Rank-Based Radiometric Calibration

Han Gong [1]*, Graham D. Finlayson* [1]*, Maryam M. Darrodi* [1]*, Robert B. Fisher* [2]

[1] *University of East Anglia, UK*

[2] *University of Edinburgh, UK*

Abstract

Raw images are more useful than JPEG images for machine vision algorithms and professional photographers because raw images preserve a linear relation between pixel values and the light measured from the scene. A camera is radiometrically calibrated if there is a computational model which can predict how the raw image is mapped to the corresponding rendered image (e.g. JPEGs) and vice versa. Our method makes use of the observation that the rank order of pixel values are mostly preserved post color correction. We show that this observation is the key for getting a compact and robust radiometric calibration model. Since our method requires fewer variables, it can be solved for using less calibration data. An additional advantage is that we can derive the camera pipeline from a single pair of raw-JPEG images. Experiments demonstrate that our method delivers state-of-the-art results (especially for the most interesting conversion from JPEG to raw).

Introduction

Many computer vision algorithms (*e.g.* photometric stereo [21], photometric invariants [12], shadow removal [16, 15], and color constancy [2]) rely on the assumption that the captured RGBs in images are linearly related to the actual scene radiance. However, mostly, the actual output of a digital camera imaging pipeline is necessarily non-linear in order to produce perceptually-pleasing photos as opposed to their physically-meaningful counterparts. In this paper, we present a compact rank-based radiometric calibration method which solves for the bi-directional mappings between the camera's raw responses and the rendered RGBs produced by digital cameras.

There is prior art in this field which models the pipeline with a large number of parameters (up to several thousand [5]) which both means a large corpus of data is required to uncover the pipeline and that there is at least tacitly the premise that the underlying pipeline is quite complex. The key insight in our approach is that post-color correction (a 3×3 matrix correction) the linear corrected raw RGBs are to the greatest extent in the same rank order as the final rendered RGBs. Based on this insight, we develop a compact rank-based radiometric calibration model that models the camera pipeline with many fewer parameters and concomitantly needs much less training data.

In Fig. 1, we illustrate a typical image reproduction pipeline which is representative of many cameras [18]. An exemplar raw image, Fig. 1a, is mapped by a 3×3 color correction matrix to give a color corrected image (Fig. 1b). The color correction matrix implements several processing steps (*e.g.* illumination correction [23, 4], display RGB mapping [1], and color preference adjustments [23]). It is well-known that a display device cannot show all captured image colors and some RGBs will fall outside

the RGB cube after mapping (*e.g.* the pixels marked in light purple in Fig. 1b). Gamut mapping is therefore required, *e.g.* [18, 5, 13], to move the colors back inside the cube as shown in Fig. 1c. Finally, the gamut mapped image is tone mapped to arrive at the final rendered output [23, 4, 18] shown in Fig. 1d. Tone mapping accounts for the display non-linearity [1], dynamic range compression and some aspects of preference [24].

In general, the camera color processing pipeline can be written as Eqn. 1.
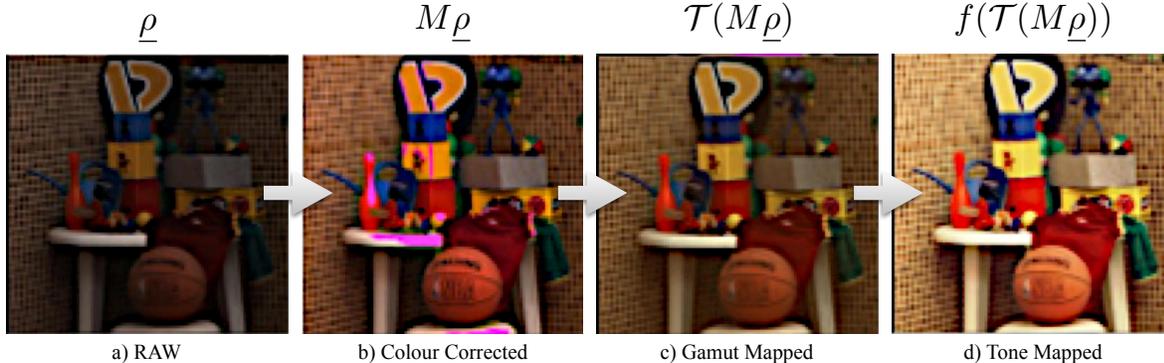
$$\underline{P} = \underbrace{f(\Gamma(M\underline{\rho}))}_{(1a)} = \underbrace{\Gamma(f(M\underline{\rho}))}_{(1b)} \approx \underbrace{\text{LUT}(\underline{\rho})}_{(1c)} \quad (1)$$

where $\underline{\rho}$ denotes a camera raw and $\underline{P}$ refers to its rendered RGB counterpart. Respectively, the 3×3 correction matrix, gamut mapping and tone mapping are denoted by the matrix M and the functions $\Gamma()$ and $f()$. The function $f()$ can implement a single or three per-channel tone curves. Since gamut mapping only implements a small change in comparison with color and tone mapping steps, the order of gamut mapping and tone mapping may be switched (Eqn. 1b & c), a property that we exploit in this paper. Equally, we can also merge three processing steps into one and directly solve for a 3-D LUT (Look-Up-Table) that maps raw to rendered counterparts. This LUT function is denoted LUT() [19] in Eqn. 1c. Readers may refer to the top row of Fig. 1 to link each mathematical function to our example processed image.

In radiometric calibration, given a set of $\underline{\rho}$ and $\underline{P}$, we solve for the parametrised pipeline parts (*e.g.* M, $\Gamma()$, $f()$ and LUT()). A disadvantage of the current best performing methods is that a great deal of data may be required to fit their assumed models. In Eqns. 1a and 1b, the gamut mapping step could be modeled by 1000s of Radial Basis functions [18, 19, 5] and in Eqn. 1c, the deployed LUT could also have several thousand control points.

Our proposed method exploits the simple observation [8] that, assuming the gamut mapping step slightly changes image colors and the tone curves are always monotonically increasing, we expect mostly the rank ordering of the rendered $\underline{P}$ to be the same as $\underline{\rho}$ multiplied by the correction matrix M. Suppose that two rendered (JPEG) responses – in the 1st red color channel – are denoted P_1^a and P_1^b and that $P_1^a > P_1^b$. The rank order of two corresponding raw red channel measurements post color correction is written as $M_1\underline{\rho}^a > M_1\underline{\rho}^b$ (where M_1 denotes the first row of M and $\underline{\rho}^a$ and $\underline{\rho}^b$ are a pair of raw RGBs). This implies that $M_1(\underline{\rho}^a - \underline{\rho}^b) > 0$ which defines a half-space constraint.

The row vector M_1 can be considered as a point in 3-space and this inequality (ranking constraint) forces the point to be located in only one half of 3-space. Because we have multiple pixels, each pair of pixels (2 raw and 2 JPEG RGBs) generates a half

$$\underline{\rho} \qquad M\underline{\rho} \qquad \mathcal{T}(M\underline{\rho}) \qquad f(\mathcal{T}(M\underline{\rho}))$$

a) RAW   b) Colour Corrected   c) Gamut Mapped   d) Tone Mapped

Figure 1. *a) a raw input image is color corrected to give image b). Non-displayable colors are highlighted in purple pseudo color. Gamut mapping, in step c), brings colors within gamut. Finally, in d), a tone mapping step results in the final rendered image. The image is taken from [10].*

space constraint and intersecting all these constraints delimits the region in which M_1 must lie. Our experiments demonstrates that a small numbers of patches suffices to estimate M accurately.

Once we have M we then find the best rank preserving tone curves $f()$. At this stage, only using M and $f()$ we have a good approximation of the pipeline. Indeed, we argue that our construction of M and $f()$ also incorporates, to a first order, gamut mapping. Now we adopt (Eqn 1b) and find a 125-parameter per channel LUT to reduce any remaining errors due to gamut mapping (higher order terms).

In Section 2, we review radiometric calibration paying special attention to methods which adopt Eqns. 1. Rank-based radiometric calibration is described in Section 3 and is shown to provide leading performance on a public dataset in Section 4. An application of one-shot radiometric calibration without the access to raw is shown in Section 5. The paper concludes in Section 6.

Related Work

Using the pipeline form of Eqn 1b, Chakrabarti *et al.* [5] first solve for M and $f()$ in iteration and then solve directly for $\Gamma()$. In their approach, $f()$ is constrained to be a 7th order increasing polynomial. They model $\Gamma()$ by the radial basis function (RBF) method of [18] where several thousands of RBFs are potentially used. A restriction of the above calibration is presented in [4] where the gamut mapping $\Gamma()$ is ignored. This less general model works tolerably well on many real pairs of raw and rendered images and this is a point we will return to later in this paper. In either version ([5] or [4]), the coupled nature of the minimization indicates that a global minimum is not guaranteed. Therefore, a random start point search is implemented to find a better set of parameters.

Kim *et al.* [18] solve for the pipeline in the form of Eqn. 1a and makes additional assumptions to decouple the optimization. They assume that images of the same scene are captured with respect to two or more exposures and their $\Gamma()$ is a multi-thousand set of RBFs. Regarding solving for $f()$, Debevec *et al.* [6] showed how relating corresponding pixels under known exposure differences suffices to solve for $f()$ (assuming there is no gamut mapping step). Importantly, in [18], it was argued that for the set of desaturated pixels (*i.e.* raws far from the RGB cube boundary), the gamut mapping step has little or no effect and can be ignored.

Relative to this assumption, $f()$ can be solved using the Debevec method. Given $f()$ then the color correction matrix M can be found (again using desaturated pixels).

We point out that for most off-shelf capture devices (*e.g.* for most mobile phones), manual exposure control is usually unavailable and the requirement of multiple exposures is impractical. We also note that, in [18], the adopted gamut mapping RBF network requires a large number of parameters and thus a large corpus of data [18, 5].

In [19], it was shown that it is possible to ignore the underlying structure of the color processing pipeline and directly solve for the best 3-D surjective function – implemented as a LUT that maps the raws to rendered RGBs (Eqn. 1c). Finally, in [20], a method is presented for solving for $f()$ by examining the edge distribution in an image. This method has the advantage that the method works for a single image (without multiple exposures) but the method is sensitive to processing steps such as image sharpening which is used extensively in mobile phone image processing.

The Rank-Based Method

In this paper, we are interested in calibrating in the most general circumstances when the amount of training data is modest and there is only a single calibration image. Assuming that the rank order of intensities are almost preserved after the camera processing steps, we present a method that solves for an accurate rank-preserving camera pipeline model. This paper extends our previously published rank-based method [17] with more details and experiments.

To make the rank-based method work we need to assume that the gamut mapping step $\Gamma()$ only changes color slightly. In fact our assumption is more nuanced. We assume that – to a first order – gamut mapping can mostly be implemented as an affine transform and that this affine transform can be folded into the color correction matrix M and the monotonically increasing tone mapping functions $f()$.

Gamut Mapping as An Affine Transform

After color correction, some colors are mapped outside the color cube and become non-displayable. To address this, gamut mapping in Eqn. 1b is applied. A Taylor expansion to model $\Gamma()$

around a point $\underline{a}$ inside the gamut is used:

$$\Gamma(M\underline{\rho}) \approx \Gamma(\underline{a}) + J(\underline{a})(M\underline{\rho} - \underline{a}) \tag{2}$$

where J is the 3×3 Jacobian (matrix of derivatives of Γ). Not only does Eqn. 2 show that, to a first approximation, gamut mapping is an affine transform it is also one of the gamut mapping algorithms proposed in [13].

$$\min_{T,\underline{o}} \Sigma_i ||TM\underline{\rho}_i + \underline{o} - M\underline{\rho}_i||^2 \ s.t. \ \underline{0} \leq TM\underline{\rho}_i + \underline{o} \leq \underline{1} \tag{3}$$

In Eqn. 3, T and $\underline{o}$ are respectively a 3×3 matrix and 3×1 off-set vector defining the affine gamut mapping algorithm, i is the index of an input RGB vector. The 3-vectors of 0s and 1s are denoted $\underline{0}$ and $\underline{1}$. Eqn. 3 is solved directly by Quadratic Programming [14]. The gamut mapping shown in Fig. 1c is the result of solving Eqn. 3.

Here, we make two important remarks about affine gamut mapping: 1) Gamut mapping and color correction combined can be represented by the single affine transform: 3×3 matrix TM and offset $\underline{o}$; 2) It follows that the rank-based method presented in the next section will actually solve for TM. The offset term can be incorporated directly in $f()$.

Our hypothesis is that the part of gamut mapping that is not described by an affine transform will be small and the remaining error can be modeled with a function that has fewer parameters (100s in contrast to the prior art 1000s).

Rank-Based Estimation for Color Correction

Denote the kth row of M as M_k. We assume that given two color corrected raws, $M_k\underline{\rho}^a$ and $M_k\underline{\rho}^b$, that the rank order is the same as for the corresponding rendered RGBs:

$$P_k^a > P_k^b \ \Rightarrow \ M_k\underline{\rho}^a > M_k\underline{\rho}^b \ \Rightarrow \ M_k(\underline{\rho}^a - \underline{\rho}^b) > 0 \tag{4}$$

Defining the difference vector $\underline{d}^j = \underline{\rho}^a - \underline{\rho}^b$:

$$M_k\underline{d}^j > 0 \tag{5}$$

where it is understood the superscript j denotes the difference vector from the jth of $\binom{n}{2}$ pairs of image pixel values (n is the total number of image pixels). Suppose that we have a vector M_k where Eqn. 5 holds, then the inequality cannot be true for $-M_k$. That is Eqn. 5 defines a half plane constraint [8, 3]. The vector $\underline{d}^j$ is perpendicular to the half-plane: any M_k less than 90 degrees to $\underline{d}^j$ is a possible solution. Given multiple difference vectors then we have multiple half-plane constraints which taken together delimit a region in 3-space where M_k must lie. Denoting the half-plane as $\mathscr{H}(\underline{d}^j)$, M_k must satisfy:

$$M_k \in \bigcap_j \mathscr{H}(\underline{d}^j) \tag{6}$$

The intersection in Eqn. 6 defines an unbounded cone, anchored at the origin, in 3-D space. Clearly, if M_k is in the intersection region defined by Eqn. 6 then αM_k (where α is a positive scalar) is another solution. Using ranking we solve for each rows of M up to an unknown scalar multiplier.

Let us visualize the computation of M_k using ranking. Without loss of generality let us assume that $M_{k,3} = 1$. We rewrite Eqn. 5 as

$$M_{k,1}d_1^j + M_{k,2}d_2^j + d_3^j > 0 \tag{7}$$

If $[a\,b\,c]$ is a solution to Eqn. 6, then $[a/c\,b/c\,c/c]$ for Eqn. 7 is also true since $M_{k,1} = a/c$ and $M_{k,2} = b/c$. Solutions for $[M_{k,1}, M_{k,2}]$ lie on one side of the line, *i.e.* the 3-D half-space constraints maps directly to a 2-D half-plane constraint. Or, if we consider the whole set of intersections, the cone in 3-D, defined by Eqn. 6, maps to a 2D convex region [7]. Denoting half-planes as $\mathscr{P}(\underline{d}^j)$ we, equivalently, solve for

$$[M_{k,1}, M_{k,2}] \in \bigcap_j \mathscr{P}(\underline{d}^j) \tag{8}$$

The intersection problem of Eqn. 8 is easily visualized. In Fig. 2a we show the intersection of 4 half plane constraints and indicate the solution set where M_k must lie.

We solve for M_k one sensor channel at a time. Due to noise or small deviations in real camera data, it is likely that no common intersection can be found that satisfies every half-plane's constraint. To solve this problem, we generate 100,000 unit length vectors that are uniformly distributed on the surface of the unit sphere [22], which is visualized in Fig. 2b. With respect to this sampling, the furthest distance between any point and its nearest neighbor is less than 1.15 degrees. Therefore, the orientation of the rows of M are found to this accuracy. For each point on the sphere (*i.e.* a possible row of M_k), we count how many half-space constraints are satisfied. The point on the unit sphere that has the largest number of satisfying half-plane constraints – or the median of multiple points if there is a tie – defines M_k.

To maintain a reasonable computational cost, we have to be careful not to generate too many half planes. We simply select 200 random RGB-JPEG pairs for estimating M. Since it is not guaranteed that these 200 random pairs are the optimum selections, we practically generate 25 random sets of 200 RGB-JPEG pairs and select the "best luck" set which gives the lowest training error. Note that the other pipeline components introduced later are also applied to evaluate the training error.

Overall, we find the M that places all the corresponding raw and rendered image RGBs in the most similar rank order. That is, if we plot the mapped raw red responses, for example, against the rendered red JPEG corresponding values then the graph should be a monotonically increasing function. How well a monotonically increasing function fits our data can be used to judge the efficacy of each M.

Rank-Preserving Optimization of Tone Curves

We now solve for the optimal per-channel tone curves which map color corrected raws to corresponding rendered RGBs. Let us denote the ith color corrected raw and rendered RGB pixel pairs for the kth channel as $(M_k\rho_{k,i}, P_{k,i})$. Then, the kth-channel rank-preserving tone curve $f_k()$ is optimized as a 7th order monotonic and smooth polynomial function as follows:

$$\min_{f_k()} \Sigma_i ||f_k(M_k\rho_{k,i}) - P_{k,i}||^2 + \lambda \int_t ||f_k''(t)||^2 dt \ \text{s.t.} \ f_k'() \geq 0. \tag{9}$$

where the first term is for data fitness, the second term is for curve smoothness and λ is a small weight (*e.g.* 10^{-5}). This polynomial fitting is solved by Quadratic Programming [14]. Note that these 7th order polynomials already include the Affine gamut mapping offset parameters described previously. In this paper, we further denote the combination of all 3-channel mappings $f_{1-3}()$ as $f()$.
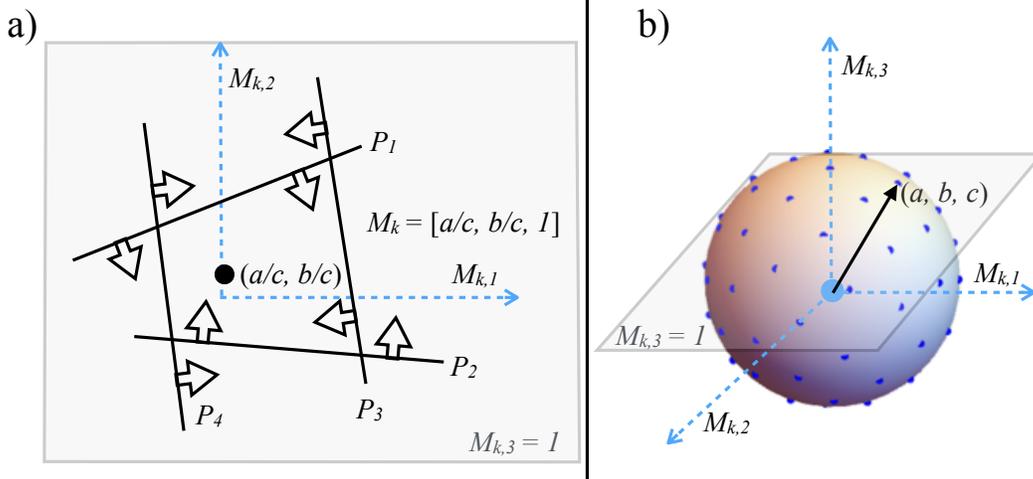
Figure 2. a) The region where 4 half-plane constraints intersect delimit the region where $[M_{k,1}, M_{k,2}]$ must lie where the black point is a feasible solution. b) On an unit sphere, each vector represented by the origin and a blue surface point is a probe for a possible solution (e.g. the black arrow). All 3-D points and constraints are projected to a 2D plane $M_{k,3} = 1$.

Gamut Correction Step

As argued previously, we propose that $f(M\underline{\rho})$ has the expressive power to implement color correction, tone correction and gamut mapping (to the first order in a Taylor expansion). However, we wish to add a further gamut mapping step for the higher order terms. But, since our hypothesis is that much of the gamut mapping will have been accounted for we are going to adopt a simple small parameter solution. Further, this additional correction is going to be carried out at the end of the process, we adopt Eqn. 1b. Specifically, we find a $5 \times 5 \times 5$ LUT by using lattice regression [11] that minimizes $\min_{LUT()} \Sigma_i ||\text{LUT}(g(f(M\underline{\rho}_i))) - \underline{P}_i||^2$ where $g()$ is a non-linear function that stretches highlights. We found empirically there was an advantage in deploying more LUT resolution in the highlight region where gamut mapping is created. We implemented this not by changing the sampling structure of the LUT (which is uniform) but by stretching our data, by applying the function $g()$ shown in Figure 3. The function $g()$ is fixed for all our experiments.
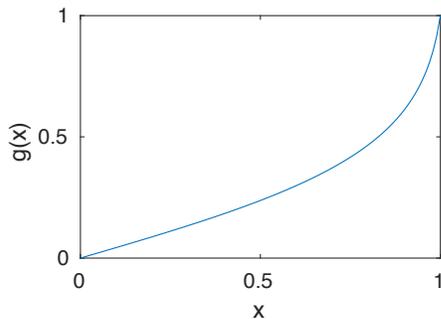
Figure 3. Plot of a non-linear function $g(x)$ where $g(0) = 0$ and $g(1) = 1$.

Rank-Based Recovery of Raw

Suppose we wish to map rendered RGBs to raws. Using the previously described method, M has already been solved in the RAW-to-JPEG forward estimation phrase. Now, in a least-squares optimal way, we use the same polynomial fitting method (Eqn. 9) to find f^{-1} by optimizing $\min_{f^{-1}()} \Sigma_i ||f^{-1}(\underline{P}_i) - M\underline{\rho}_i||$. Finally, we solve for the backward LUT() by optimizing $\min_{LUT()} \Sigma_i ||\text{LUT}(g(M^{-1}f^{-1}(\underline{P}_i))) - \underline{\rho}_i||$ where the LUT is fitted by a $5 \times 5 \times 5$ lattice regression [11].

Parameter Counting

Assuming we solve for 3 independent tone curves then our method requires 9 (for M) + 8×3 (for $f()$) + 125×3 (for $\Gamma()$) = 408 parameters which is significantly less (even an order of magnitude less) than [5, 18, 19].

Evaluation

Our evaluation is based on two challenging datasets from Harvard [5] and NUS [18]. The Harvard dataset [5] contains the RAW/JPEG intensity pairs of 140 color checker patches viewed under multiple viewing conditions. The color chart is captured by 8 cameras and under 16 illuminants across many different exposures. Compared with the Harvard dataset, the NUS dataset contains large data captured with a 24-patch color checker and 31 camera sensors though its capture conditions (*i.e.* light and exposure) are relatively limited.

We carried out the same experiment described in [5, 18]. We are interested in validating whether our method, with much reduced number of parameters can produce, similar or even better results compared with the state-of-the-art [5]. We evaluate both RAW-to-JPEG and JPEG-to-RAW. The Harvard dataset [5] captures a sort of "worst-case" viewing conditions. Normally, when we capture a picture there is a single prevailing illuminant color. In the Harvard dataset, all camera processing parameters are turned off and then the same reflectances are viewed under multiple colored lights. As Forsyth observed [9], the reddest red

camera response cannot be observed under a blue light. And, then he exploited this observation to solve for the color of the light. In real imaging conditions, the greenest green and the bluest blue do not typically appear at the same time. A pipeline that suffices for the combinations of all lights and all surfaces is unlikely needed. This means the prior art pipelines are probably more complex than they need to be. As described in [5], for each camera, we estimate the parameters of a calibration model using different subsets of the available RAW-JPEG pairs. For each subset and a selected camera, the root mean-squared error (RMSE) between the prediction and ground truth is validated by using all available RAW-JPEG pairs.

Figure 4 (top half) shows the raw-to-JPEG mapping error plot (where pixel intensities are coded as integers in the interval $[0, 255]$. In both forward and backward tests, our RB (Rank-Based) method [17] is significantly better than the independent polynomial method (IndPoly) [4]. IndPoly is a simple model which only contains 3 per-channel tone mapping (or linearization) curves and a 3×3 gamut mapping matrix. We also found that our RB's forward errors are close to the results of the state-of-the-art ProbRC [5], especially for the condition of less than 3 illuminants which are more likely to occur in the real world. Evidently, for the many illuminant case the prior art has a small advantage. Remembering that JPEGs are coded as integers in [0,255] the RMSE is typically 1 or less. Practically, when the "fits" are viewed visually (by looking at images) it is hard to see the difference. For computer vision, we are more interested in the performance of JPEG-to-RAW mapping which is shown in Figure 4 (bottom half). In ProbRC [5], a probabilistic framework for mapping rendered RGB to raw was presented. Here we take their mean estimates as the most likely raw predictions. We found that our methods generally reduce the errors of [5] by $\sim 34\%$.

We also verify the results with a wider range of camera sensors using the NUS dataset [18] by a 4-fold cross validation. Since IndPoly [4] performs significantly worse, we omit it in our NUS dataset test. The data in [18] contain an uneven number of capture modes for each camera sensors. Therefore, in Figure 5, we show the overall performance categorized by camera sensor. We found that the results in general shows a similar trend for the forward raw-to-JPEG mapping. For the backward JPEG-to-raw mapping our RB has a significant advantage over [5] as our backward mapping errors are only $\sim 10\%$ of [5].

The reader might be interested in why our simple method seems to work so well going from rendered to raw (better than [5]) but not quite as well as the prior art in the forward direction (albeit visually almost indistinguishable). Our hypothesis here is that the LUT in the forward direction is applied post the tone curve. This curve (at least for dark values) has a very high slope and, consequently, the coarsely quantized $5 \times 5 \times 5$ per-channel LUT cannot capture gamut mapping well. Yet, in the reverse direction (JPEG to RAW) the LUT is applied in linear raw where a coarse uniform quantization is more justified. The full calibration results maybe found in our supplementary materials.

Calibration Stability of Random Sample Selection

We are also interested in how the randomly selected samples for estimating the 3×3 color correction matrix affects the calibration results. We select the most common "1 illuminant + 10
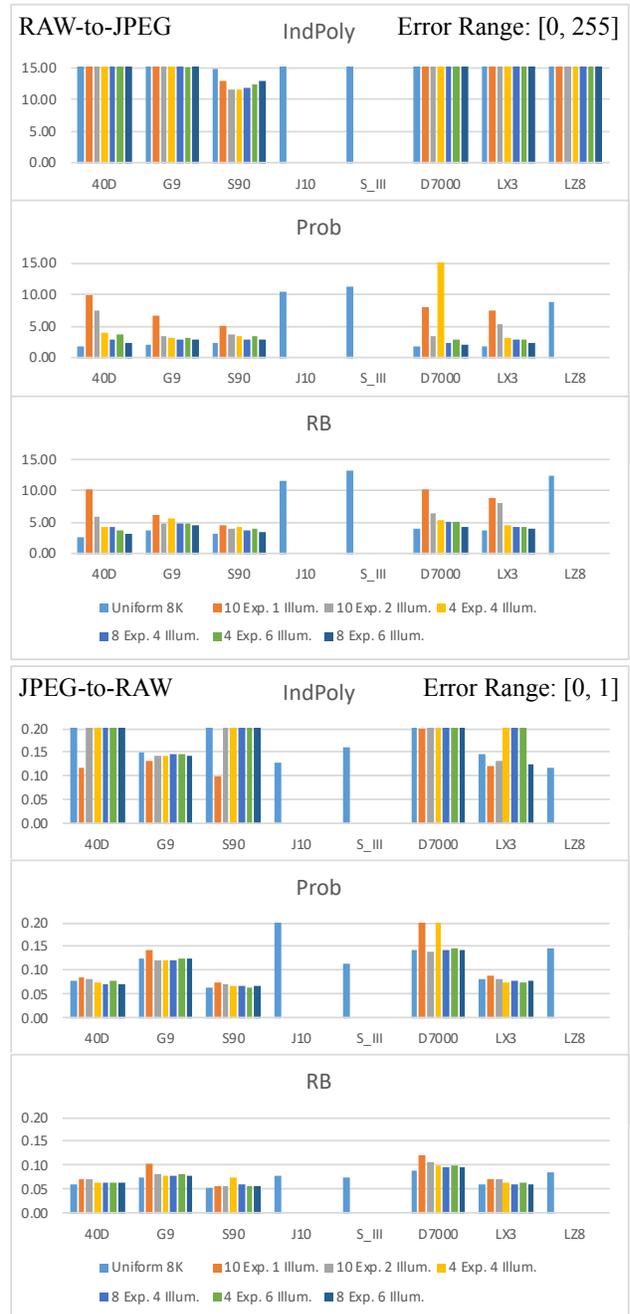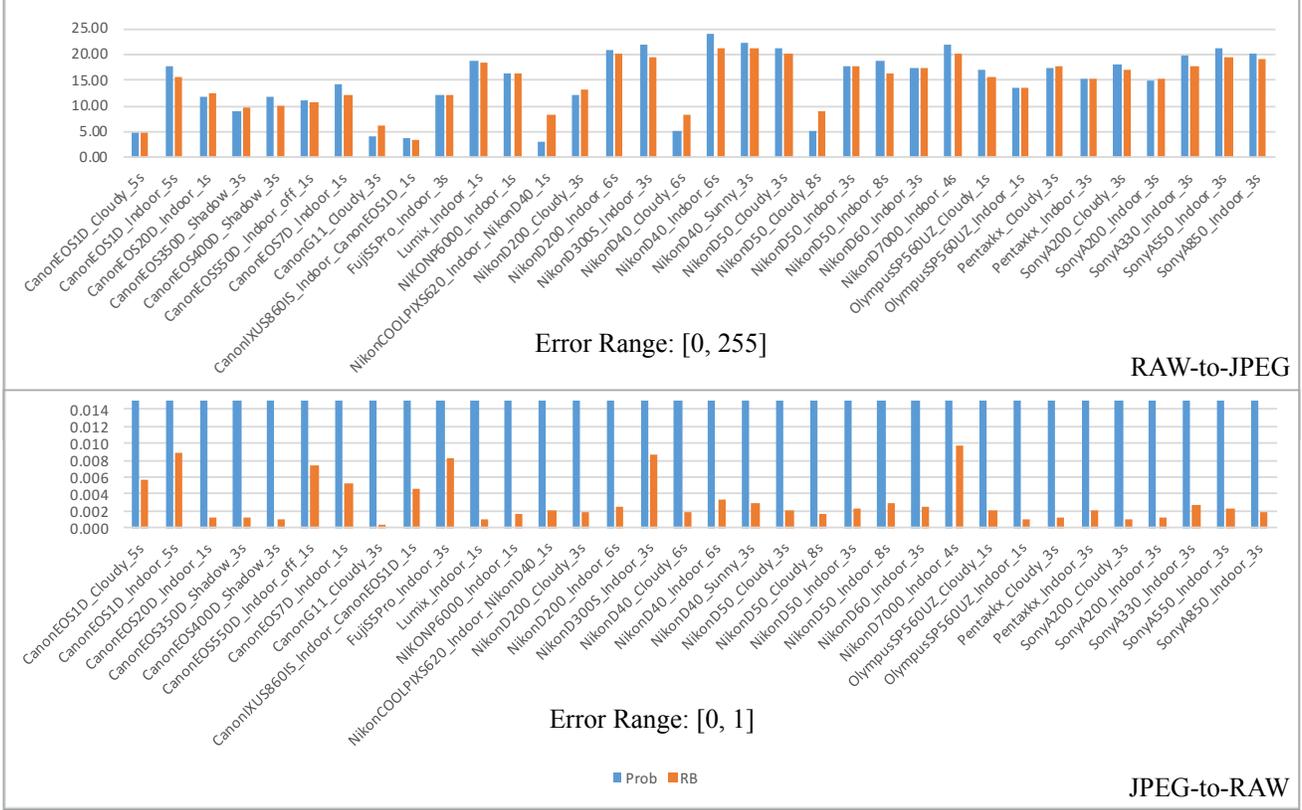

Figure 4. *Test RMSE bar chart of the Harvard dataset [5] for IndPoly [4], Prob [5], and our RB [17]. The figure shows RMSE between ground truth and prediction for bidirectional raw and JPEG mappings. "Exp." and "Illu." are respectively short for "Exposure" and "Illuminant". Each bin color indicates a capture condition. The horizontal and vertical axes indicate camera sensor and calibration error respectively. The displayed forward and backward errors are clipped at 15.0 and 0.20 respectively.*

exposures" data from the Harvard dataset and repeat the calibration experiment for 50 times. This calibration stability test results are shown in Table 1. As shown, the variation for forward estima-

Figure 5. *Test RMSE bar chart of the NUS dataset [18] for Prob [5] and our RB [17]. The figure shows RMSE between ground truth and prediction for bidirectional raw and JPEG mappings. "Exp." and "Illu." are respectively short for "Exposure" and "Illuminant". Each bin color indicates a capture condition. The horizontal and vertical axes indicate camera sensor model code and calibration error respectively. The displayed forward and backward errors are clipped at 25.0 and 0.015 respectively.*

a) RAW-to-JPEG	40D	G9	S90	D7000	LX3
mean	10.52	8.85	4.92	14.19	9.74
std	1.65	2.54	0.41	3.14	1.00
b) JPEG-to-RAW	40D	G9	S90	D7000	LX3
mean	0.073	0.106	0.058	0.124	0.071
std	0.002	0.003	0.001	0.004	0.001

Table 1: RMSE of our rank-based method between ground truth and prediction for bidirectional RAW and JPEG conversions. The results are based on 50 repeated tests.

tion is about 15% of the mean value while that of the backward estimation can be omitted.

Calibration Error Distribution

To better understand how well the proposed algorithm performs w.r.t. input hue and saturation, we visualize the distributions of forward and backward mapping errors (RMSE) over the hue and saturation gamut (HSV color space [25]) which is shown in Fig. 6. The distributions are generated for the common capture condition set – "1 illuminant + 10 exposures" – as the examples. Specifically, the complete distribution data is interpolated for a 400×400 uniform grid. This grid (as an image) is then filtered by a 41×41 Gaussian kernel with a 20 standard deviation.

In addition, in Fig. 7, we also plot the RMSE distributions for the same capture condition – "1 illuminant + 10 exposures" – w.r.t. lightness (i.e. value channel of the HSV color space). Similarly, the complete distribution data are interpolated for a 400-tick uniform 1-D space. This interpolated 1-D space data is then filtered by a 41×1 Gaussian kernel with a 20 standard deviation.

Overall, higher forward errors are observed when hue and saturation are both high. Value (of the HSV color space) does not seem to have a great impact for the forward errors although lower forward errors are found near both clipping boundaries of 0 and 1. Higher backward errors are found when value is high while hue and saturation have less impact on the backward error.

Calibration with Small Numbers of Parameters

We wished to visually validate our claim that we can calibrate with few parameters. We took 4 RAW+JPEG pairs (for different cameras) from [4]. We then uniformly selected 140 corresponding pixels from the RAW and JPEG. We solved for all the 408 parameters in our rank-based method. We then applied our model to the rest of the image. The result of this experiment for 4 images (JPEG-to-RAW) is shown in Fig. 8.
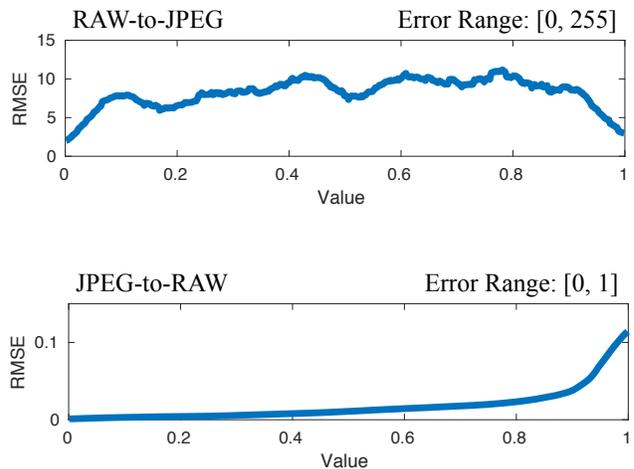
RAW-to-JPEG   Error Range: [0, 255]

JPEG-to-RAW   Error Range: [0, 1]
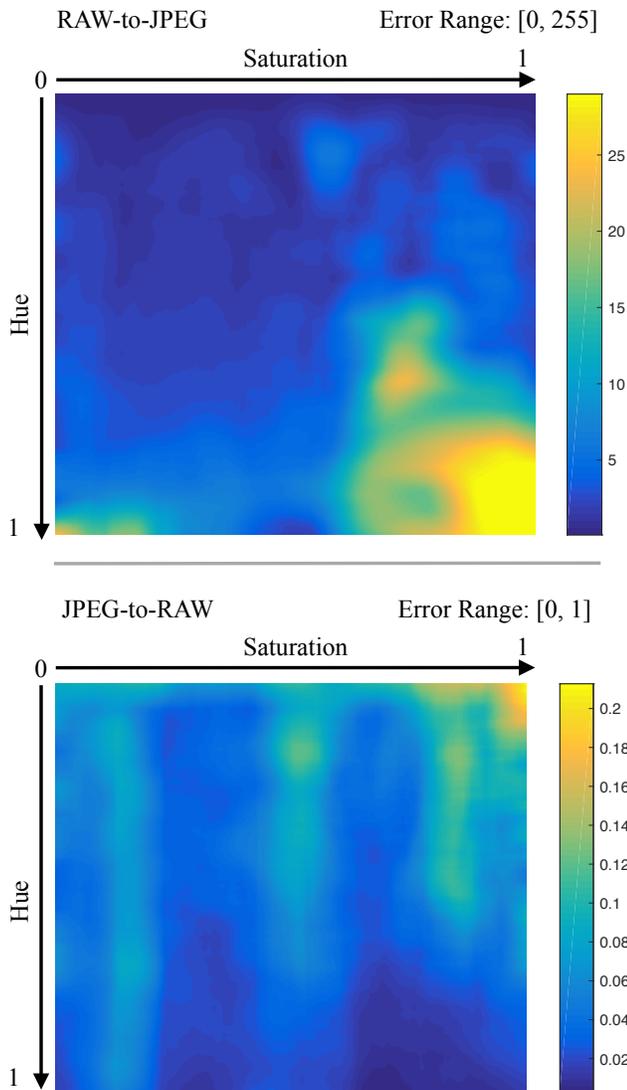
Figure 6. *Distributions of forward and backward mapping RMSE of our method over the input hue and saturation gamut. It is tested with the "1 illuminant + 10 exposures" set from the Harvard dataset [5].*

Figure 7. *Distributions of forward and backward mapping RMSE of our method w.r.t input lightness (value of the HSV color space). It is tested with the "10 Exps. and 1 illum." set from the Harvard dataset [5].*

Conclusion

In this paper we have shown how the rank order of image responses is a powerful tool for solving for the individual steps in a camera processing pipeline (color correction, gamut and tone mapping). A simple ranking argument, relating color corrected raws to corresponding rendered RGBs suffices to solve for the color correction matrix. Then, the rank-preserving tone map is found and, finally, a simple gamut correction step is derived. Compared with the prior art, our rank-based method requires the fewest assumptions and delivers state-of-the-art radiometric calibration results. Experiments also show that excellent calibration is possible given a single image exposure and limited color diversity (*e.g.* a color chart).

References

[1] M. Anderson, R. Motta, S. Chandrasekar, and M. Stokes. Proposal for a standard default color space for the internet:srgb. In *IS&T and SID's 4th Color And Imaging Conference*, pages 238–245. 1995.

[2] Kobus Barnard, Lindsay Martin, Adam Coath, and Brian V. Funt. A comparison of computational color constancy algorithms. ii. experiments with image data. *IEEE Transactions on Image Processing*, 11(9):985–996, 2002.

[3] Mark de Berg, Otfried Cheong, Marc van Kreveld, and Mark Overmars. *Computational Geometry: Algorithms and Applications*. Springer-Verlag TELOS, 3rd ed. edition, 2008.

[4] Ayan Chakrabarti, Daniel Scharstein, and Todd Zickler. An empirical camera model for internet color vision. In *British machine Vision Conference*, volume 1, page 4, 2009.

[5] Ayan Chakrabarti, Ying Xiong, Baochen Sun, Trevor Darrell, Daniel Scharstein, Todd Zickler, and Kate Saenko. Modeling radiometric uncertainty for vision with tone-mapped color images. *Transactions on pattern analysis and machine intelligence*, 36(11):2185–2198, 2014.

[6] Paul Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH*, pages 369–378, 1997.

[7] G.D. Finlayson. Color in perspective. *Transactions on Pattern analysis and Machine Intelligence*, pages 1034–1038, October 1996.

[8] Graham Finlayson, Maryam Mohammadzadeh Darrodi, and Michal Mackiewicz. Rank-based camera spectral sensitivity estimation. *Journal of the Optical Society of America A*, 33(4):589–599, 2016.

[9] David A Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–35, 1990.

[10] David H. Foster, Kinjiro Amano, Sérgio M. C. Nascimento, and Michael J. Foster. Frequency of metamerism in natural scenes. *Journal of the Optical Society of America A*, 23(10):2359–2372, Oct 2006.
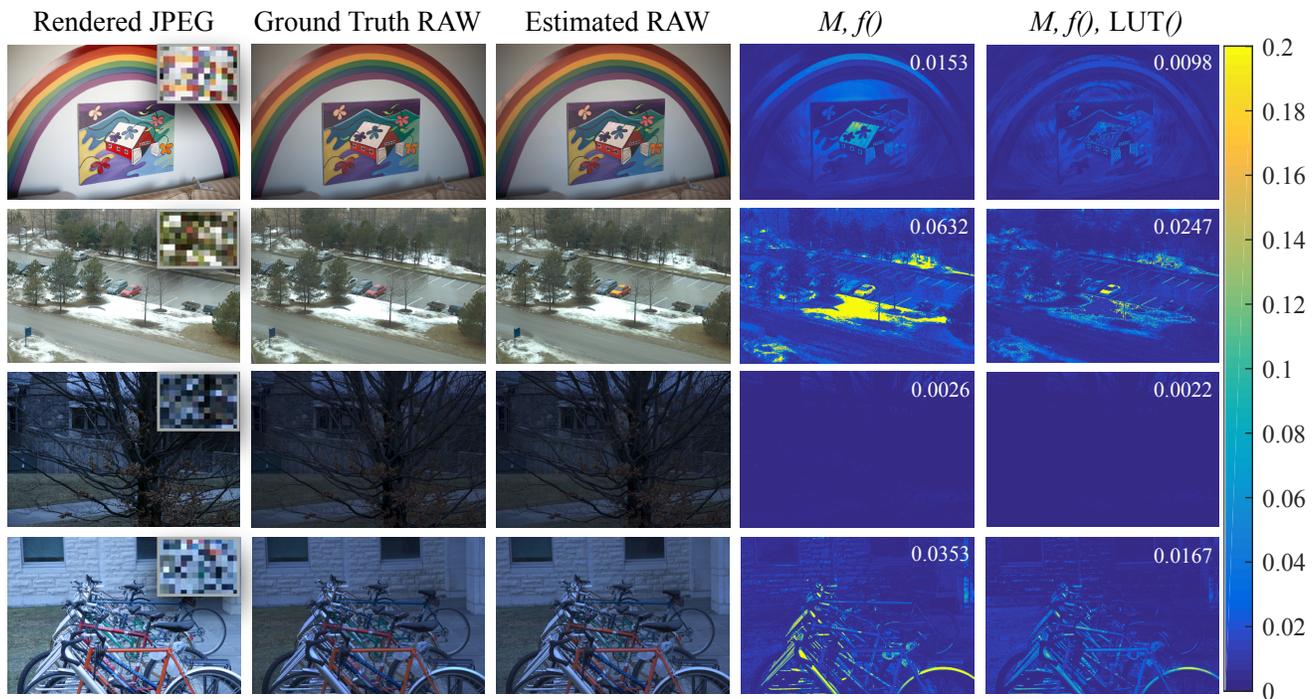
Figure 8. *Visualization of one-shot radiometric calibration through a simulated 140-patch color checker, shown at the top-right corner of each Rendered JPEG image. The error maps in the 4th and 5th columns respectively visualize the per pixel RMSE for our rank-based method with & without the gamut mapping LUT. The RMSE of each whole image is shown at the top-right corner of each error map. All raw images are shown with a 0.5 gamma.*

[11] Eric K. Garcia, Raman Arora, and Maya R. Gupta. Optimized regression for efficient function evaluation. *IEEE Transactions on Image Processing*, 21(9):4128–4140, 2012.

[12] Theo Gevers, Graham D. Finlayson, and Raimondo Schettini. Color for image indexing and retrieval. volume 94, pages 1–3, 2004.

[13] J. Giesen, E. Schuberth, K. Simon, P. Zolliker, and O. Zweifel. Image-dependent gamut mapping as optimization problem. *Image Processing, IEEE Transactions on*, 16(10):2401–2410, Oct 2007.

[14] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical optimization*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], 1981.

[15] Han Gong and Darren Cosker. Interactive removal and ground truth for difficult shadow scenes. *Journal of the Optical Society of America A*, 33(9):1798–1811, Sep 2016.

[16] Han Gong and Darren Cosker. User-assisted image shadow removal. *Image and Vision Computing*, 62:19–27, 2017.

[17] Han Gong, Graham D Finlayson, and Maryam M Darrodi. Concise radiometric calibration using the power of ranking. In *British machine vision Conference*. BMVA, 2017.

[18] Seon Joo Kim, Hai Ting Lin, Zheng Lu, Sabine Susstrunk, 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*, 34(12):2289–2302, 2012.

[19] Hai Ting Lin, Zheng Lu, Seon Joo Kim, and Michael S. Brown. Nonuniform lattice regression for modeling the camera imaging pipeline. In *European Conference on Computer Vision*, volume 7572 of *Lecture Notes in Computer Science*, pages 556–568. Springer, 2012.

[20] S. Lin, Jinwei Gu, S. Yamazaki, and Heung-Yeung Shum. Radiometric calibration from a single image. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–938–II–945 Vol.2, June 2004.

[21] A.P. Petrov. On obtaining shape from color shading. *Color research and application*, 18(4):236–240, 1993.

[22] E.A. Rakhmanov, E.B. Saff, and Y.M. Zhou. Minimal discrete energy on a sphere. *Mathematical Research letters*, pages 647–662, 1994.

[23] R. Ramanath, W.E. Snyder, Y. Yoo, and M.S. Drew. Color image processing pipeline. *Signal Processing Magazine, IEEE*, 22(1):34–43, 2005.

[24] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. *ACM Transaction on Graphics*, 21(3):267–276, July 2002.

[25] Alvy Ray Smith. Color gamut transform pairs. *ACM Siggraph Computer Graphics*, 12(3):12–19, 1978.