Low-cost visible light spectral imaging

Daniel W. Dichter

Independent Researcher, Cambridge, Massachusetts, USA
Email: daniel.w.dichter@gmail.com

Spectral imaging is an emerging technology for measuring spectral power distributions (SPDs) of electromagnetic radiation over a two-dimensional spatial domain. Within the visible light wavelength domain, spectral imaging measures light and colour with greater accuracy than digital photography. However, high cost limits its accessibility. Accordingly, a low-cost method was developed using commercially-available hardware – primarily a DSLR camera and a set of narrow bandpass filters. The quantity of filters was minimised to a total of seven, set by the dimensionality of SPDs, the spectral sensitivities of eyes and cameras, and commercial availability. Camera spectral sensitivity was measured using this same filter set, a colour chart, a spectrophotometer, and noon daylight modelled as CIE D65. The RAW photo format was used to access unprocessed sensor data. Independent SPD measurements from each colour channel were fused as a sensitivity-weighted average for efficient and continuous interpolation between colour channels with a high signal-to-noise ratio. Images were reconstructed from SPDs with standard observer functions. The method was demonstrated with a Canon 650D DSLR camera, a set of Thorlabs one-inch narrow bandpass filters, an X-Rite ColorChecker chart, and a Spectro 1 spectrophotometer. Accuracy was validated by quantitative comparison against ground truth SPD measurements and qualitative assessment of reconstructed images. The total filter cost was $715, plus $405 to measure camera sensitivity.

Received 12 August 2021; revised 06 September 2021; accepted 07 September 2021
Published online: 06 October 2021

Introduction

Background

Colour perception results from the interaction of an illuminant, optionally a reflective surface, and an observer. Humans are trichromats, having three-dimensional colour perception, over the visible light wavelength spectrum of 400 to 700 nm. Human colour perception is significantly rank-deficient even within its wavelength limits, as shown in the subsections Dimensionality of reflectance spectra and Curve reconstruction.
Whereas digital photography produces a three-dimensional RGB (red, green, blue) colour image, spectral imaging produces a k-dimensional hyperspectral datacube (HSDC), typically with $10 \leq k \leq 60$; wavelength channels replace colour channels. HSDCs are informationally complete and observer independent, providing a more accurate representation of colour, and enabling greater optionality in subsequent image processing.

Spectral scanning is the simplest of several methods for generating HSDCs, wherein a 1-channel intensity image is captured at a series of regularly-spaced, narrow, non-overlapping wavelength bands. These images are subsequently “stacked” to form the final HSDC. Though spatially high-resolution relative to other methods, spectral scanning experiences “spectral smearing” if the subjects or sensor move during the image capture process, due to the characteristically large temporal distribution of the images. By comparison, so-called “single-shot” methods use various diffraction techniques to capture the entire HSDC simultaneously, but at lower spatial resolution.

A survey of commercially-available hyperspectral cameras\(^1\) showed a price range of $20,000-25,000, whereas consumer cameras such as DSLRs and MILCs cost 1-10% of this amount.

**Related work**

Consumer cameras are not capable of spectral imaging as-is, but various researchers have demonstrated this capability nonetheless with various hardware and software methods. Such methods can generally be classified as multi-shot spectral scanning or single-shot diffraction grating.

Cosentino \([1]\) used 12 bandpass filters and a modified camera to produce HSDCs, comparing results favourably to a commercial hyperspectral camera. Berns et al. \([2]\) used a large-format camera and an optimised set of filters, performing singular value decomposition on a dataset of 2,500 reflectance spectra to reduce the quantity of filters to a total of six.

Baek et al. \([3]\) developed a single-shot method using a custom prism objective. Combined spatial and wavelength information was separated by detection of “spectral cues” present “only around object edges”. Habel et al. \([4]\) similarly developed a single-shot diffraction grating method, with reconstructed images limited to 120 $\times$ 120 pixels.

Oh et al. \([5]\) developed a novel non-diffraction single-shot method using three different synchronised digital cameras, exploiting small differences in their sensitivities. An image registration process was used to align images between cameras using planarity. Principal component analysis was performed on a database of 1,257 Munsell reflectance spectra to create a low-dimension vector space for describing SPDs as linear combinations.

The method described in this paper is believed to be novel in its low cost, high spatial resolution, use of unmodified commercially-available hardware, and independence from training data.

**Methods**

**Dimensionality of reflectance spectra**

Parkkinen et al. \([6]\) measured the reflectance spectra of 1,257 standard Munsell colour swatches over 400 to 700 nm. Noting that “the components of a color spectrum are highly correlated”, the authors performed principal component analysis on these spectra, producing a set of reflectance eigenvectors.

---

\(^1\) PixeTeq SpectroCam-VIS, BaySpec GoldenEye, Resonon Pika L, Specim IQ.
It was found that these spectra could be reconstructed as linear combinations of eight or fewer eigenvectors. This result indicates that reflectance spectra are up to eight-dimensional over the visible light spectrum, with a characteristic wavelength resolution of approximately \((700 \text{ nm} - 400 \text{ nm}) / 8 = 37.5 \text{ nm}\).

**Curve reconstruction**

Sampling an SPD by spectral scanning with narrow bandpass filters produces a sparse sample; the value of the SPD is measured only at certain regular intervals. Between these, the value of the SPD is not measured, but can be approximated due to the limited dimensionality of SPDs. This process was modelled as a curve reconstruction problem, i.e. given a sparse sampling of an unknown curve, reconstruct the curve by means of an appropriate interpolation scheme such that the reconstructed curve matches a theoretical measured curve. This concept is shown in Figure 1.

![Curve reconstruction](https://www.aic-color.org/)

**Figure 1**: Curve reconstruction for SPD of ColorChecker swatch #2 under D65 illuminant for varying sample quantity. Samples are denoted with vertical lines, with cubic interpolation between samples.

To determine the optimal filter quantity and interpolation scheme, two reflectance datasets were considered: the 8 eigenvectors developed by Parkkinen et al. [6] with 5 nm resolution, and the 24 reflectance spectra measured from an X-Rite ColorChecker Classic [7].

ColorChecker reflectance spectra were measured using a Spectro1 spectrophotometer with a domain of 400 to 700 nm and a resolution of 10 nm. Threescans were performed and averaged for each swatch, with near-perfect repeatability across trials. These spectra were validated by calculating the CIEDE2000 colour difference (denoted \(\Delta E_{00}\)) against the manufacturer’s published Lab colour values under CIE D50 illuminant with a 2° observer [7]. The error for all 24 swatches was \(1.39 \pm 1.06 \Delta E_{00}\), showing good agreement. Research has shown significant variance in ColorChecker reflectance spectra,
with standard deviations of 0.1-9.1 percentage points over 400-700 nm [8], though the manufacturer
does not publish spectra or tolerances.

Per subsections Dimensionality of reflectance spectra and Filter set design, both 30 and 40 nm
resolution were considered. The reconstruction domain was taken as 420-660 nm per subsection Filter
set design. The filters were assumed to be of sufficiently narrow bandwidth that they measure at their
exact centre wavelength (CWL). Linear, spline, and cubic interpolation were considered. The quality of
the reconstruction was calculated as the residual sum of squares (RSS) for all curves, multiplied by
the wavelength resolution, divided by the quantity of samples in the dataset. This calculation was performed
for both reflectance spectra and simulated SPDs, with the latter modelled as the element-wise product
of reflectance and the D65 illuminant scaled to the range of 0-1. This normalised error metric allowed
direct comparison between datasets.

Several observations were made. 30 nm resolution outperforms 40 nm resolution in all cases. With
40 nm resolution, cubic interpolation is optimal. With 30 nm resolution, cubic interpolation is at least
near-optimal. Optimised error at 40 nm is 1-3× larger than at 30 nm, but both are small in an absolute
sense. Reflectance (not illuminant) dominates the quality of the curve reconstruction. Thus, 40 nm
resolution and cubic interpolation were chosen for curve reconstruction.

Filter set design

Manufacturers of narrow bandpass filters include Thorlabs, Edmund Optics, and MidOpt. Within the
visible light spectrum, CWL is generally discretised as whole-number multiples of 10 nm, i.e. 400, 410,
420, ..., 700 nm.

The centroid of the visible light spectrum may be defined at the wavelength corresponding to 50% on
the cumulative density function (CDF) of the sum of the CIE 2° tristimulus observer functions \([x, y, z]\)
[9]. Rounding to the nearest 10 nm per commercial availability, this centroid is 540 nm.

It can be shown that the wavelength range of 420 to 660 nm encompasses 97% of the area under the
CIE 2° and 10° tristimulus observer functions [9], thus:

\[
\int_{420\text{ nm}}^{660\text{ nm}} (x + y + z) \, d\lambda \approx \int_{-\infty}^{\infty} (x + y + z) \, d\lambda 
\]

This range is also evenly divisible into 40 nm increments, and intersects the centroid of 540 nm.

The final specification is the full-width half max (FWHM), or bandwidth. As discussed in subsection Curve
reconstruction, reducing FWHM increases the wavelength accuracy of measurements to the
CWL. The reduction in overall transmission associated with a low FWHM was compensated by
increasing the exposure, as discussed in subsection Photographic aspects.

The set of filters chosen are shown in Figures 2 and 3, and have the following properties:

<table>
<thead>
<tr>
<th>Quantity</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacture</td>
<td>Thorlabs</td>
</tr>
<tr>
<td>CWL</td>
<td>420 to 660 nm</td>
</tr>
<tr>
<td>CWL spacing</td>
<td>40 nm</td>
</tr>
<tr>
<td>FWHM</td>
<td>10 nm</td>
</tr>
<tr>
<td>Part numbers</td>
<td>FB420-10, FB460-10, etc.</td>
</tr>
<tr>
<td>Diameter</td>
<td>One inch</td>
</tr>
<tr>
<td>Cost (total)</td>
<td>$686 (+$29 case)</td>
</tr>
</tbody>
</table>
Camera response linearity

RAW values reported from cameras may be thought of as photonic measurements. A simple under-exposure experiment found that pure black (i.e. zero photons) corresponds to RAW = 2,048, commonly referred to as the black level. In principle, over-exposure and saturation occur at the maximum value permitted by the bit depth: \(2^{14} = 16,384\). In practice, saturation was observed at values ranging from 12,000-16,384. Between these limits, response linearity was verified by photographing a ColorChecker chart under noon daylight while independently varying shutter duration and ISO. The trichromate mode of the #19 white swatch was calculated for each photo as the measurement of interest. Mode was chosen for its robustness against hot and dead pixels.

These measurements are shown in Figure 4. For ideal linearity, the RAW value is proportional to the product of shutter duration and ISO, all else equal. Accordingly, an idealised RAW value was calculated for each reported RAW value, by linearly scaling the product of shutter duration and ISO to the RAW value range of 2,048-12,000. This relationship was expressed as a set of transfer functions, shown in
Figure 5. The shape of these transfer functions indicates that the sensor has a linear response up to saturation. Similarly, varying ISO with a constant shutter duration exhibited near-perfect linearity (f/22, ISO 100-12,800, 1/2,000 sec).

**Camera spectral sensitivity**

Camera spectral sensitivity was calculated by photographing a ColorChecker chart through each filter under cloudless noon daylight modelled as D65. Noon daylight was chosen for its roughly neutral SPD, availability of standard data, and accessibility. Images were captured on 2021-02-21 at 12:22 pm in an open field in Cambridge, Massachusetts.

For an ideal linear response as shown in subsection Camera response linearity, sensitivity may be described generally as:

$$S = \frac{V_M}{V_A}$$ (2)

$V_M$ was calculated as the mode of the RAW values inside a square inset slightly from the swatch perimeter, minus the black level. $V_A$ was calculated as:

$$V_A = \sum \lambda I(\lambda) \odot R(\lambda) \odot T(\lambda)$$ (3)

with $\lambda$ = wavelength, $I$ = scene illuminant, $R$ = swatch reflectance, $T$ = filter transmission and $\odot$ = element-wise multiplication.

**Figure 6 (left):** Non-dimensional trichromatist spectral sensitivity of Canon 650D camera from ColorChecker under D65 illuminant. Faint coloured lines correspond to individual swatches; solid black lines correspond to weighted averages for all swatches.

**Figure 7 (right):** Comparison of sensitivities for various Canon cameras [11]. Curves were normalised so that the maximum sensitivity for each camera is equal to unity.

Sensitivity for all colour channels and wavelengths was calculated from each swatch, as shown in Figure 6. These per-swatch sensitivity curves were fused as a weighted average using $V_M$ to weight calculated sensitivities in proportion to their signal-to-noise ratio.
The result is a set of three sensitivity curves, \([S_R(\lambda), S_G(\lambda), S_B(\lambda)]\). As shown in Figure 7, this result is consistent with values from literature for similar Canon cameras.

**Photographic aspects**

A prime (i.e. non-zoom) lens, tripod, quick release plate, and remote shutter were used to minimise movement of the camera while capturing the photo stack. Because filter cost is proportional to filter area, the choice of lens is the primary means of cost reduction. The measurement of interest is the diameter of the objective at the outer end of the lens, as it sets the minimum filter diameter. For consumer Canon EF lenses, this ranges from approximately 20 to 60 mm. The theoretical filter cost at these extremes differs by nearly an order of magnitude; \(60^2 / 20^2 = 9\). The lens selected has an outer objective diameter of 20 mm, a focal length of 40 mm, and an aperture of f/2.8. It is shown in Figure 2.

The filters discussed in subsection *Filter set design* are unthreaded, and used by resting them against the camera lens by hand, transferring minimal force and maintaining alignment of the photo stack. The glass optics of both the lens and filters are recessed from their cases, preventing abrasion. The filters were stored in a case in order of ascending wavelength, and cycled through in sequence manually. Depending on the nature and luminance of the scene, typical camera settings were \(f/2.8-10\), ISO 100-400, and \(1/4 - 1/500\) second. Bright exposures were required to compensate for the low transmission of the filters. These settings were adjusted to “expose to the right”, i.e. fully utilise the available set of values without saturating the upper limit, thereby maximising the signal-to-noise ratio in the measured values.

The open-source dcraw utility was used to extract RAW values from the .CR2 file format [12]. The modifier string -D-4-j-t0 was used to specify that the extracted RAW values were unprocessed sensor measurements.

<table>
<thead>
<tr>
<th>-D</th>
<th>No value scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>-4</td>
<td>Linear 16-bit</td>
</tr>
<tr>
<td>-j</td>
<td>No stretching or rotating pixels</td>
</tr>
<tr>
<td>-t 0</td>
<td>No image rotation</td>
</tr>
</tbody>
</table>

The RAW sensor measurements were then demosaiced into R, G, B colour channels with a Bayer filter pattern of rggb.

**SPDs from RAW photos**

By rearranging the general expression of sensitivity in subsection *Camera spectral sensitivity*:

\[
V_A = \frac{V_M}{S}
\]  

(4)

For a linear sensor response, \(V_M \propto P\), with \(P\) denoting the RAW photo value minus the black level. By inspection, \(V_A \propto \text{SPD}\). Using proportionality and non-dimensionality, this is rearranged and substituted as:

\[
\text{SPD} = V_A = \frac{P}{ST}
\]

(5)
The denominator $ST$ is most accurately and generally expressed as a dot product over the wavelength domain rather than a scalar product at the CWL, i.e. $S(\lambda) \cdot T(\lambda)$ rather than $S(\lambda=CWL) \cdot T(\lambda=CWL)$.

At an arbitrary pixel and CWL, each colour channel produces an independent SPD measurement, collectively given as:

$$\text{SPD} = \begin{bmatrix} P_R & P_G & P_B \\ S_R T & S_G T & S_B T \end{bmatrix}$$

(6)

These measurements are theoretically equal, but in practice differ due to various sources of error. They are fused as a sensitivity-weighted average:

$$\overline{\text{SPD}} = \frac{1}{S_R + S_G + S_B} \left( \frac{P_R S_R}{S_R T} + \frac{P_G S_G}{S_G T} + \frac{P_B S_B}{S_B T} \right)$$

(7)

Calculating $\overline{\text{SPD}}$ as a sensitivity-weighted average continuously interpolates between colour channels as a function of wavelength. Since $S_R + S_G + S_B > 0$ for all $\lambda$, divide-by-zero is precluded. This formulation was generalised spatially by using matrices in place of scalars for $P$, neglecting vignetting, dark-frame effects, and other spatially-related sources of error. Subsection Lens glare includes a discussion on lens glare specifically.

The HSDC was first calculated in a sparse fashion, only at the CWLs of the filters. The full (i.e. non-sparse) HSDC was then calculated by cubic interpolation in the wavelength domain between the filter CWLs per subsection Curve reconstruction.

Results

Validation

The method was validated by comparing ColorChecker SPDs and $\Delta E_{00}$ between camera vs. spectrophotometer, with the latter taken as ground truth and modeled as $R(\lambda) \odot I(\lambda)$. Because the two sets of curves, $\text{SPD}_{\text{camera}}$ and $\text{SPD}_{\text{spectrophotometer}}$, are non-dimensional, their ranges were aligned with a single scalar gain for comparative purposes, such that the sum of the residuals was zero. This scalar gain $\alpha$ was found by the bisection method such that:

$$\sum_{\lambda} \alpha \text{SPD}_{\text{camera}} - \text{SPD}_{\text{spectrophotometer}} = 0$$

(8)

Standard equations were used to derive XYZ, Lab, and RGB colours from SPDs, with D65 as both the scaling factor and white point [13].

Results are shown in Figures 8-10. The colour error for all swatches is $2.44 \pm 1.65 \Delta E_{00}$, with more than 90% of swatches within $1.75 \pm 0.96 \Delta E_{00}$. The primary outlier is #19 white, which shows a good match in terms of normalised distribution, but a poor match in magnitude and thus luminance. This is best explained as an artefact of glare or other lighting non-uniformities, amplified by the high reflectance of the colour white. The SPDs of the matte black background showed a spatial luminance variance that matched the SPD magnitude error.
In the more general case of an unknown illuminant, colours are dimensionali
d by linear scaling in the luminance (or intensity) dimension to the limits of the colour space. A common strategy is to allow
a small amount of saturation at the limits of the colour space, thereby avoiding hot and dead pixels. Reconstructed images in this paper are scaled such that the bottom and top 1% of pixels are saturated, with black and white defined as 5% and 95% (i.e. RGB [13 13 13], [242 242 242]) respectively.

**Figure 8 (left):** Distribution of colour errors shown in Figure 10.

**Figure 9 (right):** Reconstructed image of ColorChecker chart; f/2.8, ISO 100, 1/400 sec.

**Figure 10:** Comparison of ColorChecker SPDs and colours as measured by spectrophotometer (solid, left) and camera (dashed, right). A single scalar gain is applied to align the two sets of curves. Plots are scaled 400 to 700 nm along x-axis, and non-dimensional along y-axis.
**Lens glare**

Lens glare is a phenomenon characteristic of optical imaging systems resulting from scattering of light within the lens [14]. Within the context of spectral imaging, glare is an undesirable source of spatial-spectral error, “especially for darker pixels and pixels close to light sources or bright patches”. Signoroni et al. demonstrated that hyperspectral cameras are susceptible to significant glare-induced error by comparing measurements against a spectrophotometer, taking the latter as ground truth for its absence of spatial-spectral effects. The authors showed that glare may be quantified by proxy of its attenuation of measured contrast, particularly for greyscale colours spanning black to white.

An experiment was performed to induce, isolate, and quantify glare effects for the proposed system. The ColorChecker chart was placed indoors in natural light. A 40 Watt, 340 lumen incandescent bulb was placed out of frame, with its axis directed at the camera’s lens. The swatches, lens, and bulb were roughly coplanar. This setup is shown in Figure 11. SPDs were measured per subsection SPDs from RAW photos with the bulb on and off under the same camera settings of f/2.8, ISO 400, 1/2 sec.

The brightness, or perceived luminance, was calculated for each greyscale swatch as \( \sum \text{SPD}(\lambda) \odot \bar{y}(\lambda) \), with \( \bar{y}(\lambda) \) per the CIE 2\(^{\circ}\) XYZ observer functions. These values were compared against ground-truth brightness, calculated as \( \sum \text{R}(\lambda) \odot I(\lambda) \odot \bar{y}(\lambda) \) per spectrophotometer measurement discussed in subsection Curve reconstruction, with \( I(\lambda) \) taken as CIE D65. Brightness values were normalised such that white had a brightness of 1.0 for all three datasets. Results are shown in Figure 12.

Contrast was calculated as the measured brightness difference between black and white. With the bulb off, contrast was 97.7% of ground truth, with the error attributed to glare from pixels and lighting non-uniformity. With the bulb on, contrast decreased to 82.8% of ground truth, with the further attenuation attributed to induced glare. These values are consistent with findings of Signoroni et al. [14]

As discussed in subsection SPDs from RAW photos, the assumption of spatial-spectral independence of pixels significantly simplifies SPD computation, and is consistent with the methodology of commercially-available hyperspectral cameras [14]. Detecting and normalising generalised glare effects is a complex image processing problem that is left beyond the scope of this paper. In practice, glare can typically be limited to an acceptably small degree by following standard best practices of photography.
Reconstructed images

Several scenes are shown in Figure 13 with subjects and illuminants that represent typical colour perception. Shown for each scene is the reconstructed image (left column) and a sparse sampling of the HSDC (right column). Each HSDC was sampled at an evenly-distributed 10 × 7 square mesh with 70 nodes total, showing a characteristic set of SPDs that are coloured according the reconstructed image.

![Reconstructed images](Image)

Figure 13: Reconstructed RGB images and characteristic sets of SPDs for several scenes. Plots are scaled 420 to 660 nm along x-axis, and non-dimensional along y-axis. Filter CWLs are denoted by vertical lines.

The still life (first row) exhibits blue-red contrast apparent in the bimodal distribution of the SPDs. A relatively high f-stop was needed to keep the scene in focus, which required increasing both ISO and
shutter duration. The lower signal-to-noise ratio inherent with higher ISO can be seen in the faint high-frequency noise pattern. The camera settings were f/10, ISO 400, 1/4 sec.

The landscape (second row) exhibits both colour and luminance contrast between sky and ground. Minor “spectral smearing” can be seen around the edges of the foliage due to natural movement. The sky exhibits a characteristic D65-like distribution. Higher overall luminance enabled lower ISO and shutter duration, reducing noise. The camera settings were f/2.8, ISO 200, 1/200 sec.

The industrial scene (third row) exhibits correlation between SPDs that is consistent with the yellow-red light characteristic of sunsets. The camera settings were f/4, ISO 200, 1/60 sec.

Discussion

This paper demonstrates a method for spatially high-resolution visible light spectral imaging at less than 10% the cost of a commercial hyperspectral camera, using commercially-available hardware. Results are validated quantitatively and qualitatively against ground truth. Key aspects of novelty include characterisation of SPD dimensionality, curve reconstruction from sparse samples, and fusion of trichromate measurements. This method significantly improves access to the field of spectral imaging, and enables further research into the field of spectral image processing.

Acknowledgments

The author gratefully recognises his friends, family, and colleagues for their generous and thoughtful support in shaping this work – in particular: Eric Kirchner, Trey Henderson, Bronek Dichter, Michał Dichter, and Evan Molinari.

References

7. X-Rite (2019), Colorimetric values for ColorChecker family of targets. [https://xritephoto.com—last accessed 11 April 2019]
8. BabelColor (2021), The ColorChecker Pages. [https://www.babelcolor.com—last accessed 1 August 2021]

