the average camera signals of 10 images of the Macbeth ColorChecker Chart for 3 sets of exposures (1/40, 1/20, and 1/10 s for Sigma and 1/40, 1/25, and 1/15 s for Nikon).

b The repeatability test was carried out using the 10 images taken at 1/40 s exposure with both cameras. The values in this table are relative uncertainty percentages.
ties in the regions of low response) but of also giving greatest weight to the uncertainties in the regions of highest response (since these regions make the greatest contribution to the integrated value). Thus, we can consider the uncertainties associated with the RGB values as being similar to the uncertainties associated with those regions of high response, i.e., around 2% in most cases, and around 3.5% for the Nikon red sensor. Having these low levels of uncertainties for both cameras marks the NPL measured spectral response functions as a robust source, reliable enough to be evaluated and applied as ground truth data for comparison purposes in the next section.

4. EXPERIMENTS AND RESULTS

A. Validation of NPL Measured Spectral Response Functions (Outside of NPL)

In this section we explain a validation test for the measured spectral sensitivity functions of the Nikon and Sigma cameras from NPL. We wish to test how close the real camera responses that we capture outside NPL are to those predicted using the NPL sensors and numerical integration.

Real data. The 24 × 3 raw response for n color samples is stored in matrix P. Here, this is obtained from the camera’s (Nikon and Sigma) response to an image of a Macbeth ColorChecker, which was placed approximately in the center of the floor of a VeriVide cabinet, facing upward toward the D65 illuminant. The position of the camera relative to the ColorChecker was chosen to be approximately at 45 deg. A dark reading was taken by capturing an image while the camera lens cap was on. For each patch in the checker, the camera signal was obtained by averaging the dark-corrected images for each of the red, green, and blue channels over the central area of the patch resulting in a P_{24×3} matrix for the 24 colors. This real data set is also used to evaluate our camera sensors (see Sections 4.2 and 4.3).

Estimated data. A PR650 was set in the same position as the camera explained for real data collection (approximately at 45 degrees relative to the ColorChecker placed in a VeriVide cabinet) to obtain measured C_{24×3}. Note that, rather than moving the PR650, the Macbeth ColorChecker was moved into center of PR650’s focus for each color patch measurement in order to keep the light intensity constant across all 24 samples. Then P_{24×3} is numerically calculated using Eq. (5).

To perform this validation, all rows of P and ̂P are divided by the corresponding RGB values of the white in that matrix. As a result, both matrices would have values equal to one for R, G, and B response that represent the white and the rest of the rows represent the relative RGB values to white (i.e., for each channel response we have p_j/p_{white}, j = 1, …24). Relative estimation error for each R, G, or B channel’s response p and estimated response ̂p is calculated as

$$E = 100 \times \frac{|p - ̂p|}{|p|}.$$  \hspace{1cm} (29)

Table 2 illustrates the relative prediction error percentages for the Nikon and Sigma’s validation experiment averaged over RGB channels.

The prediction error is on average 1.5% for Nikon and 2.1% for Sigma, correlating with less than 3 CIE Lab Delta E values. Thus, the difference between the response of the camera predicted by the camera sensitivity functions measured at NPL and the real measured response is small. The experiment validated the sensor sensitivities recovered at the calibration lab at the NPL.

B. Evaluation of Sensor Sensitivity Estimation Algorithms

The different algorithms described in Section 2.2 have been implemented and tested. In order to compare the methods across the Nikon and Sigma data sets, the parameters in all estimation techniques were set to those values that worked well. Seven dimensional sine basis functions [22] and 5 Fourier basis for quadratic programming were found sufficient. In Jian et al.’s study [6] 2 basis functions were found to perform well for estimating the Nikon camera sensor estimation. We do not use the PCA for Sigma sensor estimation, as they are not well modeled.

We propose that the uncertainty levels measured at NPL for each wavelength can be used as a standalone spectral response function over the spectrum. Here we treat the upper and lower uncertainty envelopes as two separate spectral response functions, and we denote them as UB and LB, respectively. UB is the measured sensor plus 2 standard deviations and LB is the measured sensor minus 2 standard deviations. In essence, we just take the points at the top and bottom bars of Figs. 4 and 5. The UB and LB sensors represent a worse case measurement scenario where all the uncertainty manifests itself in sensor departure. The UB and LB sensors provide a reasonable goal to which spectral estimation algorithm should aim (i.e., recovery modulo the worse case sensor measurement uncertainty).

Here we also introduce an additional noise-free synthetic data set to accompany the spectral estimation using real data. The synthetic version of matrix C_{24×3} is a result of applying previously measured D65 light source and 24 reflectance of the Macbeth ColorChecker to Eq. (3). Then with q applied

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>Synthetic</th>
<th>Real</th>
<th>Synthetic</th>
<th>Real</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tikhonov</td>
<td>0.7</td>
<td>1.4</td>
<td>0.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Tikhonov (derivatives)</td>
<td>0.5</td>
<td>1.4</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>Sine basis-7</td>
<td>0.3</td>
<td>1.1</td>
<td>0.7</td>
<td>1.0</td>
</tr>
<tr>
<td>PCA basis-2</td>
<td>1.0</td>
<td>2.1</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>QB-2</td>
<td>0.6</td>
<td>1.2</td>
<td>0.4</td>
<td>0.7</td>
</tr>
<tr>
<td>UB</td>
<td>0.2</td>
<td>1.4</td>
<td>0.3</td>
<td>2.1</td>
</tr>
<tr>
<td>LB</td>
<td>0.2</td>
<td>1.6</td>
<td>0.4</td>
<td>2.3</td>
</tr>
</tbody>
</table>

*Columns represent synthetic and real data from the Nikon and Sigma cameras. Numbers accompanying the estimation technique names represent the number of basis used.*
to Eq. (5), synthetic data (with zero noise) is derived. The results of the estimation algorithms are reported in Table 3 for real and synthetic data in which the values indicate the averaged relative prediction error percentages over the RGB channels of the Nikon and the Sigma.

In comparison with Table 2 we can see the estimation algorithms provide plausible predictions for the camera response. In most cases the prediction error for the synthetic data is less than real data. Of course it is possible to produce sensor responses using a sensor whose spectral shape is far from the actual sensors.

C. Spectral Recovery Error

In order to compare the estimated spectral shapes with the measured spectral sensitivity functions from NPL, spectral error for each of the R, G, and B channels is calculated using the following:

$$SE = 100 \times (\|\hat{q} - q\|/\|q\|),\quad (30)$$

where for each channel, $q$ is $m \times 1$ and denotes a single sensor’s measured spectral sensitivity (from NPL), and $\hat{q}$ denotes the estimated spectral sensitivity rescaled regarding to $q$ using the following:

$$\hat{q} = \hat{q}(\max(q)/\max(\hat{q})).\quad (31)$$

The averaged SE values over the R, G, and B sensors are illustrated in Table 4. Spectral error is a pessimistic measure, which equivalently measures the predicted and actual responses for monochromatic stimuli across the visible spectrum [30]. Vora and Trusell’s measure calculates the difference in spectral response for all possible color signals using (see [27] for discussion)

$$Vora = 100 \times (1 - (trace(QQ^t\hat{Q}\hat{Q}^t))/3),\quad (32)$$

in which $\hat{Q}^t$ indicates the Moore–Penrose inverse [24] illustrated in Eq. (7). Vora values close to 0 indicate estimated sensors close to measured ground truth. Vora values are reported in Table 5.

The Vora values inform us that, for example, Tikhonov estimates of the Nikon sensor on real data perform well in predicting data outside the measuring data set. Notice that spectral error and Vora values of UB and LB sensors are much lower than the sensitivity estimation algorithms. There is considerable room for improvement in spectral estimations.

D. Discussion

From Tables 4 and 5 we observe the differences in the performance of estimation methods across the different data sets by comparing the SE and Vora values between real and synthetic data for each camera. Over all, sensor estimation using sine basis functions and Tikhonov based on derivatives perform better than the rest across all data sets. However, the objective in this paper is not to rank the estimation algorithms with one another but rather to benchmark their performance using the ground truth spectral sensitivity functions. Also, we can see that the values of the UB and LB rows are smaller than other estimated sensors in Tables 4 and 5. The prediction error in Table 3 is sometimes worse for LB and UB. This indicates the peril of over fitting data (and the instability of using predicted response to determine how well spectral estimation algorithms work).

In Figs. 6 and 7 we choose to illustrate the estimated Nikon’s red sensor for the real data of the Nikon and Sigma cameras using sine basis functions. The shape of the estimated spectral sensitivity function is similar to the NPL ground truth data. However, we can see that it does not completely overlap the ground truth sensitivities from NPL over

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>Nikon</th>
<th>Sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic</td>
<td>Real</td>
<td>Synthetic</td>
</tr>
<tr>
<td>Tikhonov</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>Tikhonov (derivatives)</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>Sine basis-7</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>PCA basis-2</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>QP-5</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>UB</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LB</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

*Columns represent synthetic and real data from the Nikon and Sigma cameras. Numbers accompanying the estimation technique names represent the number of basis used.*

Fig. 6. Illustration of the red channel spectral response function recovered from NPL (solid line) and estimated Nikon sensor sensitivities recovered by 7 sine basis (dashed) using real data from the Macbeth ColorChecker.
the entire spectrum. For the Nikon’s red sensor, as an example, we have found that the estimated sensor using sine basis exceeds the uncertainty boundaries region between UB and LB, by 97% of the measured spectrum. Thus, further development of these estimation techniques is required in order to enhance their performance in terms of the shape of the spectral recovery.

5. CONCLUSION
In this paper, we have described the methodology for measuring the spectral response functions for two cameras (Nikon D5100 and Sigma SD1 Merrill). We have undertaken this measurement at the National Physical Laboratory (NPL). We explain how the spectral response functions can be measured using the calibration system and a reference detector with a calibration that is traceable to SI units. We determined approximately a 2% uncertainty value for the recovered spectra per wavelength. The measured spectral response functions from NPL for both the Nikon and the Sigma were validated in an experiment, reporting plausible estimation errors.

Sensor sensitivity estimation techniques implemented from the literature are Tikhonov, Tikhonov based on derivatives, linear combination of basis functions, and quadratic programming. The results of their evaluation show that most techniques predict the camera responses very close to the real ones. Additionally, we used two error metrics (spectral error and Vora’s error) to compare the estimated sensors with the NPL ground truth measurements. Results show that 7 sets of sine basis and Tikhonov based on derivatives estimates the spectral sensitivity functions consistently well across all our data sets. Since most estimated sensor sensitivity functions differ from ground truth sensitivities in terms of recovered shape across the spectrum, we conclude that the estimation techniques do not provide recoveries as accurate as the spectral response functions measured at NPL. Furthermore, no algorithm returns estimates with measurement error. We hope that the ground truth data made available will aid the development of the estimation techniques as well as provide a basis for comparisons.

APPENDIX A

Fig. 7. Illustration of the red channel spectral response function recovered from NPL (solid line) and estimated Sigma sensor sensitivities recovered by 7 sine basis (dashed) using real data from the Macbeth ColorChecker.

Fig. 8. Figures of the linearity test for the Nikon camera. Points represent the camera signal in response to the input mean reflectance for 1/40, 1/25, and 1/15 s exposure timings.

Fig. 9. Uncertainty budget percentage for the Nikon’s red channel spectral response function.

ACKNOWLEDGMENTS
Special thanks to the National Physical Laboratory and colleagues Subrena Harris (NPL), Roberto Montagnia, and Hans Rivertz for their support. As our final goal, the measured spectral sensitivity functions from NPL and data sets from the evaluation experiment for both the Nikon and the Sigma are made available for the community to download at http://spectralestimation.wordpress.com/data/.

REFERENCES