Additivity Constrained Linearisation of Camera Calibration Data

Casper F. Andersen, Ivar Farup, Jon Y. Hardeberg

***Abstract*—When characterising a digital camera spectrally or colourimetrically, the camera response to a generally diffusely reflecting colour chart is often employed. The recorded responses to the light incident from each colour patch are typically not linearly related to the power of the irradiance on the chart, and the irradiance varies with position on the chart. This necessitates a linearisation of the responses. We present a new single image colour chart-based estimation method of responses, that are *linearly related* to camera response values known as *ground truth*. The method estimates the spatial geometry of the irradiance incident on the chart attenuated by lens vignetting and compensates individually for volumetric and per colour channel non-linearities, including compensation for physical scene and camera properties in a pipeline of successive signal transformations between the estimated linear and the given recorded responses. The estimation is controlled by introducing a novel *Additivity Principle* of linear responses, which is derived from the spectral reflectances of the coloured surfaces on the colour chart, observing that linear relations of the spectral reflectances are equal to the relations of the corresponding linear responses. Crucially, the additivity principle is not subject to metamerism. The method is fundamentally solely reliant on a one-shot set of one triplet of response values sampled from each patch of a colour chart with known spectral reflectances, where rendition level, gray scale, illuminant, camera sensor curves, irradiance geometry, vignetting, moderate specular reflection, colour space, colour correction, gamut correction and noise level are unknown.**

***Index Terms*—Additivity principle, colour chart, camera calibration, response linearisation, rendered responses, tone curves, irradiance geometry, colour correction, specular reflection, noise ...**

I. Introduction

KNOWLEDGE of linear colour channel values in digital images pertaining to *linear* recorded sensor responses to spectral stimuli in a physical scene is a prerequisite when estimation of spectral components of or colourimetric operations on said image are necessary [1]. By *linear colour channel values* we refer to ideal (*ground truth*) responses that result from integrating the per wavelength multiplied spectral power distribution of the illuminant, the spectral reflectance of a surface and the per channel spectral sensitivity of the camera. This simulates a spatially uniform illumination irradiating spatially arranged reflective surfaces recorded by a sensor, which is defined exclusively by a set of spectral filters each pertaining to a colour channel. They are thus exposure independently related to the according colourimetric values. The *ground truth* responses and their according colourimetric responses, i.e. XYZ form the data sets needed in the colour characterisation of the camera. In the important cases, when calibrating [2], colour image re-targeting [3] or colour characterizing [4], [5], [6], [7] a digital recording device (e.g. a digital camera) based on its responses to a set of spectrally calibrated surfaces on, say a colour chart, such values are not necessarily available in practice or given tentatively as raw responses in raw capable cameras. The raw responses are often used in in-camera colour characterisation as ground truth.

C. F. Andersen, I. Farup, and J. Y. Hardeberg are with the Norwegian Colour and Visual Computing Laboratory, Department of Computer Science, Norwegian University of Science and Technology (NTNU), Gjøvik, Norway

The recorded raw responses to the colour patches on the chart need correction for elements in the optoelectronic calibration function (OECF) [8] including dark noise, shot noise and read-out noise, sensor non-linearities, lens glare and vignetting [9], [10], unknown effects of low-level camera signal processing e.g. demosaicing etc. The according linearisation corrections may be part of an in-camera signal processing [10], in which vignetting and noise might be compensated in the raw-responses. For our purposes, this correction is not required. Furthermore, to truly become proportional to the incident irradiant power on the chart, its spatial geometry, which depends on the capture geometry and not the camera itself, must also be known and compensated for. In cases of spatial non-uniformity of the illumination and specular reflections, linearised raw data do not suffice as *ground truth* data. We note that vignetting like the spatial geometry of the irradiant power of light attenuates (by shading) the camera responses spatially [9]. They are thus controlling the spatial irradiance on the sensor together. It is thus considered a part of the irradiance throughout the method model. However, the effect of vignetting may be considered small around the centre of the image [11].

In cases where the images of calibration charts contain rendered colour channel values, which typically are highly non-linear, the linearisation is complicated further as the data results from a pipeline of transformations of raw data, including white balancing, tone mapping, colour correction, gamut operations [12] [13], sharpening and quantisation.

In this paper, we introduce a novel approach to linearisation of camera calibration data responses subsampled from a single shot image of calibrated colour charts. Although not restricted to, only one triplet of responses per patch is necessary. We do not present an exhaustive comparison with other methods, as the aim here is to present the additivity principle and its implementation along with a regularisation method of choice, which in itself is novel in conjunction with camera linearisation.

It is assumed that the transformation from *ground truth* responses, to camera responses at any stage of rendition, consists of a successive series of smooth corrections that are

at least C^1, i.e., continuously differentiable (excluding gamut wall cut-offs and patches with a conspicuously high level of specular reflectance). The method estimates these *individual corrections*, which, when subsequently applied inversely to the camera responses, obtains estimated linear responses that are *linearly related to ground truth* responses (i.e. related within a matrix multiplication). Crucially, *the estimation is thus not subject to metamerism*. The estimation incorporates a regularisation according to the aforementioned smoothness, and is not constrained to monotonicity. The method will estimate gamut-clipped responses and responses to patches on charts with an offset consisting of noise and a (small and evenly distributed) specular reflection added to the ground truth responses.

Section II provides background on existing methods for camera characterisation, calibration, and linearisation. As there is no other method, to our knowledge, that truly encompasses the full inversion of the transformation pipeline from rendered or raw data to linearly related *ground truth* calibration data, given the limitations of available subsampled colour patch-data that our method presumes, we will briefly present the highlights of the various methods and discuss their potential.

The methods discussed below have goals that are not necessarily pointed towards the linearisation of colour chart responses exclusively as ours, but may encompass the estimation of raw which, depending on the camera in question, may or may not be fully linearised [10] in advance or pointed at a linearisation of the camera responses to the full image plane on the CCD, i.e. a radiometric calibration of said camera. Methods that include inversion of rendered data typically presume a parameterised rendering model of reduced degree of freedom [14] to regularise the inversion. In these methods, there is no mention of estimation of gamut edge cut-offs (except [12]) as our method includes. In table I, an overview of a selection of state-of-the-art methods is presented.

In Section III, we describe the fundamentals of colour formation and rendition pertaining to digital camera recording, with a specific view to camera calibration based on colour charts. We further develop a model of camera response from linear *ground truth* over raw to rendered colour channel values, that will subsequently be the foundation for the linearisation method presented. A graphic overview of said pipeline is given in Figure 1.

In Section IV our linearisation method is defined. To enable implementation, the presentation is comprehensive.

In Section V we present the measured and (semi-)synthetic sets of data at our disposal that we use in the validation of the method, involving the digital camera, colour and gray chart, spectroradiometer and capture settings and environment.

Section VI outlines our typification of the conducted tests and according origin of the response data, by referring to specific *rgb-types* in accordance with measured, variations of semisynthetic and ground truth data.

In Section VII the results of the application of the method to the typified camera data are presented and discussed. We will also compare our results with the classical method of normalisation to a gray chart, when the image data permits. We will provide a table of results pertaining to the real camera samples from each individual recorded image along with EXIF-data and according raw responses to the white patch on the colour chart. These results may be compared to the reported results from literature, which are (roughly) outlined in table I.

Section VIII contains a conclusion.

II. Background on Camera Linearisation

State-of-the-art methods of calibration chart response linearisation target different elements of the linearisation process depending on whether the data is raw or rendered. In the pipeline from rendered to *ground truth*, we choose to define two distinct categories. First, the raw to rendered process consisting of colour correction, gamut mapping and per channel tone curve correction, which contains gamma and contrast operations upon which a quantising JPEG compression is applied. This process can contain proprietary transformation [15]. Second, the raw to *ground truth* connection which compensates for the fact that the raw data may not be perfectly linearly related to the incident power of light due to lens transmission (e.g. vignetting), noise and sensor non-linearities, if the camera is not linearised otherwise before. The raw data extracted from an image of spatially distributed colour patches on a chart, may furthermore include spatial variation of irradiance on the chart surface, glare, and spatially dependent reflective properties like specular reflections originating from incongruent camera recording and spectral measuring geometry of the patch surfaces.

We consider methods of camera linearisation with a view to the specific task of colour chart linearisation. The methods are based on a variety of input data ranging from full image analysis at pixel level or subsampled and averaged areas of interest, employment of a linear proxy as substitute for raw, imagery with or without an embedded colour chart, precalculated linearisation parameters, such as image or device database support and one or multiple captures. Common for these methods is the estimation of raw from rendered through linearisation. With a few exceptions, *this constitutes the main difference to our method, as it further linearises the step from raw to ground truth, but not raw itself and furthermore, relies solely on per patch subsampled camera values*.

A. Raw to Rendered

1) Based on a Single Image: In [16] pairs of raw and rendered (srgb/JPEG) colour values are employed to estimate a linear colour transformation of raw and a subsequent channel tone mapping. This approach is not dependent on a colour chart, just as the method in [17], where a colour edge analysis from selected parts of a single image reveals the radiometric pipeline consisting of tone mappings and a linear colour transformation matrix.

Using an image of a colour chart and assuming monotonicity of the relationship between ranked camera responses and their equivalent raw responses, the methods in [18] and [19] can estimate the irradiance intensity across the surface of the colour chart which attenuates the raw responses and the tone mapping of the intensity corrected and linearly colour

transformed rendered data. The methods require pairs of raw and rendered data, but raw can be approximated by a choice of a canonical (or proxy) colour space [19] or in-camera linearised raw data from another camera recording of the same scene [14].

2) Based on Multiple Exposures: Methods relying on analysing a multiplicity of exposures of a scene with a view to constructing a high dynamic range image through radiometric calibration are presented in [11] and [20]. These methods are not reliant on a colour chart. In [8] an imaging model for radiometric calibration enables conversion from rendered srgb images to the corresponding raw values. In [21] the *Opto-electric Conversion Function* (OECF) curves in each channel are estimated by mapping camera response as a function of incident irradiant power to linearize the responses.

3) Based on Gray scale: Instead of gathering multiple images of the same scene, tone mapping can be estimated by analysis of a series of known in-image neutral patches ranging from high to low reflectance [22].

4) Based on Mixed Methods: In [23] an optimisation framework based on a rank minimisation technique is presented by which full images, edge analysis and multiple exposures can be used to estimate the radiometric non-linearity in the colour channels of the camera pertaining to the raw data.

An alternative to using a gray-scale when linearising channel responses is to gather multiple in-register images of the same scene at varying and known exposures [11], [20].

Based on exploring information from a prior statistical analysis of three-dimensional reflective properties in a multitude of images in photo collections pertaining to a camera sensor, radiometric calibration methods are proposed in [14], [24]. This analysis is based on a large body of prior image information and a full image analysis and aims at raw recovery.

A comprehensive variation of methods employed by digital camera manufacturers of handling noise, exposure, flat-field compensation, sensor non-linearities, rendering techniques etc. are assembled and analysed in [10], [25]. Here, it is proposed to gather a structured set of information from each manufacturer in order to standardize spectral and radiometric calibration of cameras. When raw data is not available, a canonical (proxy) colour space is introduced in [18], [26] and [27] that serves as basis for illuminant and sensor independent colour space as an approximate set of *ground truth* raw responses, that enables estimation of the radiometric operations imposed to reach a set of rendered data. In [28] an alternating least squares method estimates the irradiance geometry over a known colour chart by pairing the rendered with the proxy data. It assumes that tone curves are linear. Including unknown tone-curves, this method is close to [29] where an image corrective irradiance geometry and a von Kries type illuminant colour correction on the image colour values are found in an iterative linear least squares optimisation.

B. Raw or Rendered to Ground Truth

In [1], a radiation intensity-based linearisation technique is presented, where ground truth is estimated. The method assumes a spatially even irradiance over the centre of the capture area in which 14 neutral patches are located on top of a regular colour chart, similar to the method presented in [22].

In [26] a method to estimate the *ground truth* responses to a specific in-image colour chart is based on a database of images of the colour chart, captured under a variety of illuminants.

To compensate for the very likely non-uniform irradiance over a chart surface, without prior knowledge of the magnitude of the variation of the irradiation in each response sample, a method of colour characterisation that takes advantage of the intensity independent chromaticity values is presented in [30]. Specular reflection is not considered in these methods.

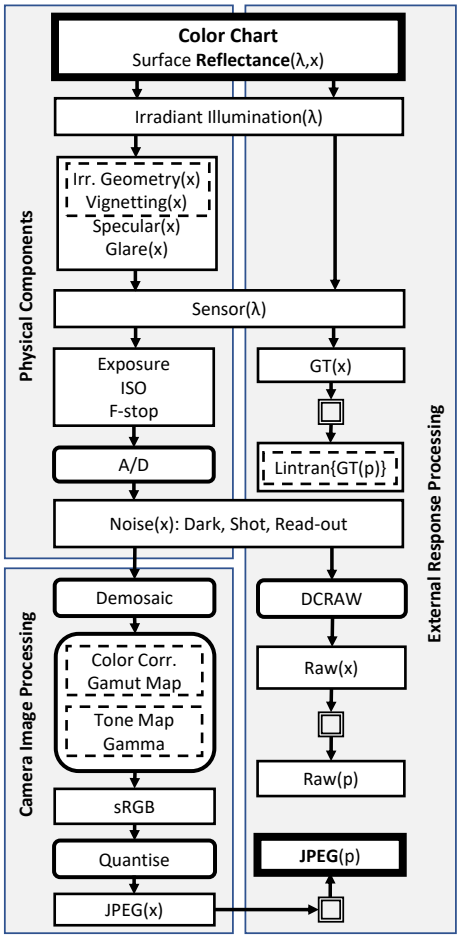

Fig. 1. The radiometric pipeline and connection with ground truth, within which our method lumps in per channel 1D corrections: specular and noise offsets, glare, tone mapping and gamma, in a 2D spatial correction: illuminant irradiance geometry (non-uniformity) and vignetting and in an all channels 3D correction: gamut mapping and nonlinear part of colour correction. The wavelength is λ, pixel position (x) and sampling in patch (p). Sampling is depicted by a double square box and registers averaged channel values from a small area on (p) with an according image position at the centre of the sampling area. Rectangular boxes are data, rounded boxes are processing, and thick rectangular boxes are input data to the method. Note, Raw(p) can substitute JPEG(p) if the estimation is based on raw instead of JPEG patch sampled image data. The spectral functions form the linear *ground truth* camera responses in Equation 1. The dashed boxes signify elements to estimate in (p) and surround each of the three lumped corrections and Lintran{...} which signifies a linear transformation of {...}. GT(p) signifies *ground truth* method evaluation responses in (p), derived from sampling of GT(x).

C. $\mathcal{X}$ to Ground Truth

In the situation we address here, there is no requirement for access to the raw or proxy camera recording and no additional in-register image of a gray chart. It is not a priori known whether the data is raw or rendered, or somewhere between. There is neither access to auxiliary camera-specific information nor a full image of the chart. The scene environment surrounding the chart, i.e., illuminant, specular reflectance and viewing geometry etc., is likewise not known. We will briefly refer to the unknown stage to which the data has been transformed in the rendition process as $\mathcal{X}$.

The spectral reflectances and their according geometrical positions on the chart are assumed to be known. Furthermore, there is only access to one shot of the target and only one set of RGB values for each patch gathered from a small area around the position where the known spectral reflectances originate. Within practical limits, the spectral measuring geometry of the colour chart need not be identical to the camera recording geometry. It is assumed that the prevailing illumination across the entire chart surface consists of a single spectral composition, and that the spatial irradiation geometry does not contain abrupt variations in intensity. It is further assumed, that any specular component in the reflection from the patches is distributed over the chart patches without abrupt spatial variation (for example a few scattered patches with high specular reflection, as on the *X-Rite Colour Checker DC*). Should such patches exist, they are omitted from the analysis. The estimation of their linear response can be conducted afterwards using the results pertaining to the chosen patches. For this data setting, no method is found in the literature.

III. Formulation of Colour Formation

Here we introduce the mathematical description and the nomenclature for digital camera colour formation. The terminology will be used for describing the proposed method.

A. Colour Formation

Following the model of colour formation, the *ground truth* sensor responses from each colour channel $\underline{\rho}$ for a given set of K sensors ($K = 3$ in contemporary camera systems) with spectral sensitivities $\underline{\sigma}(\lambda) = \left[\sigma^1(\lambda), \sigma^2(\lambda), ..., \sigma^K(\lambda)\right]$, to a spectral reflectance from a surface $\varsigma(\lambda)$ under a spatially constant spectral illumination $\varepsilon(\lambda)$ where λ is wavelength in the wavelength interval ω, are given by:

$$\underline{\rho} = \int_{\omega} \varepsilon(\lambda)\varsigma(\lambda)\underline{\sigma}(\lambda)d\lambda \tag{1}$$

and $\omega = \{\lambda \in \mathbb{R} | 400 \leq \lambda \leq 700\}$ measured in $[nm]$. Obtaining $\varsigma(\lambda)$ on the surface of a patch, generally means spectrophotometric or spectroradiometric measurement averaged over a suitably small but finite sample area on each patch. The method theoretically requires, that ω should encompass all wavelengths, in which the camera responds to any of the chosen colour patches given the illuminant. This may include extensions of the interval, beyond the visible range. For convenience, we have assumed that the camera responses beyond ω are negligible. The measurement areas of the patches are spatially positioned in a $N \times 2$ matrix of coordinates $\mathbf{x} = [\underline{x}, \underline{y}]$, where N is the number of patches on the colour chart. Even if patch surfaces on colour charts are held as uniformly reflecting and diffuse as possible, a small amount of specular reflection is expected. The amount of specular reflection from a patch depends largely on angular subtense between the measuring device and the light relative to the target.

We define $e(\lambda)$ the spectral specular reflection from a patch surface area in the measuring direction. Here, we simplify the spectral distribution of the specular reflection to assume that it is equal to the perfect reflecting diffuser and thus giving the specular component the colour of the illuminant [31].

In real situations, the irradiance intensity very likely varies over the entire chart area. This leaves the actual recorded camera responses $\underline{\rho}^{raw}$ in a patch dependent on spatial position on the chart. The power of the irradiation incident on the sample area of each individual patch is, however, as an approximation, considered to be varying linearly over its spatial area when it is suitably small. So, in the geometrical centre position of the sample area, the averaged responses from that area, defines the responses from said patch.

Furthermore, in a raw camera signal there will be a noise contribution, mainly composed of dark noise, readout noise and image dependent shot noise. We designate noise ν and consider readout noise as a part of dark noise. In an n^{th} patch on the chart, denoted with subscript $(...)_n$ and spatially positioned in $\underline{x}_n = [x_n, y_n]$, the raw camera response $\underline{\rho}_n^{raw}$ under the spatially dependent irradiation (including vignetting) over the chart $\varepsilon_n(\lambda)$ to the equivalently spatially dependent camera response $\underline{\rho}_n$, becomes:

$$\underline{\rho}_n^{raw} = \int_{\omega} \varepsilon_n(\lambda)(\varsigma_n(\lambda) + w_n e(\lambda))\underline{\sigma}(\lambda)d\lambda + \underline{\nu}_n \tag{2}$$

where $w_n \ll 1$ for $n \in \mathcal{n}$ where $\mathcal{n} = \{1, 2, ..., N\}$ scales the contribution of specular reflection down to a realistically small amount considering the mostly diffuse reflection properties of colour charts. Let $w_n\underline{e}_n$ be the according camera response to the specular reflection:

$$\underline{e}_n = \int_{\omega} \varepsilon_n(\lambda)e(\lambda)\underline{\sigma}(\lambda)d\lambda \tag{3}$$

and define an irradiance function in patch n:

$$\underline{I}_n(\rho_n) = \int_{\omega} \varepsilon_n(\lambda)\varsigma_n(\lambda)\underline{\sigma}(\lambda)d\lambda \tag{4}$$

then in patch n, when considering noise:

$$\underline{\rho}_n^{raw} = I_n(\underline{\rho}_n + w_n\underline{e}_n + \underline{\nu}_n^s) + \underline{\nu}_n^d \tag{5}$$

where $\underline{I}_n(\underline{\nu}_n^s)$ is the shot noise contribution attenuated by the power of the irradiance and $\underline{\nu}_n^d$ is the dark noise contribution. We assume that shot noise is *Poission* distributed and thus largely will average out in the patch sampling. Dark and readout noise is also assumed to be *Poission* distributed and thus will average to a constant offset throughout the image and so in the sampled patch values. Otherwise, noise modelling is not specifically part of the method definition. We also assume that quantization errors are dampened in the sampling procedure.

TABLE I

THE ROWS PERTAIN TO A SUBSET OF THE REFERENCED METHODS THAT ROUGHLY RELIES ON ONE SHOT ONLY AND OFFERS A PARTIAL COMPARISON WITH OUR METHOD. IN THE COLUMNS, THE REQUIRED INPUT-DATA, THE OUTPUT-DATA AND THE CHARACTERISTICS OF THE MAIN LINEARISATION ELEMENTS ARE SPECIFIED IN THE THIRD ROW. THESE PERTAIN TO EACH METHOD LISTED IN THE ROWS BELOW. THE QUANTITATIVE COMPARISON IN THE LAST COLUMN MUST TAKE THE DIFFERENCES OF THE METHOD CHARACTERISTICS INTO ACCOUNT. THE X'S DENOTE THAT A METHOD CHARACTERISTIC IS INCLUDED. F/I: FULL IMAGE ANALYSIS. CC: ONLY COLOUR CHART VALUE-BASED ANALYSIS. GC: GRAY-CHART-BASED IRRADIANCE COMPENSATION. VIGN: ONLY VIGNETTING CORRECTION. RMSE RESULTS (RELATIVE TO BIT-DEPTH): PERTAIN TO THE RECOVERY GOAL FROM EITHER INPUT RAW OR INPUT JPEG VALUES, I.E. RAW/JPEG. ΔE: MEANS ACCURACY REPORTED IN CIELAB-VALUES. (JPEG): MEANS RMSE-VALUES MEASURED IN JPEG VALUES. TRAINING: IMPLIES AUXILIARY DATA INPUT LIKE CAMERA-SPECIFIC DATABASE SUPPORT. (+): GAMUT EDGE CUT-OFFS RE-ESTIMATED.

Method Comparison Overview of relying on one shot images for calibration									
	output	input	input	input	input	output	output	output	RMSE
Method	Recovery Goal	Proxy	RGB	SpecR	Training	Gamut	Tone	Irr	raw/JPEG
AP	$\rho^M(+)$	-	CC	CC	-	X	X	X	0.015/0.023
[8]	ρ^{raw}	X	F/I	-	X	X	X	-	-/0.01
[9]	$\rho(Vign)$	-	F/I	-	X	-	X	Vign	0.06(JPEG)
[12]	$\rho^{raw}(+)$	-	F/I	-	X	X	X	-	-/0.03
[13]	ρ^{raw}	X	F/I	-	-	X	X	-	0.04(JPEG)
[14]	ρ^{raw}	X	F/I	-	X	X	X	-	-/0.01
[16]	ρ^{raw}	-	F/I	-	X	X	X	-	-/0.003
[17]	ρ^{raw}	-	F/I	-	-	-	X	-	-/0.012
[18]	ρ^{raw}	X	F/I	-	-	X	X	X	-/0.013
[25]	ρ^{prx}	-	F/I	-	-	X	X	GC	-/0.016
[28]	ρ^{raw}	X	CC	CC	-	-	-	X	ΔE

Here it is implied that the specular reflection $w_n\underline{e}$ is attenuated by the irradiant geometry in the raw responses. When formalizing the spatially dependent specular contribution to the response in the n^{th} patch by:

$$\underline{W}_n(\underline{\rho}_n) = \underline{\rho}_n + w_n\underline{e}_n \tag{6}$$

and inserting Equation (6) in Equation (5), then the camera raw response is fully described by:

$$\underline{\rho}_n^{raw} = \underline{I}_n(\underline{W}_n(\underline{\rho}_n) + \underline{\nu}_n^s) + \underline{\nu}_n^d \tag{7}$$

B. From $\mathcal{X}$ to Ground Truth

The transformation from camera raw responses $\underline{\rho}_n^{raw}$ to rendered camera responses, $\underline{r}_n$ in the n^{th} patch, typically involves five operations.

1) Initially, attempts to correct the raw data for dark noise are made. This typically involves pixel-by-pixel subtracting a raw image recorded with the lens cap on or an in-camera procedure in which non-irradiated pixels on the CCD serve as dark noise reference values [10].

2) A correction for the irradiance geometry $I(...)$ is conducted, if the geometry (i.e. including vignetting) is not *a-priori* known to be spatially even. This is often achieved by pixel-by-pixel normalising the responses in the chart image with the according pixel-by-pixel responses from a spatially perfectly registered image of a neutral surface.

Now, the raw responses can be considered to be linearised and ready to be used in a colour characterisation of the camera.

Here, however, we *omit* these preliminary signal corrections. They are precluded by the goal of the method, as there are no in-register gray chart, in-camera linearisation steps nor lens-cap on images at hand.

3) The colour correction $\mathcal{C}(...)$ is applied to the raw data. Very often this colour correction is linear and therefore based on a 3 by 3 matrix $\mathbf{M}_{cc}$, relating raw responses to colourimetric values, but can also be non-linear.

Depending on illuminant changes and knowledge of it, a chromatic adaptation matrix $\mathbf{M}_{ca}$ may be applied in the colour correction operation. The resulting estimated colourimetric values are then linearly transformed by a 3 by 3 matrix $\mathbf{M}_{cs}$ to an output working colour space, such as defined in the srgb.

4) It is quite likely that colour correction and colour space transformations will yield out-of-gamut values. In that case, a gamut mapping operation $\mathcal{G}(...)$ on these values is applied, which in its simplest implementation per channel cuts channel values to the gamut edge. More advanced methods compress the colour values in the vicinity of the out-of-gamut colour in order to prevent sharp colour changes [32]. Assuming that $\mathcal{G}(...)$ leaves in-gamut colours untouched (i.e. only moves out-of-gamut colour values to the gamut edge), it is observed that:

$$\underline{\mathcal{G}}(\underline{r})\mathbf{M} = \underline{\mathcal{G}}(\underline{r}\mathbf{M}) \tag{8}$$

or simply:

$$\underline{\mathcal{G}}(\underline{r}) = \underline{r} \tag{9}$$

where $\underline{r}$ is an arbitrary $1\times K$ colour vector unaltered by $\underline{\mathcal{G}}(...)$. In practice, if $\mathcal{G}(...)$ incorporates gamut compression involving in-gamut colours, it is difficult to separate the individual contributions of the colour and the gamut correction when analysing their combined transformation of a signal. It is, however, fair to assume that colour channel values on the gamut edge are results of gamut clipping. Thus, easily identifiable, they can be omitted from the input data. Their linear responses can however ultimately be estimated based on the estimated linear responses of the included in-gamut colours. This estimation may be challenged when the compression is image-dependent, in which case there may be no unique intra-image nor inter-image relation between colour values before and after gamut compression [33]. We will not consider this possibility, and thus fuse $\mathcal{G}(\mathcal{C}(...))$ into $\mathcal{C}(...)$.

5) Denoting a vector value in a k^{th} channel $(...)^k$, where $k \in \mathcal{K}$ where $\mathcal{K} = \{1, 2, ..., K\}$, which corresponds to a column value, a non-linear per channel transformation $\Gamma^k(...)$

is applied to optimize the colour values to the available bit-depth. A contrast enhancement correction curve is considered to be incorporated in $\Gamma^k(...)$.

The pipeline of signal modifications and corrective steps from *ground truth* camera responses to rendered values can be summarised as follows: Set $\mathbf{M} = \mathbf{M}_{cc} \times \mathbf{M}_{ca} \times \mathbf{M}_{cs}$. The rendered colour value r_n^k of the k^{th} channel in the n^{th} patch becomes:

$$r_n^k = \Gamma_n^k(\mathcal{G}_n^k(\underline{\mathcal{C}}_n(\underline{\rho}_n^{raw}\mathbf{M}))) \tag{10}$$

which, with insertion of Equation (7), leads to:

$$r_n^k = \Gamma_n^k(\mathcal{G}_n^k(\underline{\mathcal{C}}_n((\underline{I}_n(\underline{W}_n(\underline{\rho}_n) + \underline{\nu}_n^s) + \underline{\nu}_n^d)\mathbf{M}))) \tag{11}$$

For readability, we introduce the matrix M as superscript $(...)^M$ when, say a $1 \times K$ vector $\underline{v}$, is linearly transformed by the matrix $\mathbf{M}$, so that:

$$\underline{v}^M = \underline{v}\mathbf{M} \tag{12}$$

Employing Equations (9) and (12) on Equation (11) then, for all K channels:

$$\underline{r}_n = \underline{\Gamma}_n(\underline{\mathcal{C}}_n(\underline{I}_n(\underline{\rho}_n^M + w_n\underline{e}^M + \underline{\nu}_n^{s,M}) + \underline{\nu}_n^{d,M})) \tag{13}$$

The goal is, *within a linear transformation* $\mathbf{M}$, to estimate the *ground truth* camera responses to the diffuse reflection, illuminated by a spatially uniform light, from each patch. These values will pair exposure independently to the according colourimetric values referred to the actual illuminant used in a camera characterisation. That means solving Equation (13) by estimating $\underline{\rho}_n$ given $\underline{r}_n$ and ς_n for $n \in \mathcal{n}$, while compensating for noise and specular reflection and simultaneously to discover estimates of the transformation functions $\underline{\Gamma}_n(...)$, $\underline{\mathcal{C}}_n(...)$ and $\underline{I}_n(...)$, each of which identifies with an according transformation function *type*. Since each transformation function has operated on the colour channel values resulting from the preceding function, the solution can be found by individual function inversion in accordance with the reverse order of their successive application in Equation (13).

IV. Method

Our linearisation method relies on establishing near-additive relations between the spectral reflection functions on each of the patches on the target. These relations are explored to define an *additivity principle* in which a set of N linear spectral estimations of each of the N colour patches, found by the additive combination of the remaining $N-1$ colour patches, define the linear relation between the equivalent linear camera responses. The key is that these *additive combinations are the same in spectral and camera response space, when the responses are linearly related to the reflectance* from the target. The basic idea is to observe that the *additivity principle is broken if the camera responses are non-linear*. The method estimates the per channel multiplicative transformation functions of the non-linear responses necessary to linearise them, at which point the additivity principle is fulfilled. The additivity principle crucially makes the estimation of *ground truth* independent of any canonical set of responses, as it is *invariant to metamerism*.

A. Multiplicative Correction for non-Linearity

Define three real analytic functions $f, g, h : \mathbb{R}^p \setminus 0 \to \mathbb{R} \setminus 0$ where p is a positive integer. If $p = 1$ and temporarily give a variable ϑ as an argument, where $\vartheta \in \mathbb{R} \setminus 0$ then define:

$$f(\vartheta)g(\vartheta) = h(\vartheta) \tag{14}$$

In Equation (14) $f(\vartheta)$ works as a relative function of ϑ, which maps $g(\vartheta)$ to the value of $h(\vartheta)$. Multiplying $h(\vartheta)$ with $f^{-1}(\vartheta)$ will map $h(\vartheta)$ back to $g(\vartheta)$. Introducing an arbitrary constant $\alpha \in \mathbb{R} \setminus 0$, Equation (14) can be rearranged to:

$$\alpha f(\vartheta)g(\vartheta) \cong h(\vartheta) \tag{15}$$

Irrespectively of the value of α, $\alpha g(\vartheta)f(\vartheta)$ is linearly related to $h(\vartheta)$ (i.e. the linear relation being invariant to scaling). $f(\vartheta)$ is thus known as a *linearisation function* pertaining to $g(\vartheta)$ and $h(\vartheta)$.

Generalising to vector notation, Equation (14) can be expanded to:

$$\underline{f}(\underline{\vartheta}) \circ \underline{g}(\underline{\vartheta}) = \underline{h}(\underline{\vartheta}) \tag{16}$$

where $\circ$ denotes the *Hadamard* matrix product [34], by which, matrices of the same dimension are multiplied entry-wise. This multiplicative correction procedure is used in [35] in conjunction with per channel non-linearity mapping and a monotonicity constraint.

B. Modelling the Transformation Functions

Employing entry-wise multiplication and the mapping defined in Equation (14) or entry-wise addition, the transformation functions in Equation (13) can be formulated as mapping functions using linearisation functions. Let a $1 \times K$ vector $\underline{\vartheta}_n$ denote image colour channel values in an n^{th} patch in position $\mathbf{x}_n$, at any state of rendition.

The per channel tone curve correction is defined as a function of a channel value, which attenuates the said channel value, so that:

$$\Gamma : \vartheta_n^k \mapsto \Gamma_n^k\vartheta_n^k, \quad \text{where:} \quad \Gamma_n^k = \Gamma^k(\vartheta_n^k) \tag{17}$$

The geometrical shape of the spatial irradiance on the chart surface is defined as a function of the spatial position of a patch on the chart, which, according to the irradiant power incident on the chart, attenuates the reflection, so that:

$$I : \vartheta_n^k \mapsto I_n\vartheta_n^k, \quad \text{where:} \quad I_n = I(\mathbf{x}_n) \tag{18}$$

The geometrical shape of the spatial distribution of the specular reflection on the chart surface is defined as a function of spatial position on the chart, which, in an n^{th} patch on the chart, adds to the reflection, so that:

$$W : \vartheta_n^k \mapsto \vartheta_n^k + W_n^k, \quad \text{where:} \quad W_n^k = W^k(\mathbf{x}_n) \tag{19}$$

In each n^{th} patch, the colour correction function is defined as a function of spatial position with coordinates $\underline{\vartheta}_n$ in the K dimensional colour channel space, which maps each channel value, so that:

$$\mathcal{C} : \vartheta_n^k \mapsto \mathcal{C}_n^k\vartheta_n^k, \quad \text{where:} \quad \mathcal{C}_n^k = \mathcal{C}^k(\underline{\vartheta}_n) \tag{20}$$

Upon discretisation, each function operates as a multiplicative or additive set of values, each of which are arranged in a look-up table of dimension $N \times K$.

Applying the multiplicative vector notation in Equation (16) to Equation (13) we obtain:

$$\underline{r}_n = \underline{\Gamma}_n \circ \underline{\mathcal{C}}_n \circ (\underline{I}_n \circ (\underline{\rho}_n^M + w_n \underline{e}_n^M + \underline{\nu}_n^{s,M}) + \underline{\nu}^{d,M}) \quad (21)$$

or in full $N \times K$ matrix form:

$$\mathbf{r} = \mathbf{\Gamma} \circ \mathcal{C} \circ (\mathbf{I} \circ (\rho^M + \mathbf{W} + \nu^{s,M}) + \nu^{d,M}), \quad (22)$$

where $\mathbf{W} = [\underline{w}, \underline{w}, ..., \underline{w}] \circ \mathbf{e}^M$, $\mathbf{I} = [\underline{I}, \underline{I}, ..., \underline{I}]$ and the $N \times K$ matrix $\mathbf{e}^M = [\underline{e}^M; \underline{e}^M; ...; \underline{e}^M]$.

C. Additive Relations of Spectral Reflectances

Observing that any spectral reflectance function in a given n^{th} patch $\varsigma_n(\lambda)$, where $n \in \mathcal{n}$ can be estimated to $\widehat{\varsigma}_n(\lambda)$ by a linear combination of the remaining $N-1$ spectral reflectance functions on the chart by:

$$\widehat{\varsigma}_n(\lambda) = \sum_{i=1}^{n-1} a_i \varsigma_i(\lambda) + \sum_{i=n+1}^{N} a_i \varsigma_i(\lambda) \quad (23)$$

where the weighting factors $a_i \in \mathbb{R}$ for $i \in \mathcal{n} \setminus n$. The accuracy of this estimation can be expressed in terms of a spectral additivity residual $\varepsilon_n(\lambda)$ so that:

$$\widehat{\varsigma}_n(\lambda) = \varsigma_n(\lambda) + \varepsilon_n(\lambda) \quad (24)$$

The Euclidean distance $\|\varepsilon_n(\lambda)\|^2$ between the real and the estimated spectral reflectance functions of the approximation in each of these N equations depends on the dimensionality of the set of N spectral reflectances on the chart. It is noted that $\varepsilon_n(\lambda)$ is solely dependent on the measured chart spectra and thus subject to spectral measuring error. It is further assumed that the spectra are generally broadband and overlapping.

D. Chart-Specific Additivity Matrix

Upon discretisation, each of the N spectral reflectances $\varsigma_n(\lambda)$ is expressed as a *dim* $\times$ 1 vector $\underline{\varsigma}_n$, where *dim* $= 31$ (based on available tabulated data) and subsequently by following Equation (23), an $N \times N$ *Chart-Specific Additivity Matrix* $\mathbf{A}$ in which:

$$\mathbf{A} = [\underline{a}_1, \underline{a}_2, ..., \underline{a}_N]^T \quad (25)$$

can be defined, where $\underline{a}_n$ is a $N \times 1$ vector for $n \in \mathcal{n}$. The values in $\mathbf{A}$ are found by solving:

$$\underline{a}_n^\star = (\mathbf{S}_n^T \mathbf{S}_n)^{-1} \mathbf{S}_n^T \underline{\varsigma}_n \quad (26)$$

where $\mathbf{S}_n = [\underline{\varsigma}_1, \underline{\varsigma}_2, ..., \underline{\varsigma}_{n-1}, \underline{\varsigma}_{n+1}, ..., \underline{\varsigma}_N]$ is an *dim* $\times (N-1)$ matrix and $\underline{a}_n^\star$ is an $N-1$ vector, so that:

$$\underline{a}_n = [(\underline{a}_n^\star)_1, (\underline{a}_n^\star)_2, \ldots, (\underline{a}_n^\star)_{n-1}, -1, (\underline{a}_n^\star)_n, \ldots, (\underline{a}_n^\star)_{N-1}]^T$$

$\mathbf{A}$ is non-singular, though poorly conditioned and requires regularisation in conjunction with inversion. Following Equation 24, a vector of the spectral additivity residuals $\varepsilon(\lambda)$ for the N reflectance spectra pertaining to our colour chart can be defined as $\mathcal{E}_\lambda = [\varepsilon_1(\lambda); \varepsilon_2(\lambda); ...; \varepsilon_N(\lambda)]$. Discounting measurement error in $\varsigma(\lambda)$, this error is small[1], as $\mathrm{RMSE}(\mathbf{A}^T \mathcal{E}_\lambda) \simeq 10^{-7}$. Note, that once calculated, $\mathbf{A}$ *replaces the reflectances as method input for subsequent chart-specific captures*.

E. Capture-Specific Additivity

Observing, by referring to Equation (5), that a set of linear combinations in spectral space is the same in linear colour space [4], it is readily seen that Equation (23) is equivalent to:

$$\underline{\rho}_n = \sum_{i=1}^{n-1} a_i \underline{\rho}_i + \sum_{i=n+1}^{N} a_i \underline{\rho}_i - \underline{\delta}_n, \quad (27)$$

where $\underline{\delta}_n^k$ is the camera response to the spectral additivity residual in the k^{th} channel of the n^{th} patch. Then for $i \in \mathcal{n} \setminus n$ and $k \in \mathcal{K}$:

$$\underline{a}_i^T \underline{\rho}^k = -\delta_i^k \quad (28)$$

or generalising Equation (28) by channel:

$$\mathbf{A}^T \underline{\rho}^k = -\underline{\delta}^k \quad (29)$$

For normalised *ground truth* responses $\mathrm{RMSE}(\mathbf{A}^T \underline{\rho} / \underline{\rho}_{max})) \simeq 10^{-10}$ pertaining to our camera sensors. The rmse-value is not likely to vary much for other sensors and illuminants.

Henceforth, multiplying an additivity matrix with a set of responses, as in Equation (29) will be known as applying the *Additivity Principle* (*AP*).

The additivity principle is invariant to linear transformations of responses. Multiplying a 3×3 matrix $\mathbf{M}$ on both sides of Equation (29), yields:

$$\mathbf{A}^T \rho^M = -\delta \mathbf{M} \quad (30)$$

Considering that, except for a simple scaling, $\delta \simeq \delta \mathbf{M}$, and using Equation (12), it is readily seen that applying the additivity matrix is quasi-invariant to linear transformation:

$$\mathbf{A}^T \rho \simeq \mathbf{A}^T \rho^M \quad (31)$$

F. Solution

We wish to calculate the values of ρ^M in Equation (22). By inversion, where each of the $N \times K$ elements of the transformation functions have been inverted entry-wise $(\mathbf{\Gamma}^{-1})_n^k = (\Gamma_n^k)^{-1}$, $(\mathcal{C}^{-1})_n^k = (\mathcal{C}_n^k)^{-1}$ and $(\mathbf{I}^{-1})_n^k = (I_n^k)^{-1}$, we obtain:

$$\rho^M = \mathbf{r} \circ \mathbf{\Gamma}^{-1} \circ \mathcal{C}^{-1} \circ \mathbf{I}^{-1} - \mathbf{W} - \nu^{s,M} - \nu^{d,M} \circ \mathbf{I}^{-1} \quad (32)$$

Assembling the unknown noise in $\nu = \nu^{s,M} - \nu^{d,M} \circ \mathbf{I}^{-1}$ and concomitantly combining it with $\mathbf{W}$, a total *Offset* denoted as a $N \times K$ matrix $\mathcal{O}$ is defined, where:

$$\mathcal{O} = \nu^{s,M} + \nu^{d,M} \circ \mathbf{I}^{-1} + \mathbf{W} \quad (33)$$

From Equation (33) it can be seen that $\mathcal{O}$ is dependent on the variation of the geometry of the irradiance. We assume that the effect is negligible, as the dark noise is typically relatively small compared to the total camera response and that, based

[1]Pertaining to X-rite ColourChecker SG chart excluding gray rim patches

on experience with our data, the irradiance variation in our data sets $\mathbf{I}_{max}/\mathbf{I}_{min} < 1.2$, and thus attenuates the dark noise with less than 20%. It should be noted that vignetting can be a considerable contributor to $\mathbf{I}$ and is reported to be up to 279% [10]. This is however over the full field of view. Our samples originate from about 20% of the full field of view belonging to the centre of the images. To investigate the impact of vignetting, we have included a test of severe vignetting simulated on ground truth and real raw in Section VII.B.

Equation (32) simplifies to:

$$\rho^M = \mathbf{r} \circ (\mathbf{\Gamma}^{-1} \circ \mathcal{C}^{-1} \circ \mathbf{I}^{-1}) - \mathcal{O} \tag{34}$$

or equivalently:

$$\mathbf{r} = (\rho^M + \mathcal{O}) \circ (\mathbf{\Gamma}^{-1} \circ \mathcal{C}^{-1} \circ \mathbf{I}^{-1})^{-1} \tag{35}$$

Using the multiplicative linearisation function definition in Equation (16) we seek the linearisation function $\mathbf{L}$, given in a $N \times K$ matrix, such that the rendered camera response $\mathbf{r}$ can be mapped to the *ground truth* linear responses ρ^M by:

$$\rho^M = \mathbf{r} \circ \mathbf{L} \tag{36}$$

Combining Equation (36) with Equation (35) yields:

$$\rho^M = (\rho^M + \mathcal{O}) \circ (\mathbf{\Gamma}^{-1} \circ \mathcal{C}^{-1} \circ \mathbf{I}^{-1})^{-1} \circ \mathbf{L} \tag{37}$$

which demonstrates the impact of $\mathcal{O}$ on the estimation of $\mathbf{L}$ and consequently of $\mathbf{\Gamma}^{-1}, \mathcal{C}^{-1}$ and $\mathbf{I}^{-1}$ as:

$$\mathbf{L} = \rho^M / (\rho^M + \mathcal{O}) \circ (\mathbf{\Gamma}^{-1} \circ \mathcal{C}^{-1} \circ \mathbf{I}^{-1}) \tag{38}$$

In an ideal case when noise and specular reflection are absent from $\mathbf{r}$ denoted $\mathbf{r}^{(ideal)}$, then $\mathcal{O}$ becomes $N \times K$ matrix of zeros and Equation (32) reduces to:

$$\rho^M = \mathbf{r}^{(ideal)} \circ \mathbf{\Gamma}^{-1} \circ \mathcal{C}^{-1} \circ \mathbf{I}^{-1} \tag{39}$$

which would make an according multiplicative linearisation function $\mathbf{L}^{(ideal)}$ in Equation (38) consist exclusively of three multiplicative transformations:

$$\mathbf{L}^{(ideal)} = \mathbf{\Gamma}^{-1} \circ \mathcal{C}^{-1} \circ \mathbf{I}^{-1} \tag{40}$$

G. Solving by Additivity Principle

To control the linearisation of $\mathbf{r}$ by $\mathbf{L}$ to estimate ρ^M and $\mathbf{L}$ in Equation (36) we employ the *Additivity Principle* by:

$$\mathbf{A}^T(\mathbf{r} \circ \mathbf{L}) = -\delta\mathbf{M} \tag{41}$$

and use Equation (36) to obtain ρ^M. In Equation (41) the right side $-\delta\mathbf{M}$ is unknown, constant and generally considered negligible. We therefore seek a modified estimate of $\widehat{\mathbf{L}}$ as solution to:

$$\mathbf{A}^T(\mathbf{r} \circ \widehat{\mathbf{L}}) = \mathbf{0} \tag{42}$$

thereby obtaining an estimated linear camera response;

$$\widehat{\rho}^M = \mathbf{r} \circ \widehat{\mathbf{L}} \tag{43}$$

To get an optimal solution, we propose to use constrained linear least squares minimisation and thus for a k^{th} channel, by minimising:

$$\phi^k = \min_{\underline{\widehat{L}}^k} (\|\mathbf{A}^T(\underline{r}^k \circ \underline{\widehat{L}}^k)\|^2) \quad \text{s.t.} \quad \sum_{i=1}^{N} \widehat{L}_i^k = N \tag{44}$$

so that, for $n \in \mathcal{N}$ and $k \in \mathcal{K}$, we avoid the obvious zero solution by conveniently constraining $\widehat{\mathbf{L}}$ in order for $\widehat{L}_n^k = 1$ when $\underline{r}^k$ is linear. We obtain estimated values of the *ground truth* linear camera responses except for a linear $K \times K$ transformation $\underline{\widehat{\rho}}^{M,k}$, where:

$$\underline{\widehat{\rho}}^{M,k} = \underline{r}^k \circ \underline{\widehat{L}}^k \tag{45}$$

H. Regularisation

As $\mathbf{A}^T$ and thus likely $\mathbf{A}^T\underline{r}^k$ are poorly conditioned, a Tikhonov type regularisation is included in Equation (44). We choose to employ regularisation of high physical justification, by utilising the assumption of smoothness (discounting high-frequency noise) in the mapping from ρ to $\mathbf{r}$ described in Equation (22) of functions $\Gamma(...)$, $\mathcal{C}(...)$ and $I(...)$ and consequently of their inverted counterparts $\Gamma^{-1}(...)$, $\mathcal{C}^{-1}(...)$ and $I^{-1}(...)$.

In accordance with Equations (17), (18) and (20), the linearisation function $\mathbf{L}$ is dependent on camera response ρ and spatial position $\mathbf{x}$. We suggest measuring smoothness in terms of second order derivatives of each linearisation function within $\underline{L}^k$.

To facilitate regularisation pertaining to each channel k and each linearisation function type, we need to split $\underline{L}^k$ into separate multiplicative linearisation contributions, each of which corresponds to the domain of $\Gamma(...)$, $\mathcal{C}(...)$, $I(...)$. The gamma-function is smooth as a function of ranked response, neither of the remaining two functions are. A constant noise contribution to the irradiance function is smooth as a function of spatial position of the patches, neither of the remaining two are. However, the irradiance function is multiplicative and cannot compensate for an additive contribution from the specular component and vice versa, the specular function is additive and cannot compensate for a multiplicative contribution from the irradiance function. It is possible, though, that a constant average noise contribution added to all patches can be compensated in the gamma function. The volumetric colour correction function is not smooth as a function of either spatial position or (typically) capable of compensating for an additive contribution. Colour correction functions (e.g. polynomials) are typically mapping as a function of added weighted higher order mixtures of the raw responses, each of which (except one) lacks one of the channel responses. This means that the colour correction function is not necessarily smooth as a function of ranked responses. It is expected that a small amount of cross-talk between the estimated linearisation functions exist. This is caused by quantisation, local intervals of the domains in which similar mapping between functions exist, noise and the built-in inaccuracy within the additivity principle given by $-\delta$ in Equation (29).

With a view to successively solving for each individual linearisation function in Equation (44), we split $\widehat{\underline{L}}^k$ into:

$$\widehat{\underline{L}}^k = \widehat{\underline{L}}^k_\Gamma \circ \widehat{\underline{L}}^k_{\widehat{\mathcal{C}}} \circ \widehat{\underline{L}}^k_I, \tag{46}$$

where $\widehat{\underline{L}}^k_\Gamma$ operates on the colour channels (1D), $\widehat{\underline{L}}^k_I$ in the spatial domain of the colour patches (2D), and $\widehat{\underline{L}}^k_{\widehat{\mathcal{C}}}$ in the colour space (3D).

The second order derivatives of $\widehat{L}^k_{(...)}$ cannot be computed by the standard finite difference method since they are all defined on irregular grids: digital camera responses, spatial patch coordinates, and colour coordinates, respectively. Instead, we propose to use the graph Laplacian. For the 2D and 3D cases, the Delaunay triangulation is used, and for the 1D case, the graph is simply a linked list of the ordered camera responses. Define subscript t to indicate the type of either transformation function $\Gamma(...)$, $\mathcal{C}(...)$ or $I(...)$. Let G_t denote the resulting graph, and D_{G_t} and A_{G_t} denote the corresponding degree and adjacency matrices of G_t, respectively. For the regularisation of $\widehat{L}^k_t$, the graph Laplacian matrix, $L^k_{G_t} = A^k_{G_t} - D^k_{G_t}$ is used to define a substitute for the finite difference Laplacian as follows:

$$\boldsymbol{\Delta}^{2,k}_t \widehat{L}^k_t = L^k_{G_t} \widehat{L^k}_t \tag{47}$$

As an alternative to the simple graph Laplacian matrix, distance weighting can be introduced. This is achieved by using the weighted versions of the degree and adjacency matrices, where the inverse distances between the vertices are used as weights. Experimentally, we found that distance weighting improved the results for the 1D and 3D regularisation, but not for 2D, since the colour charts with a regular spatial layout were used. For the 1D graph Laplacian $L^k_{G_\Gamma}$, better results were obtained when the rows were normalized to have 2 on the diagonal in analogy with the classical Laplacian operator. For the boundary values of the 1D graph Laplacian, the values of the nearest internal node were used in place of the values obtained directly from the graph Laplacian.

With these definitions of the regularisers, Equations (44) and (45) can be reformulated to:

$$\begin{aligned} &\phi^k_t(\lambda) = \min_{\widehat{\underline{L}}^k_t} (\|\mathbf{A}^T(\underline{r}^k \circ \widehat{\underline{L}}^k_t)\|^2 + {\lambda^k_t}^2 \|\boldsymbol{\Delta}^{2,k}_t \widehat{\underline{L}}^k_t\|^2) \\ &\text{s.t.} \quad \sum_{i=1}^{N} \widehat{L}^k_{t,i} = N, \end{aligned} \tag{48}$$

where λ^k_t is the Tikhonov regularisation parameter pertaining to t and k. The regularisation parameters are optimized heuristically. The magnitudes of the parameters decide the weight of the regularisations, i.e. allowance for curvature. Changes in their relative sizes can redistribute smoothness between the functions.

It was found that the amount of smoothness in a series of tests encompassing simulated curved irradiance functions with a shading ration of up to around 3 (to approximate heavy vignetting [10]), linear and second order colour correction and tone curves approximating a channel gamma value of up to 2, it was adequately covered with simply setting $\lambda_\Gamma = 10$, $\lambda_I = 5$ and $\lambda_\mathcal{C} = 0.1$ irrespective of channel k.

In theory, it would be possible to optimize also for the λ_t parameters. However, this would require a cost function that would have to balance the relative importance and the absolute size of the smoothness in the three respective domains, so it would essentially just shift the task of selecting the λ_t parameters to a new domain and allow more or less curvature in the estimated functions. We have kept said λ values fixed throughout.

In [16] an example of a similar λ value is hardcoded based on experience.

I. Resulting Algorithm

Ultimately,

$$\widehat{\underline{\rho}}^{M,k} = \underline{r}^k \circ \widehat{\underline{L}}^k_\Gamma \circ \widehat{\underline{L}}^k_\mathcal{C} \circ \widehat{\underline{L}}^k_I \tag{49}$$

We solve Equation (48) in an iterative scheme, whereby, after it iterations, Equation (48) becomes equivalent to:

$$\begin{aligned} &\phi^{k,(it)}_t({\lambda^k_t}^2) = \\ &\min_{\widehat{\underline{L}}^{k,(it)}_t} (\|\mathbf{A}^T(\widehat{\underline{r}}^{k,(it)} \circ \widehat{\underline{L}}^{k,(it)}_t)\|^2 + {\lambda^k_t}^2 \|\boldsymbol{\Delta}^{2,k,(it)}_t \widehat{\underline{L}}^{k,(it)}_t\|^2); \\ &\widehat{r}^{k,(it)}_{t,i} = r^k_{t,i} \prod_{\iota=1}^{it-1} \widehat{L}^{k,(\iota)}_{t,i}, \quad \underline{\Delta}^{2,k,(it)}_{t,i} = \underline{\Delta}^{2,k}_{t,i} \prod_{\iota=1}^{it-1} \widehat{L}^{k,(\iota)}_{t,i} \\ &\text{s.t.} \sum_{i=1}^{N} \widehat{L}^{k,(\iota)}_{t,i} = N \quad \text{for} \quad i = [1:N] \end{aligned} \tag{50}$$

in which case Equation (49) can be written:

$$\widehat{\rho}^{M,k,(it)}_i = r^k_i \prod_{\iota=1}^{(it)} \widehat{L}^{k,(\iota)}_{\Gamma,i} \widehat{L}^{k,(\iota)}_{\mathcal{C},i} \widehat{L}^{k,(\iota)}_{I,i} \tag{51}$$

The iterative scheme employed to solve Equation (51), is, in accordance with successive reverse order inversion as mentioned in connection with Equation (13), arranged as shown in Algorithm 1.

The iterative scheme is solved by successive inversion of each function type in Equation (13), as each type has operated on the colour channel values resulting from the preceding function. Out-of-gamut colour channel values (i.e. clipped colour values) are excluded in the linearisation, as they will not be linearly related to the in-gamut colour channel values upon multiplicative linearisation. As it is assumed that only a few colour patches are out-of-gamut, this omission will have an insignificant impact on the accuracy of the linearisation.

J. Estimating Omitted Colours

When the estimation of the linearisation functions is complete, it is however possible to estimate the linear values pertaining to the omitted colour channel values by employing the additivity principle. The additivity matrix is used based on all the spectral reflections of the colour patches with non-clipped colour values along with their linear estimates. In Equation (27), $N-1$ patches will refer to the number of in-gamut estimated linear colour values and the n^{th} colour values to the colour values of an out-of-gamut patch. The effect of this estimation is most visible in Figure 5, where mainly colour values in the blue channel have been clipped to zero.

Algorithm 1 Linearisation by Additivity

function LINEARIZE($\mathbf{r}, tol$)
  $it \leftarrow 0$ ▷ initialise
  $\widehat{\underline{\rho}}^{(it)} \leftarrow \mathbf{r}$
  $\widehat{\underline{L}}_{\Gamma}^{k,(0)} \leftarrow [1,1,...,1]^T$
  $\widehat{\underline{L}}_{\mathcal{C}}^{k,(0)} \leftarrow [1,1,...,1]^T$
  $\widehat{\underline{L}}_{I}^{k,(0)} \leftarrow [1,1,...,1]^T$
  compute $\mathbf{\Delta}_{\Gamma}^{2,k}$, $\mathbf{\Delta}_{\mathcal{C}}^{2,k}$ and $\mathbf{\Delta}_{I}^{2,k}$
  repeat ▷ iterate until convergence
    $it \leftarrow it + 1$
    estimate $\widehat{\underline{L}}_{\Gamma}^{k,(it)}$ by Equations (50) and (51)
    $\widehat{\underline{\rho}}^{k,(it-\frac{2}{3})} \leftarrow \widehat{\underline{\rho}}^{k,(it-1)} \circ \widehat{\underline{L}}_{\Gamma}^{k,(it)}$
    $\widehat{\underline{L}}_{\Gamma}^{k,(it)} \leftarrow \widehat{\underline{L}}_{\Gamma}^{k,(it-1)} \circ \widehat{\underline{L}}_{\Gamma}^{k,(it)}$
    estimate $\widehat{\underline{L}}_{\mathcal{C}}^{k,(it)}$ by Equations (50) and (51)
    $\widehat{\underline{\rho}}^{k,(it-\frac{1}{3})} \leftarrow \widehat{\underline{\rho}}^{k,(it-\frac{2}{3})} \circ \widehat{\underline{L}}_{\mathcal{C}}^{k,(it)}$
    $\widehat{\underline{L}}_{\mathcal{C}}^{k,(it)} \leftarrow \widehat{\underline{L}}_{\mathcal{C}}^{k,(it-1)} \circ \widehat{\underline{L}}_{\mathcal{C}}^{k,(it)}$
    estimate $\widehat{\underline{L}}_{I}^{k,(it)}$ by Equations (50) and (51)
    $\widehat{\underline{\rho}}^{k,(it)} \leftarrow \widehat{\underline{\rho}}^{k,(it-\frac{1}{3})} \circ \widehat{\underline{L}}_{I}^{k,(it)}$
    $\widehat{\underline{L}}_{I}^{k,(it)} \leftarrow \widehat{\underline{L}}_{I}^{k,(it-1)} \circ \widehat{\underline{L}}_{I}^{k,(it)}$
  until $\|\widehat{\underline{\rho}}^{k,(it)} - \widehat{\underline{\rho}}^{k,(it-1)}\|^2 \leq tol$
  return $\widehat{\underline{\rho}}^{k,(it)}$, $\widehat{\underline{L}}_{\Gamma}^{k,(it)}$, $\widehat{\underline{L}}_{\mathcal{C}}^{k,(it)}$, $\widehat{\underline{L}}_{I}^{k,(it)}$
end function

K. Computing Overhead

The convergence criteria tolerance tol defined in Algorithm 1 is set to $tol = 10^{-7}$, resulting in occasionally up to around 700 iterations or up to around 20 seconds, are registered using a contemporary Windows Intel i7-11850H PC/Matlab setup. Optimisation of the code and size of tol against estimation accuracy can probably lower these values.

L. Evaluating Linearity

As the estimated linear responses are only *linearly related* to *ground truth*, then in order to compare the data, an optimized linear transformation from estimate to *ground truth* is performed for all estimates. We include a definition of the estimation error of the linear colour values ΔRGB based on Equation (43):

$$\Delta RGB_{i,j} = \|(\widehat{\rho}^M - \mathbf{r} \circ \widehat{\mathbf{L}})_{i,j}\| \tag{52}$$

where:

$$\mathbf{M} = (\widehat{\rho}^T\widehat{\rho})^{-1}\widehat{\rho}^T\rho \quad \text{s.t.} \quad \rho_w = \widehat{\rho}_w\mathbf{M} \tag{53}$$

and $(..)_w$ denotes values of the white patch. The estimated values are thus mapped to *ground truth*. The values of $\widehat{\rho}^M$ are located on the ordinate axis in Figures 4, 5, 6 and 7. If a perfectly linear relation exists between *ground truth* and estimate, then $\Delta RGB = 0$ and paired values of ρ and $\widehat{\rho}^M$ will be on the grey line. The maximum channel value of $\widehat{\rho}^M$ in the white patch is scaled to coordinates $(1, 1)$.

V. MEASURED AND SYNTHETIC DATA

To validate the method, we have at our disposal a set of simulated and real camera data. We have tested the method on real camera data, both raw and rendered, gathered in-doors and out-doors under various combinations of natural light (sun and sky), LED, WLED, fluorescent, and tungsten. We have further simulated camera data at rendering levels $\mathcal{X}$ using these measured lights combined with either noiseless *ground truth* or the according raw data. See entries on the abscissa in Figures 13 and 14.

A *X-rite Digital ColourChecker SG* (SG) chart (After 2014 production) was used as colour chart. A supplemental *X-rite ColourChecker, White balance* (GC) gray chart has been included in some of the images for irradiance geometry reference. The SG chart has been measured with the spectro-radiometer (SpectraScan® Spectroradiometer PR-655) in an integrating sphere including the barium tile in order to enable calculation of the spectral reflection function in each colour patch.[2] The gray patches on the outer rim of the colour chart are omitted from the analysis to remove any dependence in the irradiance geometry estimation. To avoid spatial variation of the patch recordings of the SG, the spectroradiometer was held in a fixed position and the chart was moved patch by patch to the same measuring position. To tentatively test the robustness of the method, the SG chart spectral reflectance functions are also given as a set of functions found on the internet with no further data attached except that the data pertains to the chart production after 2014 (X-Rite does not provide the spectral data, but refers to data measured by third party). The data has probably been measured by a GretagMacbeth spectrophotometer. We refer to the data as *factory*, indicating that the measurement acts as a stand-in for a real set of X-Rite factory endorsed measurements.

The results are validated by using raw and rendered data from the camera of both the colour chart (and a gray chart in perfect registration for irradiance geometry comparison) and a barium sulphate tile within each image where a spectroradiometric measurement of the illuminant pertaining to each shot is performed. The colour chart and the gray chart are shot with the same exposure and in immediate order. The *ground truth* model of the camera has been established by using a camera (Nikon D50 DSLR) with known measured (National Physical Laboratory, UK) spectral sensor functions.

In each in-situ camera recording, the SG was placed within a distance of 2 meters from the camera and centred in the image. The chart was either laying flat or tilted to face the lens. Adjacent to the chart, the tile was placed. No special care was taken to expose the images optimally. Some are seriously under-exposed with only about 10 % of the available bit depth. See per channel raw responses to the white patch in columns 6, 7, and 8 in Table II.

A spectroradiometer was placed directly beside the camera to record the light from the tile, when the image of the chart and tile was recorded by the camera. An according in-

[2]The measurements were performed in the context of the project 'Rank based spectral estimation', EP/J005223/1, https://gow.epsrc.ukri.org/NGBOViewGrant.aspx?GrantRef=EP/J005223/1.

register image of the GC was taken immediately after in order to enable a gray-chart irradiation-geometry normalisation for comparison. The measured *ground truth* camera sensor functions enable a calculation of Equation (1) of the *ground truth* camera responses for each measured spectral distribution of the illumination.

Captures of the GC in geometric register to measure the relative irradiance incident on the colour chart and a simultaneous spectroradiometric measurement of an adjacent barium sulphate tile is provided for each capture in order to obtain the spectral composition of the prevailing illuminant power distribution. Camera raw data[3] as well as the corresponding rendered data (JPEG/srgb) are supplied by the camera.

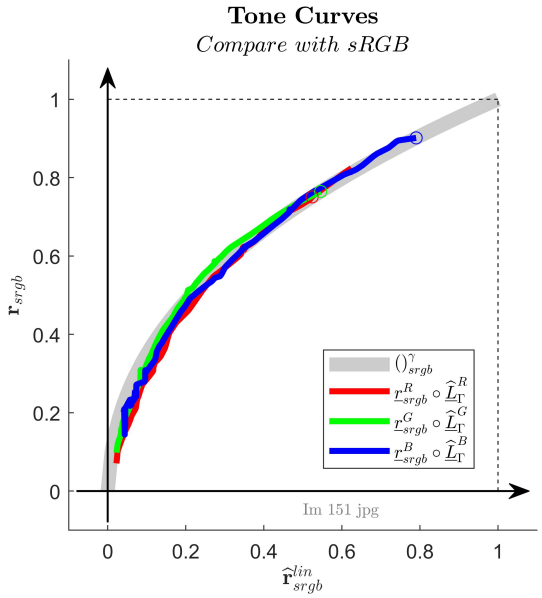

Fig. 2. Estimated tone curves from JPEG $\mathbf{r}$ vs linearised estimated srgb $\widehat{\mathbf{r}}^{lin}_{srgb}$. Here $\mathbf{r}_{srgb} = \mathbf{r}$.

A standalone Matlab-program was developed to load the images and align the picking areas of averaged colour values pertaining to each coloured surface to the patch centres comprising 20 by 20 pixels. The areas were chosen to avoid edge variations, i.e. JPEG artefacts in the colour values between the coloured pigments and the black frames covering about 20 % patch area.

VI. Tests Performed on Data

We examine a set of real, synthetic and combined data. Among the many combinations, the synthetic data is created following the structure of Equation 10. Referring to each set as a *rgb-type* we choose rgb-types covering real data: ρ^{raw} and JPEG $\mathbf{r}$. For rgb-types covering synthetic data Equation: ρ, $\mathbf{I} \circ \rho$, $\mathrm{srgb}(\rho)$ and $\mathrm{srgb}(\mathbf{I} \circ \rho)$ and for rgb-types covering raw with synthetic mapping: $\mathrm{srgb}(\rho^{raw})$, $\mathrm{srgb}(\mathbf{I} \circ \rho^{raw})$ and $\mathbf{r}_r$, where the re-engineered JPEG values (i.e. a simplified version of the actual JPEG values) $\mathbf{r}_r$ are found by using Equation 54:

$$\mathbf{r}_r = \mathrm{srgb}((\mathrm{srgb}^{-1}(\mathbf{r}))\mathbf{M}_r) \tag{54}$$

[3]The raw data is developed externally in *dcraw*, using control parameters -v -r 1 1 1 1 -M -W -o 0 -q 3 -4 -T, which excludes gain, colour-matrixing, white-balancing and dark image subtraction (https://www.dechifro.org/dcraw/)

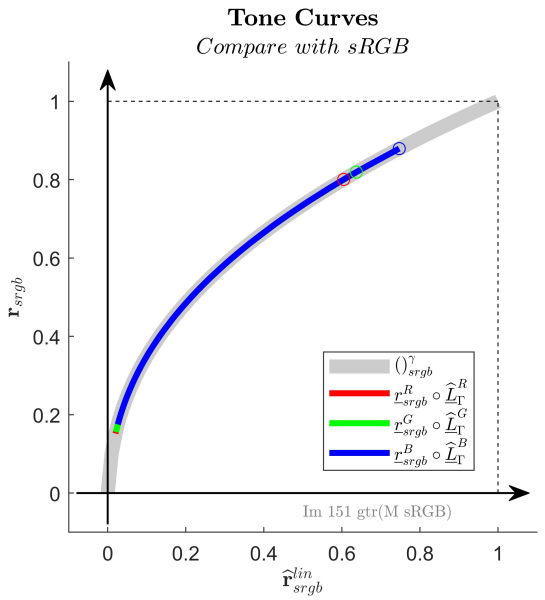

Fig. 3. Estimated tone curves from synthetic $\mathbf{r}_{srgb} = \mathrm{srgb}(\rho)$ compared to linearised estimated srgb $\widehat{\mathbf{r}}^{lin}_{srgb}$.

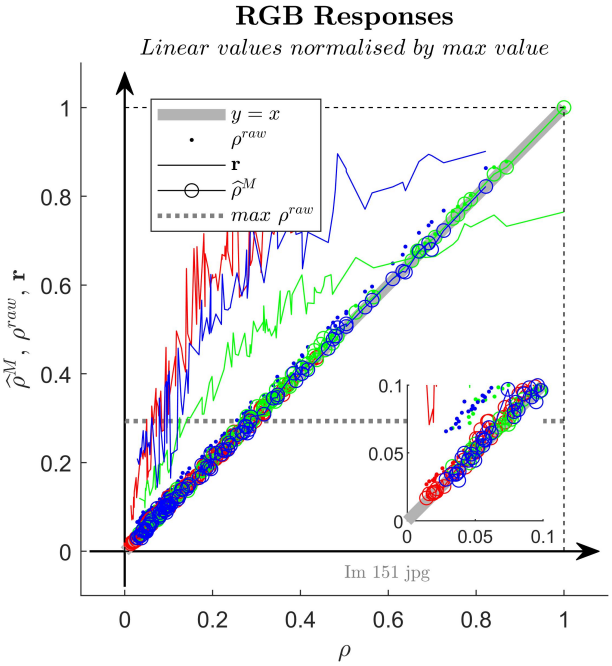

Fig. 4. Estimated $\widehat{\rho}^{M}$ from rendered input JPEG $\mathbf{r}$ vs ρ.

where $\mathbf{M}_r$ denotes a $K \times K$ matrix found by $\min_{\mathbf{M}_r} \|\rho^{raw}\mathbf{M}_r - \mathrm{srgb}^{-1}(\mathbf{r})\|$. The estimated ground truth from matrixed canonical data is represented by $\widehat{\rho}^{M}_{xyz}$ referred to the actual illuminant, for XYZ-based proxy comparison found by Equation 53 substituting $\widehat{\rho}$ with XYZ.

Per pixel normalisation of raw by the gray chart is defined by: $\rho^{gc} = \rho^{raw} \circ (\mathbf{I}^{gc})^{-1}$.

The results for the estimated linearisation functions are shown in the figures, pertaining to linear responses, tone curves and irradiance functions. The estimated perceptual differences between LAB-values pertaining to ρ and estimated linear

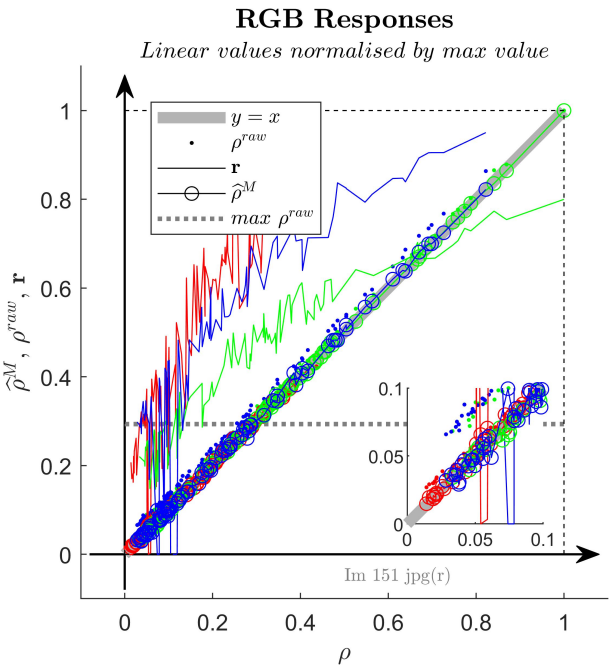

Fig. 5. Estimated $\widehat{\rho}^M$ from reverse engineered JPEG $\mathbf{r}_r$ vs ρ. Here rendered input $\mathbf{r} = \mathbf{r}_r$.

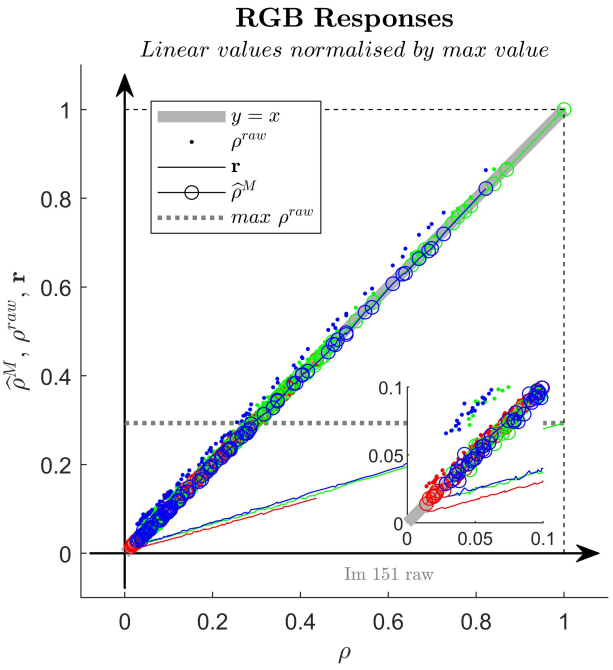

Fig. 6. Estimated $\widehat{\rho}^M$ from ρ^{raw} vs ρ. Here rendered input $\mathbf{r} = \rho^{raw}$.

responses are shown in Figure 11. The white normalisation in the LAB values comes from the XYZ values of the white patch referred to the actual illuminant.

VII. Results and Discussion

As $\widehat{\rho}^M$ is only known within a least squares optimised linear transform of ρ (Equation 53), the functions shown in the figures 2 to 7 can only be plotted as a function of ρ because it is known for evaluation purposes. In real situations, only $\widehat{\rho}^M$ is known through the estimation.

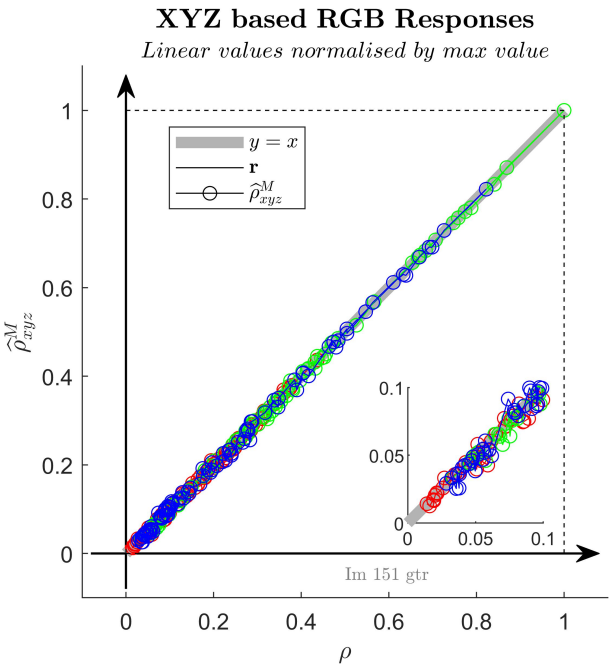

Fig. 7. Estimated Linear RGB from xyz-sensor values based on actual illuminant vs *ground truth*. $\widehat{\rho}^M_{srgb}$ refers to xyz ground truth data ρ_{xyz} linearly transformed to ρ by estimated matrix $\mathbf{M}$.

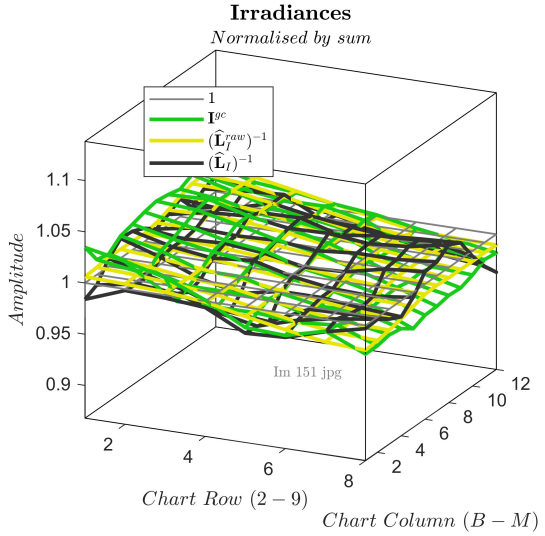

Fig. 8. Estimated irradiance $(\widehat{\mathbf{L}}_\Gamma)^{-1}$ from JPEG $\mathbf{r}$, $(\widehat{\mathbf{L}}_\Gamma^{raw})^{-1}$ from ρ^{raw} along with GC measured $\mathbf{I}^{gc}$.

In figures 2 and 3 the estimated tone curves (thick coloured curves) are plotted on top of the known srgb tone curve (gray curve) for reference. In figure 2, the white point (coloured circles) found in the rendered images are used to scale the tone curve values to the according *ground truth* values of white calculated from reversing srgb values to *ground truth*. In figure 3 the *ground truth* of the white point has been scaled down 20% to avoid out-of-srgb-gamut values in the simulation of the rendered data. Here $\widehat{\mathbf{r}}^{lin}_{srgb}$ denotes estimated $\widehat{\rho}^M$ linearly transformed to the srgb space without its gamma applied. $\mathbf{r}_{srgb}$ are values pertaining to ρ transformed to srgb values. The resulting relation between these to data sets are the per channel

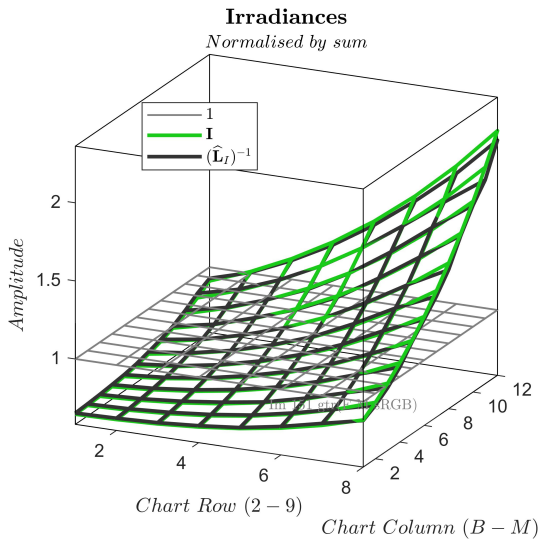

Fig. 9. Estimated irradiance $(\widehat{\mathbf{L}}_\Gamma)^{-1}$ from synthetic vignetting image $\mathrm{srgb}(\mathbf{I} \circ \rho)$, compared to synthetic irradiance $\mathbf{I}$.

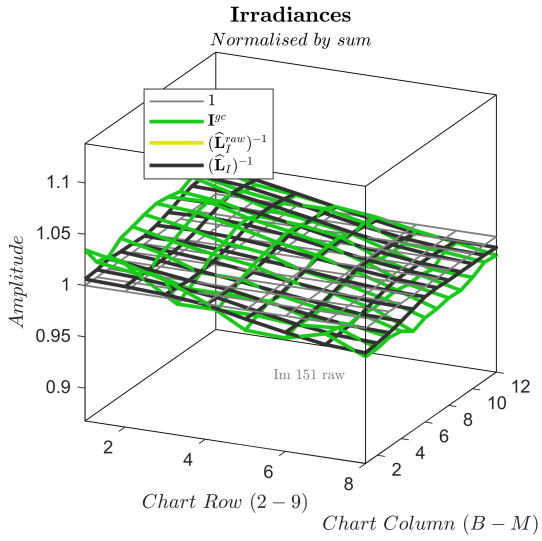

Fig. 10. Estimated irradiance $(\widehat{\mathbf{L}}_\Gamma)^{-1}$ and $(\widehat{\mathbf{L}}_\Gamma^{raw})^{-1}$ both from raw ρ^{raw}, compared to GC measured irradiance $\mathbf{I}^{gc}$.

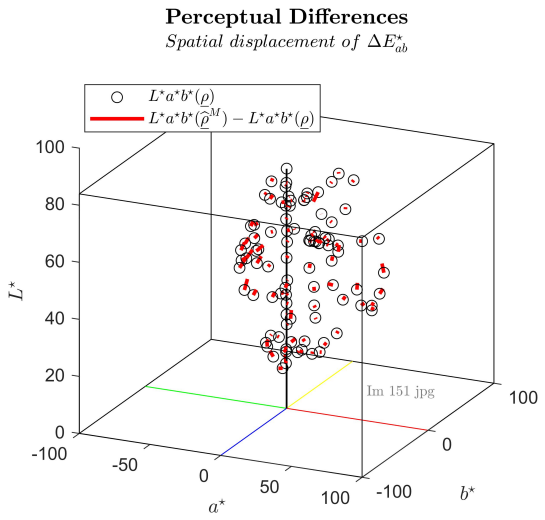

Fig. 11. Estimated Perceptual Differences from JPEG image to *ground truth*

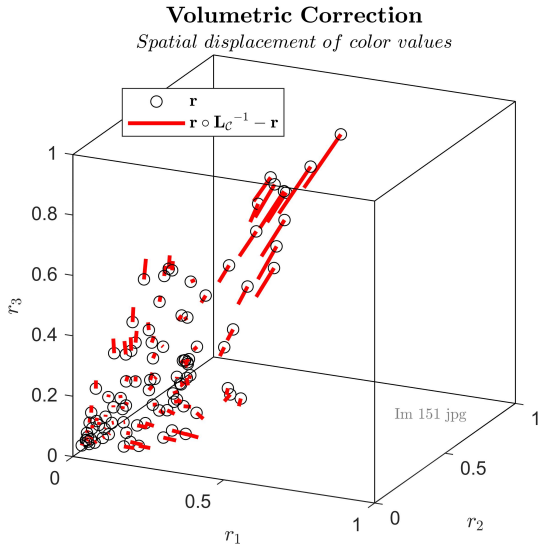

Fig. 12. Estimated volumetric $(\widehat{\mathbf{L}}_\mathcal{C})^{-1}$ displacements of JPEG $\mathbf{r}$. An expected tendency for the gamut mapping to have been contracting the colour space in the saturated colours. As the distribution of scaling between $(\widehat{\mathbf{L}}_\mathcal{C})^{-1}$,$(\widehat{\mathbf{L}}_\mathcal{C})^{-1}$ and $(\widehat{\mathbf{L}}_\mathcal{C})^{-1}$ is unconstrained, a part of said scaling is also visible in the contraction.

estimated gamma curves. The green and the red hide behind the blue.

In figures 4, 5 and 6 the estimated linear responses pertaining to JPEG, reverse engineered JPEG and raw are shown (thin coloured lines with circles). The reverse engineering of JPEG resulted in out-of-gamut colours. Their linear estimates are visibly found. Estimated linear responses can be compared with the raw data (coloured dots) which visibly contains an offset, see inset. The data is normalized to the maximum response in the white patch. The rendered response input for linearisation are shown (thin coloured lines) for reference. Clearly, and perhaps not surprisingly, the estimated linear responses pertaining to JPEG, which incorporates the most challenging rendering manipulations, are the most inaccurate (i.e. noisy). The dotted black line indicates the maximum value of the raw image in the white patch, and thus the exposure level. See Table II for overall image exposure levels.

In figure 6 the tone curves pertaining to raw data are shown. Comparing figure 6 with 7 an example of how the accuracy of the raw-based ground truth estimations compares to the linearly optimized canonical ground truth estimates $\widehat{\rho}^M_{xyz}$, they seem to be almost on par. In the box plot in Figure 13 the averaged accuracy for all images can be found showing the same tendency. So, the linearisation performed on real noisy raw data is on par with the accuracy of theoretical noiseless and physical linear canonical (XYZ proxy) estimated linear values. This points to the impact on accuracy stemming from sensor metamerism. Adding illuminant metamerism will probably only aggravate the problem.

The perceptual displacements between CIELAB values pertaining to the linearised JPEG data (pertaining to Capture 151) and *ground truth* can be seen in Figure 14. Except in the saturated greenish colours, the perceptual differences seem

small. Refer to Figure 14 for overall results.

A. Resulting colour Accuracy

To permit a maximum signal-to-noise influence in the estimates, the evaluation of the linearised results are quantified in absolute response differences between raw and *ground truth* scaled (by gaining) to the maximum *ground truth* channel response of the white patch. Furthermore, these results are calculated in terms of RMSE to enable approximate comparison with other methods, see Table I. The maximum raw channel responses to the white patch (i.e. around 5 to 47% of the total 16 bit exposure range in the camera), are shown in the figures (i.e. Capture 151, 29%) for visualisation and listed in detail in Table II. A rough inspection of the results indicates (a part from two outliers in capture 97 and 303) that only FLT-illuminated (fluorescent light) results seem to have elevated RMSE-values. They stem from exposures of only up to 36%. Such RMSE-values are, however, also found among captures under other illuminants. Since FLT illuminants are quite spiky, small changes in the reflectance can lead to large variations in sensor responses and thus noise sensitivity.

Furthermore, the estimates are related to colour differences in a perceptually uniform representation in the CIELAB-values. Here, the scaling of the estimated and the *ground truth* responses are related directly to the colourimetric values calculated from the srgb values of the white patch in the rendered (JPEG) images.

For all the patches and rgb-types, the predicted RGB and CIELAB values were compared to the known *ground truth* values, and the colour differences ΔRGB and ΔE were computed using the measured reflectances of the colour charts. The CIELAB values use the colourimetric values estimated from the images. The resulting colour differences between *ground truth* and its linear estimate, are shown as box plots in Figures 13 and 14, respectively. The green lines correspond to the median values, the upper and lower ends of the boxes to the quartiles, the whiskers to the full range of the data excluding outliers, and the separate dots to outliers.

For the *ground truth* data, both spectroradiometrically measured reflectances and colour chart spectral reflectances *factory* were tested. The ones measured in the laboratory are supposed to be more accurate than the provided ones, given that they relate to the particular chart that was used, including its state of wear and tear, pigment variations in production and a viewing geometry that mimics the camera recordings pertaining to some of the possible specular reflection. The predicted colour differences and response differences between using the two different *ground truths* were found to be very small. For the raw images, the RMSE was 0.015 for the measured charts, and 0.023 for the factory provided reflectances, with corresponding average ΔE values of 1.86 and 4.26, respectively. For the JPEG images, the RMSE was 0.023 for the measured charts, and 0.033 for the factory provided reflectances, with corresponding average ΔE values of 3.62 and 9.40, respectively. The differences are statistically significant with vanishing p-values according to the Wilcoxon test and with a size in line with the precision in state-of-the-art colour characterisation methods for real data.

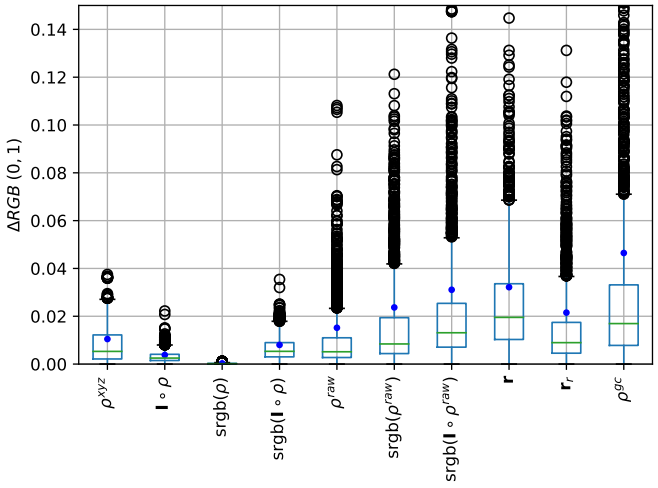

Fig. 13. Box plot of the RGB response difference of all the patches for the different rgb-types for the R measurements. The boxes range from the 25% to the 75% quartile, the green line marks the median and the blue dot is the RMSE. The extent of the whiskers is the full range of the data excluding the outliers, defined as points that are outside 1.5 times the interquartile range outside the box.

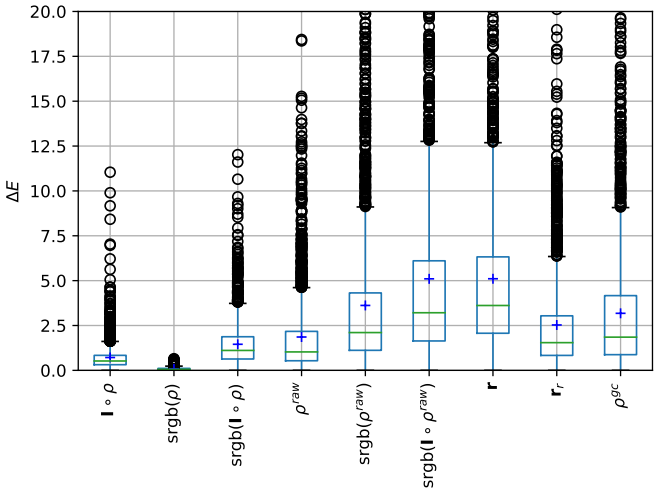

Fig. 14. Box plot of the ΔE colour difference in accordance with Figure 13.

B. Resulting Accuracy of the Irradiance Function

Under the assumption that the offset $\mathcal{O}$ can be corrected in Equation (49) and particularly in $\widehat{\mathbf{L}}_\Gamma$ it makes sense to validate the estimation of the irradiance function on its own. In Figure 9 is shown an estimation of a synthetic irradiance function $\mathbf{I}$ negotiating the exaggerated curvature pertaining to a third order spatial polynomial function with a shading ratio of 3.3. The irradiance function is applied to ρ and so defines rgb-type $\mathrm{srgb}((\mathbf{I} \circ \rho))$.

Without the srgb non-linearity included, the shading compensation method presented in [28] aligns with rgb-type $\mathbf{I} \circ \rho$. Simulated data are generated by applying a synthetic smooth shading on linear rgb-responses from spectrally known colour chart patches. Given the linear responses linearly transformed by a matrix $\mathbf{M}$ and the shaded rgb-responses, the method solves for $\mathbf{M}$ and the shading function. On synthetic data with

a shading ratio of 3.3, they report a mean $\Delta E = 3.0$ and a max $\Delta E = 22.4$. Our method reports $\Delta E = 0.72$ and a max $\Delta E = 11.04$. Note that $\mathbf{M}$ is not estimated by our method.

In Figure 8 the estimated irradiance functions pertaining to $\mathbf{r}$, ρ^{raw} and the in-register camera responses from the GC. When inspecting Figure 8 it can be seen that the linearisation method estimates almost the same irradiance function irrespective of rendition level, measured, raw or re-rendered.

The data results for the *factory* measured reflectances are not presented here in terms of specific results and figures, as that would cover a full parallel set of results. However, the overall influence of the interchanging spectroradiometric with the spectrophotometric measurements of the CG pertaining to the estimates of the irradiance functions is minute. The median error in the $\widehat{\underline{L}}_I^k$ correction factor is on the order of 0.4 %. It varies only insignificantly between the various rgb-types and *ground truth* data based on the two reflectance sets. Although impossible to evaluate precisely as the origin of the *factory* measurement is largely unknown, this alludes to the robustness of employing the additivity principle to data which is not necessarily the exact colour chart and whether the spectral measurement is radiometric or photometric.

Although a thorough investigation of the impact on the accuracy of the estimations is beyond the scope here, we briefly point to possible reasons for shortcomings of the presented method: The additivity matrix $\mathbf{A}$ is subject to the quality of the measurements (e.g. noise) of $\varsigma(\lambda)$ and their overlaps, the patch sampling procedure (e.g. image blurring, edge response contamination, quantisation, and lens glare), in-situ spatially varying SPD of the illuminant and its spectral selectivity, inadequate exposure, reduction of N and spectral dimensionality of reflectance spectra (may be alleviated in the regularisation) and discretization resolution and bandwidth of wavelength interval ω.

VIII. Conclusion

This paper proposes a digital image colour channel value linearisation method of sampled calibration data based on a novel additivity principle. The additivity principle is invariant to metamerism and apart from regularisation parameters, exclusively based on knowledge of the spectral reflectance functions of a colour target. The method linearizes camera responses to one-shot digital images of colour calibration charts from which response values of unknown rendition level are gathered. Only one triplet of sampled responses per patch is necessary. In the process estimates of irradiance geometry, channel non-linearity and non-linear volumetric transformation are calculated. Although, the method does not constitute a radiometric camera calibration, it estimates and performs parallel compensations whilst stepping further to compensate for physical interactions between the camera sensor and the reflectances. It is reiterated that no existing method encompasses the full linearisation from unknown rendition level to estimated colour values that are linearly related to *ground truth*, as defined in Equation 1. This complicates a fully satisfying comparison with state-of-the-art. However, the results indicate that the error in the ground truth estimates are on par with competing methods that only estimate raw and typically need much more data support.

Acknowledgment

This research was funded by VisioTrue ApS, Denmark and the Research Council of Norway over the project 'Individualised colour Vision-based Image Optimisation', grant number 287209. We would like to thank the editor and the anonymous reviewers for constructive comments on the original manuscript.

Author Contributions

Conceptualization, CFA; methodology, CFA and IF; software, CFA; validation, CFA, IF, and JYH; formal analysis, CFA and IF; investigation, CFA and JYH; resources, CFA, IF, and JYH; writing—original draft preparation, CFA, IF and JYH; writing—review and editing, CFA, IF and JYH; supervision, JYH. All authors have read and agreed to the published version of the manuscript.

References

[1] S. He, "Study of linearization for spectral imaging device," in *International Conference on Information Computing and Applications*. Springer, 2012, Conference Proceedings, pp. 125–132.

[2] J. Jiang, D. Liu, J. Gu, and S. Süsstrunk, "What is the space of spectral sensitivity functions for digital color cameras?" in *2013 IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE, 2013, Conference Proceedings, pp. 168–179.

[3] A. Artusi, F. Banterle, T. O. Aydın, D. Panozzo, and O. Sorkine-Hornung, *Image content retargeting: maintaining color, tone, and spatial consistency*. CRC Press, 2016.

[4] C. F. Andersen and D. Connah, "Weighted constrained hue-plane preserving camera characterization," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4329–4339, 2016.

[5] T. Johnson, "Methods for characterizing colour scanners and digital cameras," *Displays*, vol. 16, no. 4, pp. 183–191, 1996.

[6] G. Hong, M. R. Luo, and P. A. Rhodes, "A study of digital camera colorimetric characterization based on polynomial modeling," *Color Research & Application*, vol. 26, no. 1, pp. 76–84, 2001.

[7] G. D. Finlayson, M. Mackiewicz, and A. Hurlbert, "Color correction using root-polynomial regression," *IEEE Transactions on Image Processing*, vol. 24, no. 5, pp. 1460–1470, 2015.

[8] H. Lin, S. J. Kim, S. Süsstrunk, and M. S. Brown, "Revisiting radiometric calibration for color computer vision," in *2011 International Conference on Computer Vision*. IEEE, 2011, Conference Proceedings, pp. 129–136.

[9] S. J. Kim and M. Pollefeys, "Robust radiometric calibration and vignetting correction," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, no. 4, pp. 562–576, 2008.

[10] O. Burggraaff, N. Schmidt, J. Zamorano, K. Pauly, S. Pascual, C. Tapia, E. Spyrakos, and F. Snik, "Standardized spectral and radiometric calibration of consumer cameras," *Optics express*, vol. 27, no. 14, pp. 19 075–19 101, 2019.

[11] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149)*, vol. 1. IEEE, 1999, Conference Proceedings, pp. 374–380.

[12] Y. Xiong, K. Saenko, T. Darrell, and T. Zickler, "From pixels to physics: Probabilistic color de-rendering," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012, Conference Proceedings, pp. 358–365.

[13] A. Chakrabarti, Y. Xiong, B. Sun, T. Darrell, D. Scharstein, T. Zickler, and K. Saenko, "Modeling radiometric uncertainty for vision with tone-mapped color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2185–2198, 2014.

[14] S. J. Kim, H. T. Lin, Z. Lu, S. Süsstrunk, S. Lin, and M. S. Brown, "A new in-camera imaging model for color computer vision and its application," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 12, pp. 2289–2302, 2012.

TABLE II

ILL TYPE: INC: INCANDESCENT, FLT: FLUORESCENT, WLED: WHITE LED, DAY: SUN-SKY-CLOUD AND MIX: LED BASED TUNABLE LIGHT SOURCE. CAPTURE: IMAGE REFERENCE NAMING. *Ground truth* (GTR) VALUES, TO WHICH THE RMSE CALCULATIONS REFER, ARE NORMALISED TO THE MAXIMUM GTR VALUE OF THE WHITE PATCH. RAW AND RENDERED VALUES ARE NORMALISED TO BIT-DEPTH. THEY ARE THUS THE ACTUAL VALUES PERTAINING TO THE EXIF-PARAMETERS. THE ISO VALUE WAS HELD AT 100 TO ENSURE CONTROL OF GAIN. RAW RMSE $R(\rho^{raw})$ REFERS TO ESTIMATED GTR FROM RAW DATA, AND RENDERED RMSE $R(\mathbf{r})$ TO ESTIMATED GTR FROM JPEG DATA. THE CAPTURE NUMBERS REFER TO AN IMAGE NAMING, NOT A SPECIFIC CAPTURE AMONG 1200.

Per capture real image data: White patch responses, EXIF-data and RMSE $R(...)$ results														
Ill_type	Capture	f-stop	Exp.	ρ_r	ρ_g	ρ_b	ρ_r^{raw}	ρ_g^{raw}	ρ_b^{raw}	$\mathbf{r}_r$	$\mathbf{r}_g$	$\mathbf{r}_b$	$R(\rho^{raw})$	$R(\mathbf{r})$
INC	118	5.3	1/13	0.86	1.00	0.38	0.09	0.11	0.04	0.83	0.51	0.30	0.0068	0.0397
INC	301	5.3	1/13	0.78	1.00	0.42	0.15	0.19	0.08	0.78	0.67	0.59	0.0179	0.0350
MIX	302	5.3	1/13	0.56	1.00	0.50	0.15	0.28	0.14	0.82	0.76	0.73	0.0126	0.0203
MIX	18	5.6	1/2	0.68	1.00	0.38	0.23	0.34	0.14	1.00	0.80	0.54	0.0039	0.0312
WLED	9	5.6	1/3	0.46	1.00	0.76	0.20	0.42	0.33	0.99	0.86	0.97	0.0029	0.0178
FLT	303	5.3	1/13	0.51	1.00	0.72	0.24	0.47	0.34	0.96	0.88	0.92	0.0423	0.0479
FLT	1011	5.3	1/13	0.62	1.00	0.46	0.06	0.09	0.05	0.67	0.49	0.33	0.0051	0.0432
FLT	12	5.6	1/3	0.74	1.00	0.37	0.23	0.32	0.12	1.00	0.77	0.51	0.0045	0.0256
FLT	15	5.6	1/3	0.62	1.00	0.49	0.22	0.36	0.19	1.00	0.81	0.68	0.0259	0.0533
FLT	211	5.6	1/60	0.69	1.00	0.42	0.06	0.09	0.04	0.71	0.48	0.29	0.0307	0.0365
FLT	811	5.6	1/15	0.60	1.00	0.53	0.05	0.08	0.05	0.60	0.46	0.37	0.0055	0.0212
DAY	41	13	1/125	0.46	1.00	0.78	0.14	0.29	0.22	0.86	0.77	0.81	0.0036	0.0277
DAY	42	13	1/125	0.46	1.00	0.78	0.13	0.27	0.21	0.85	0.76	0.80	0.0039	0.0404
DAY	45	13	1/125	0.47	1.00	0.76	0.18	0.38	0.28	0.94	0.85	0.90	0.0035	0.0412
DAY	50	13	1/80	0.45	1.00	0.81	0.11	0.23	0.18	0.77	0.72	0.78	0.0133	0.0303
DAY	54	13	1/80	0.45	1.00	0.81	0.12	0.28	0.22	0.80	0.76	0.84	0.0051	0.0157
DAY	70	13	1/160	0.45	1.00	0.79	0.06	0.13	0.10	0.57	0.57	0.63	0.0058	0.0145
DAY	97	13	1/160	0.43	1.00	0.83	0.03	0.07	0.06	0.39	0.42	0.47	0.0214	0.0336
DAY	128	5.3	1/13	0.47	1.00	0.76	0.10	0.22	0.17	0.83	0.70	0.73	0.0044	0.0313
DAY	151	14	1/13	0.44	1.00	0.82	0.12	0.29	0.25	0.82	0.76	0.90	0.0065	0.0183
DAY	611	5.6	1/10	0.39	1.00	0.85	0.03	0.07	0.06	0.35	0.39	0.48	0.0146	0.0192
DAY	1200	5.3	1/13	0.42	1.00	0.87	0.16	0.39	0.35	0.92	0.84	1.00	0.0052	0.0233

[15] O. Burggraaff, M. Werther, E. S. Boss, S. G. Simis, and F. Snik, "Accuracy and reproducibility of above-water radiometry with calibrated smartphone cameras using raw data," *Frontiers in Remote Sensing*, vol. 3, 2022.

[16] R. M. Nguyen and M. S. Brown, "Raw image reconstruction using a self-contained srgb-jpeg image with only 64 kb overhead," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, Conference Proceedings, pp. 1655–1663.

[17] S. Lin, J. Gu, S. Yamazaki, and H.-Y. Shum, "Radiometric calibration from a single image," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 2. IEEE, 2004, Conference Proceedings, pp. II–II.

[18] G. Finlayson, "Method and system for determining parameters of an image processing pipeline of a digital camera," Sep. 21 2017, uS Patent App. 15/528,981.

[19] H. Gong, G. D. Finlayson, and M. M. Darrodi, "Concise radiometric calibration using the power of ranking," *arXiv preprint arXiv:1707.08943*, 2017.

[20] P. E. Debevec and J. Malik, *Recovering high dynamic range radiance maps from photographs*. ACM, 2008, pp. 1–10.

[21] J. E. Garcia, A. G. Dyer, A. D. Greentree, G. Spring, and P. A. Wilksch, "Linearisation of rgb camera responses for quantitative image analysis of visible and uv photography: a comparison of two techniques," *PLoS One*, vol. 8, no. 11, p. e79534, 2013.

[22] S. Lin and L. Zhang, "Determining the radiometric response function from a single grayscale image," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2. IEEE, 2005, Conference Proceedings, pp. 66–73.

[23] J.-Y. Lee, Y. Matsushita, B. Shi, I. S. Kweon, and K. Ikeuchi, "Radiometric calibration by rank minimization," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 144–156, 2012.

[24] M. Díaz and P. Sturm, "Radiometric calibration using photo collections," in *2011 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2011, Conference Proceedings, pp. 1–8.

[25] A. Chakrabarti, D. Scharstein, and T. E. Zickler, "An empirical camera model for internet color vision," in *BMVC*, vol. 1. Citeseer, 2009, Conference Proceedings, p. 4.

[26] D. Cheng, B. Price, S. Cohen, and M. S. Brown, "Beyond white: Ground truth colors for color constancy correction," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, Conference Proceedings, pp. 298–306.

[27] G. D. Finlayson, H. Gong, and R. B. Fisher, "Color homography color correction," in *Color and Imaging Conference*, vol. 2016. Society for Imaging Science and Technology, 2016, Conference Proceedings, pp. 310–314.

[28] G. D. Finlayson, M. Mohammadzadeh Darrodi, and M. Mackiewicz, "The alternating least squares technique for nonuniform intensity color correction," *Color Research & Application*, vol. 40, no. 3, pp. 232–242, 2015.

[29] G. D. Finlayson, B. Schiele, and J. L. Crowley, "Comprehensive colour image normalization," in *European conference on computer vision*. Springer, 1998, Conference Proceedings, pp. 475–490.

[30] P. Bastani and B. Funt, "Simplifying irradiance independent color calibration," in *Color Imaging XIX: Displaying, Processing, Hardcopy, and Applications*, vol. 9015. International Society for Optics and Photonics, 2014, Conference Proceedings, p. 90150N.

[31] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, 1985.

[32] J. Morovič, *Color gamut mapping*. John Wiley & Sons, 2008, vol. 10.

[33] J. Giesen, E. Schuberth, K. Simon, P. Zolliker, and O. Zweifel, "Image-dependent gamut mapping as optimization problem," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2401–2410, 2007.

[34] E. Million, "The hadamard product," *Course Notes*, vol. 3, no. 6, 2007.

[35] M. Kang, U. Yang, and K. Sohn, "Spectral sensitivity estimation for emccd camera," *Electronics letters*, vol. 47, no. 25, pp. 1369–1370, 2011.

Casper Find Andersen holds an M.Sc. in Mechanical Engineering and Computational Fluid Dynamics from the Danish Technical University and a PhD in Computer Science from the University of East Anglia, UK. He is currently a researcher with the NTNU – the Norwegian University of Science and Technology, on the Norwegian Research Counsel Project of Individual Colour Vision-based Image Optimisation (ICVIO).

Ivar Farup received an M.Sc. in technical physics from the Norwegian Institute of Technology in 1994 and a PhD in applied mathematics from the University of Oslo, Norway in 2000. He is currently a professor of computer science at NTNU – the Norwegian University of Science and Technology, mainly focusing on colour science and image processing.

Jon Yngve Hardeberg Jon Y. Hardeberg received his sivilingeniør (M.Sc.) degree in signal processing from the Norwegian Institute of Technology (now NTNU) in Trondheim, Norway in 1995, and his PhD from Ecole Nationale Supérieure des Télécommunications in Paris, France in 1999. He is now Professor of Colour Imaging at NTNU, Department of Computer Science, The Norwegian Colour and Visual Computing Laboratory, Gjøvik, Norway.