Practical photoquantity measurement using a camera

Julien Dehos, Eric Zéghers, Laurent Sarry, François Rousselle, Christophe Renaud

Practical photoquantity measurement using a camera

Julien DEHOS, Éric ZÉGHERS, Laurent SARRY
Université d'Auvergne,
Image Science for Interventional Techniques,
Faculté de médecine
28 place Henri Dunand, F-63001 Clermont-Ferrand
julien.dehos@gmail.com, eric.zeghers@u-clermont1.fr, laurent.sarry@u-clermont1.fr

François ROUSSELLE and Christophe RENAUD
Université du Littoral Côte d'Opale,
Laboratoire d'Informatique, Signaux et Images de la Côte d'Opale
Maison de la Recherche Blaise Pascal
50, rue Ferdinand Buisson, BP 719, F-62228 Calais
roussell@lisic.univ-littoral.fr, renaud@lisic.univ-littoral.fr

Abstract

An image output by a camera is generally not a faithful representation of the real scene, because it undergoes a series of radiometric disturbances during the imaging process. This paper proposes a method for obtaining a more accurate measure of the light seen by a camera. Our method requires no specific calibration apparatus and only minimal supervision. Nevertheless, it is quite comprehensive, since it accounts for response function, exposure, vignetting, spatial non-uniformity of the sensor and colour balancing. Our method works in two steps. First, the camera is calibrated off-line, in a photoquantity sense. Then, the photoquantity of any scene can be estimated in-line. Our method is therefore geared to a wide range of computer vision applications where a camera is expected to give a measurement of the visible light. The paper starts by presenting a photoquantity model of the camera imaging process. It then describes the key steps of calibration and correction method. Finally, results are given and analyzed to evaluate the relevance of our approach.

1 Introduction

Computer vision handles various kinds of data, from videos to 3D medical images. This paper solely deals with 2D stills typically obtained from a computer-controlled colour camera. Such images can be used, for example, to reconstruct scene geometry [5, 22], estimate the material reflectances [19] or compute a wide-field view by image mosaicing [11]. To achieve these tasks accurately, input images generally have to be a quite faithful measure of the light seen by the camera. However, this is not the case since the imaging process introduces geometric and radiometric distortions.

In this paper, we only deal with radiometric distortions. Our objective is to deploy a comprehensive yet practical method which can reduce radiometric distortions and thus give a quite faithful measure of the light seen by the camera.

We make several assumptions about the camera:

- electronic gain and optical aperture are constant,

- exposure time is known and controlled,

- optics can be temporarily removed from the camera.

Our method comprises two steps. First, the camera is calibrated. This step has to be done a single time, off-line. The correction step then computes the photoquantity of any static scene, in-line, based on the characterization of the camera given by the calibration step.

Our method accounts for:

- non-linear response functions (one response for each colour channel),

- colour balancing (global adjustment of the colour channels),

- spatial non-uniformity of the sensor (due to imperfections in the building process),

- vignetting (spatial non-uniformity due to the internal geometry of the camera).

Furthermore, since the method performs high dynamic range imaging (HDRI) through multiple exposures, it can be used with highly contrasted scenes and possesses some inherent robustness to noise.

The main contribution of this paper is a practical method to characterize a camera and then measure photoquantity. It also proposes an extension of the Goldman-Chen vignetting model and a calibration scheme based on this model.

The remainder of the paper is organized as follows. Section 2 reviews some related work. Section 3 presents a photoquantity model of the imaging process. Section 4 details the model inversion, which gives a photoquantity measurement method. Section 5 details the calibration step. Section 6 presents and discusses some of our results and Section 7 concludes.

2   Related work

There is a wealth of literature on radiometric camera correction. However, work is generally limited to a specific kind of radiometric distortion (such as, vignetting or camera response), application (mosaicing, HDRI) or technology (CCD, CMOS).

In [17], Tarel gives a fairly comprehensive study of radiometric camera distortion, but only considers CCD cameras leading to a specific characterization. Based on the study, he also explains how to radiometrically correct camera acquisitions. However, no practical method is given, as it was not the purpose of his paper.

High dynamic range imaging is a related field which has gained a lot of attention. In [10], Mann and Picard propose a method to estimate camera response from multiple exposures. Their method works as follows. Let us consider a point in the scene. This point emits a constant light flux. Therefore, by comparing the exposures and the corresponding values output by the camera, we can characterize the camera response for the measured light. Thus, by considering several points in the scene, the full response can be measured.

In [2], Debevec and Malik propose another method to estimate the camera response from multiple exposures. Their method consists in considering the values of several points under different exposures. This makes it possible to define a minimization problem. The resolution of this problem gives the camera response.

In [16], Robertson et al. propose a similar method to Debevec-Malik. However, they use a different formalization, starting out from a statistical approach. They obtain a minimization problem, and propose a relaxation method to solve it.

Finally, in [14], Mitsunaga and Nayar propose a polynomial model of camera response. Using this model, they formulate the response characterization as an optimization problem and propose an iterative algorithm to solve it.

There has been a recent surge in work on vignetting correction. The methods proposed are generally based on a vignetting model. For example, in [20], Yu proposes a hypercosine model of vignetting and an antivignetting method based on wavelet denoising and decimation. In [21], Zheng et al. use the Kang-Weiss model [8] to propose a single-image vignetting correction method.

Vignetting correction is often associated with HDRI and image mosaics. In [3], Goldman and Chen propose a method to estimate vignetting, response function and exposures from an image mosaic. They propose a 6th-order polynomial vignetting model and use the empirical model of response given in [3] to formulate the problem as a non-linear optimization.

In [9], Kim and Pollefeys use the same vignetting and response models as Goldman and Chen. However, they decouple vignetting from the response function and propose a method to successively estimate response function and vignetting by solving two linear least-square problems.

Finally, in [6, 7], Litvinov and Schechner propose a method to estimate vignetting, response function and exposures from a frame sequence. Their method is not based on a model. Instead, they use the fact that regions are represented on several frames to state and resolve a non-parametric linear least-square problem.

In contrast to previous methods, ours is both general and practical. First, it accounts for the main radiometric distortions occurring throughout the camera imaging process. Second, it is not limited to a specific kind of camera or application. Finally, it is practical in the sense that it requires no specific calibration apparatus and only little supervision.

3   Photoquantity model

In this section, we start with a recap on how a camera creates an image. This enables us to explain the main radiometric distortions, and thus to establish a photoquantity model of the imaging process.
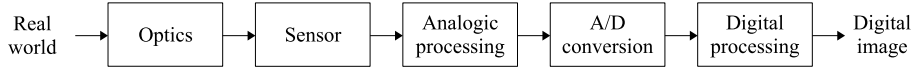
Figure 1: Camera acquision radiometric process.

3.1 Image formation

The digital imaging process is illustrated Fig. 1. First, the light travels from the real scene onto the optics of the camera. It then hits the cells of the light sensor where it is converted into electric signals. These signals are then amplified, digitized and possibly post-processed. This gives the output image.

In the case of colour imaging, the sensor is composed of several kinds of cell. Each cell type has its own spectral response. The colour of a pixel is given by associating cells of different kinds.

3.2 Radiometric distortions

The imaging process introduces radiometric distortions in the output image. The main distortion sources are described below.

3.2.1 Spatial non-uniformity

Spatial non-uniformity means that for the same input light, the camera can output different values according to the position reached in the image. This phenomenon is sometimes called vignetting, but has several forms and origins [3]. In this paper, we distinguish "sensor non-uniformity" and "vignetting".

We refer to "sensor non-uniformity" as the variation in sensitivity between cells of the sensor. This stems from the fact that the industrial sensor building process is not perfect, which means that two different cells exposed to the same irradiance can output different values.

We refer "vignetting" as the variation in light reaching a cell according to the position of this cell. Vignetting is due to the geometry of the camera optics, which are generally aligned to the center of the sensor. Thus, at the sensor center, the aperture of the optical system is maximal, enabling a lot of light to be collected, whereas in the corners of the sensor, the aperture is reduced, and only allows a little light to be collected, hence the well-known corner darkening.

3.2.2 Exposure and response

Another important radiometric distortion is the response of the imaging system. This means that the value of a pixel is not linear in relation to the light collected by the cells of this pixel. There are two reasons for this phenomenon.

Firstly, sensor cells have a limited dynamic range. If the input light is too low, a cell detects nothing while if the light is too high, the cell outputs the maximum value. Thus, the input light can only be quantified if it maps between these two thresholds.

Secondly, the cells may not give a linear response. This means that there is no linear function that maps the input light to the output value of a cell.

3.2.3 Colour balancing

As stated earlier, a sensor can associate different kinds of cells (colour channels) to quantify colour. Consider a pixel. We call l the spectral distribution function of the light reaching this pixel and r_λ the spectral response of the colour channel λ of the camera. Thus, the value l_λ output by the camera for this colour channel is :

$$l_\lambda = \int r_\lambda(x)l(x)dx$$

where x denotes wavelength.

This means that a camera uses its own, generally 3-dimensional, colour system. Therefore, the values output by a camera are not radiometric values because we cannot deduce the light spectrum from the values output by the camera. Furthermore, since the spectral responses of the camera generally do not exactly match the spectral responses of the CIE standard observer, the camera values cannot be expressed faithfully in a standard CIE colour system. For this reason, in [12], camera values are called *photoquantities*.

The problem is that we cannot exactly transform the photoquantities output by a camera in a standard device-independent colour system. As a result, the colours in the output image may be different to the colours label we can really see.

3.2.4 Noise

Finally, as with any measurement devices, cameras are subject to noise. Noise can be seen as a random distortion of the output values. There are roughly two kinds of noise in a camera imaging system. The first is thermal noise, which depends on heat.

It can be seen as a biased additive noise and is therefore relatively important at low output levels. The second is measurement noise, which depends on the output level.

3.3 Model

Here is the photoquantity model we use. First, light comes from the scene to the camera. Then, the camera optics produce vignetting and the non-uniform sensitivity of the sensor introduces high-frequency spatial variations. Finally, the sensor responses and exposure introduce non-linearity between the incoming light and the output image. As it converts light into photoquantities, the sensor introduces colour balancing in the output image.

Let l_p be the radiance coming from the scene to a pixel p in the output image. The camera measures this radiance using the corresponding sensor cells. Let r_λ be the spectral response of the cells for the colour channel λ. The matching photoquantity $l_{p\lambda}$ is therefore:

$$l_{p\lambda} = \int r_\lambda(x) l_p(x) dx$$

Let $c_{pk\lambda n}$ be the value given by the camera for the pixel p, the exposure k of exposure time t_k and during the acquisition n. This gives:

$$c_{pk\lambda n} = a_\lambda f_\lambda (u_{p\lambda} v_{p\lambda} t_k l_{p\lambda}) + b_{pk\lambda n} \tag{1}$$

where $v_{p\lambda}$ is the vignetting at p, $u_{p\lambda}$ the sensor sensitivity at p, f_λ the response of the colour channel λ, a_λ the colour balancing factor for λ, and $b_{pk\lambda n}$ the noise.

The radiometric distortions are modeled as follows:

- the sensor responses fit the non-parametric model of Debevec and Malik,

- colour balancing is modeled by a colour-channel-dependent factor,

- sensor non-uniformity is modeled by a look-up table (LUT) giving the sensitivity attenuation of each pixel,

- vignetting follows an extension of the Goldman-Chen polynomial model.

4 Model inversion

In the previous section, we gave a photoquantity model of the imaging process. We now give a camera correction method by inverting the photoquantity model. The problem is to compute the photoquantity $l_{p\lambda}$ assuming the camera output $c_{pk\lambda n}$ is known.

4.1 Denoising

First, we have to remove noise ($b_{pk\lambda n}$). There is a wealth of research into how to denoise images [15]. Here, we exploit the fact that we can easily control the camera. Indeed, we do not explicitly know the $b_{pk\lambda n}$, but if we assume that thermal noise is negligible and that measurement noise is unbiased, then the denoised value of a pixel can be estimated by computing its mean value. Thus, we can compute the denoised value $\bar{c}_{pk\lambda}$ from N values $c_{pk\lambda n}$ of the same scene (with constant exposure time and view direction) with:

$$\bar{c}_{pk\lambda} = \frac{1}{N} \sum_{n=1}^{N} c_{pk\lambda n} \tag{2}$$

Note that a greater N value translates better noise reduction. However, this requires more image acquisitions. In practice, we can use $N = 1$ to focus on speed (during the correction step for example) and $N > 1$ to focus on accuracy (during the calibration step for example).

We can thus rewrite Equation 1 as:

$$\bar{c}_{pk\lambda} = a_\lambda f_\lambda (u_{p\lambda} v_{p\lambda} t_k l_{p\lambda}) \tag{3}$$

In the following, we consider $\bar{c}_{pk\lambda} = c_{pk\lambda}$ and just write $c_{pk\lambda}$.

4.2 Linearization

Now, we have to cancel the response functions and colour balancing. Our statement is:

$$c_{p\lambda} = u_{p\lambda} v_{p\lambda} l_{p\lambda} \tag{4}$$

Therefore, Equation 3 is equivalent to:

$$c_{pk\lambda} = a_\lambda f_\lambda (c_{p\lambda} t_k)$$

We cannot consider that any sensor is linear but we can consider that the more input light there is, the higher the output value of the sensor will be. We can thus assume that f_λ is monotonous and write:

$$f_\lambda^{-1} \left(\frac{c_{pk\lambda}}{a_\lambda} \right) = c_{p\lambda} t_k$$

i.e.

$$\ln f_\lambda^{-1} \left(\frac{c_{pk\lambda}}{a_\lambda} \right) = \ln(c_{p\lambda}) + \ln(t_k) \tag{5}$$

We use the method proposed by Debevec and Malik in [2] to robustly compute $c_{p\lambda}$:

$$c_{p\lambda} = \exp \left(\frac{\sum_{k \in K} w(c_{pk\lambda}) \left[\ln f_\lambda^{-1} \left(\frac{c_{pk\lambda}}{a_\lambda} \right) - \ln t_k \right]}{\sum_{k \in K} w(c_{pk\lambda})} \right) \tag{6}$$

where w is a weighting function that discards under-exposed and over-exposed values:

$$w(c) = \begin{cases} c - c_{\min} & \text{if } c \leq \frac{1}{2}(c_{\min} + c_{\max}) \\ c_{\max} - c & \text{otherwise} \end{cases}$$

$c_{\min}$ and $c_{\max}$ are respectively the minimum and the maximum values the camera is able to provide.

We may notice that the photoquantity $c_{p\lambda}$ is computed from several camera values $c_{pk\lambda}$. Thus, $c_{p\lambda}$ has values in $\mathbb{R}$ where as $c_{pk\lambda}$ has values in $[c_{\min}, c_{\max}]$. In other words, the images output by the camera are low dynamic range (LDR) images where as the images obtained after linearization are high dynamic range (HDR) images.

4.3 Uniformization

Finally, we have to cancel spatial non-uniformity. From Equation 4, we have:

$$l_{p\lambda} = \frac{c_{p\lambda}}{u_{p\lambda} v_{p\lambda}}$$

which gives us the photoquantity $l_{p\lambda}$ we are looking for.

5 Calibration

In the previous section, we described a method to compute photoquantity based on the photoquantity model of the camera. In this section, we explain how to estimate the model parameters for a given camera. Our calibration method consists in acquiring a set of images from which response functions, colour balancing, sensor spatial non-uniformity and vignetting are then successively estimated.

5.1 Image acquisition

5.1.1 Acquisition scenes

To characterize a camera using our model, we have to take images of several scenes:

1. a scene to estimate the response functions,

2. a scene containing a white chart (such as a simple sheet of paper) to estimate colour balancing,

3. a scene captured from a number V of slightly different view directions to estimate vignetting.

4. a "irradiance scene" captured by the sensor, with no optics, to estimate sensor spatial non-uniformity,

In practice, we can place a white chart in the real environment and take images from the V view directions. This enables us to estimate response functions, vignetting and colour balancing[1] (items 1, 2 and 3). Finally, we have to use one more image from the scene, with no optics, to estimate sensor spatial non-uniformity (item 4). This means a total of $V + 1$ scenes is required.

---

[1]The light reflected by the white chart depends on the material and on the incident light. The chart should therefore be placed at a position where the light can be considered as a reference.

5.1.2 Exposures

To estimate the camera response functions and measure the photoquantity of a scene, we have to capture this scene using several exposure times. Here, we consider that a scene is captured K times and we call t_k the exposure time of the kth capture. In fact, K and the t_k do not have to be the same for all the scenes considered.

5.1.3 Denoising

Since the calibration step is done once and its result is used many times, it can be profitable to take some time to improve calibration accuracy by reducing noise. This can be done by taking each image N times (with $N > 1$) and computing Equation 2.

5.2 Response functions

As stated earlier, we use the method proposed by Debevec and Malik in [2] to estimate the response functions f_λ. Knowing the response functions enables us to linearize the pixel values and then to estimate the other factors of radiometric distortion. Thus, we first consider each colour channel independently, which means that we can set $a_\lambda = 1$. We set $g_\lambda(c) = \ln f_\lambda^{-1}(c)$ and rewrite Equation 5 with:

$$g_\lambda(c_{pk\lambda}) = \ln c_{p\lambda} + \ln t_k \tag{7}$$

From this equation, Debevec and Malik propose to minimize the objective function:

$$\mathscr{O}_\lambda = \sum_{\substack{p \in P^- \\ k \in K}} \left(w(c_{pk\lambda}) [g_\lambda(c_{pk\lambda}) - \ln c_{p\lambda} - \ln t_k] \right)^2 + \beta \sum_{c=c_{\min}+1}^{c_{\max}-1} [w(c) g_\lambda''(c)]^2$$

where P^- is a subset of the set of image points and g_λ'' is the second derivative of g_λ defined by:

$$g_\lambda''(c) = g_\lambda(c-1) - 2g_\lambda(c) + g_\lambda(c+1)$$

The objective function $\mathscr{O}_\lambda$ can be expressed as an over-determined quadratic system which can be solved by singular value decomposition. The first term of $\mathscr{O}_\lambda$ forces the unknowns g_λ and $c_{p\lambda}$ to match the measured values $c_{pk\lambda}$ and t_k in Equation 7. The second term of $\mathscr{O}_\lambda$ smoothes the computed response function. Finally, the parameter β defines the weight of the smoothing.

5.3 Colour balancing

As explained before, the colours given by the camera are influenced by the spectral responses of the different kinds of cell and by the spectral density of the light. In our method, we adjust colour channels according to the reference white point given by the white chart we have placed in one of the captured scenes. The corresponding pixel should have approximately the same values for all colour channels. If this is not the case, we select a channel as a reference (for example, the green channel) and apply a specific correcting factor to each other channel. Let $[R_w \quad G_w \quad B_w]^T$ be the linearized colour (i.e. the colour given by Equation 6) of the white point. We define the green-based colour balancing factors a_λ by:

$$a_r = G_w/R_w$$
$$a_g = 1$$
$$a_b = G_w/B_w$$

Note here that the colour balancing factors can be included into the camera response functions. Thus, inverting response functions also performs colour correction.

5.4 Sensor spatial non-uniformity

Sensor spatial non-uniformity can be calibrated using the irradiance scene captured with no optics. Indeed, in these conditions, sensor cells are exposed to an overall irradiance, which remains roughly constant over the sensor.

Thus, to estimate sensor spatial non-uniformity, we can compute linearized colour-corrected colours in the scene captured with no optics. After normalization, this gives us the $u_{p\lambda}$, which we can store in a LUT.

5.5 Vignetting

The last part of the model we have to calibrate is vignetting. We assume the vignetting $v_{p\lambda}$ of a pixel $p = (p_x, p_y)$ is:

$$v_{p\lambda} = 1 + m_{1\lambda} R_\lambda(p) + m_{2\lambda} R_\lambda(p)^2 + m_{3\lambda} R_\lambda(p)^3$$

where

$$R_\lambda(p) = m_{4\lambda}(p_x - m_{5\lambda})^2 + (p_y - m_{6\lambda})^2$$

This vignetting model is an extension of the polynomial model proposed by Goldman and Chen in [3]. This model presents an axis-aligned elliptical symmetry. The parameters $m_{1\lambda}$, $m_{2\lambda}$ and $m_{3\lambda}$ are the polynomial factors. The parameters $m_{5\lambda}$ and $m_{6\lambda}$ define the center of the symmetry. Finally, $m_{4\lambda}$ is the scale factor of the elliptical symmetry.

The vignetting model parameters can be determined by considering a scene from different view directions. Let s be a point of the scene and $l_{is\lambda}$ the photoquantity going from s to the camera in the view direction i. Since we know all the components of the photoquantity model of the camera except vignetting, we have:

$$c_{is\lambda} = v_{is\lambda} l_{is\lambda}$$

where $c_{is\lambda}$ is the colour coming from s, seen by the camera in the view direction i, after linearization, colour correction and sensor uniformization. The corresponding vignetting factor is $v_{is\lambda}$.

Similarily, in another view direction j, we have:

$$c_{js\lambda} = v_{js\lambda} l_{js\lambda}$$

Since the V scenes we use to estimate vignetting are seen from the same camera viewpoint but from different camera directions, we can consider that $l_{ks\lambda}$ is constant for all directions k. Therefore, we have $l_{is\lambda} = l_{js\lambda}$ and:

$$c_{is\lambda} v_{js\lambda} - c_{js\lambda} v_{is\lambda} = 0$$

We cannot estimate the vignetting model from this equation alone. However, if we consider several points and view directions, we have the following optimization problem:

$$\arg\min_{m_{1\lambda}, m_{2\lambda}, m_{3\lambda}, m_{4\lambda}, m_{5\lambda}, m_{6\lambda}} \left(\sum_{\substack{i<j \\ s \in S_i \cap S_j}} \| c_{is\lambda} v_{js\lambda} - c_{js\lambda} v_{is\lambda} \| \right)$$

where S_k is a set of points visible from the camera in the direction k. To solve this optimization problem we use the Levenberg-Marquardt method [13].

Note that after calibration, the vignetting model can be evaluated for all the pixels and included in the spatial non-uniformity LUT to optimize the correction step.

Note also that this vignetting calibration method requires us to identify the position of a scene point in the different images (taken from different view directions). This can be done automatically (using feature detection or geometric calibration) or manually.

6 Results and analysis

In this section, we present and analyze some of our results. We first consider the calibration step, then the correction step, and finally, we analyze our approach from different standpoints. We also give the computation times we obtained with our non-optimized Matlab implementation. We used a 1024x768 CCD colour camera (MATRIX VISION mvBlueFox-121C).

6.1 Calibration

Image acquisition

We recall that the calibration method requires us to consider $V + 1$ scenes. Each scene is captured using K different exposure times. For one scene and one exposure, N images are taken to reduce noise. Thus, the calibration method implies taking $NK(V + 1)$ images. In the experiment presented here, we used exposure times ranging from 1 ms to 200 ms and took a total of 320 images ($N = 10$, $K = 8$ and $V = 3$). This may seem a lot, but with a computer-controlled camera, the calibration (image acquisition and model computation) can be performed in less than 5 minutes. Furthermore, with $N = 4$ and $K = 4$, we only need 64 images while still getting good results.

Fig. 2 presents selected denoised images of the calibration scene under different exposure times and view directions, and some denoised images of the irradiance scene (taken with no optics). With $N = 10$, denoising all our images takes less than 1 s.

Response Functions

Using K exposure times and one view direction, the calibration method then computes the camera response functions (10 s using 500 reference points) (Fig. 3) and the corresponding linearized high dynamic range image (10 s per image) (Fig. 4). Computing the response functions and the HDR images therefore takes less than 1 minute.

Note that a CCD sensor has a quasi-linear response function, which is not the case for a CMOS sensor.

Colour balancing

Using the white chart (the sheet of paper to the left of the rolls of adhesive tape in Fig. 4), the method then computes colour factors (1 ms). These factors are then included in the response functions (Fig. 3) to give the colour-corrected images (Fig. 4) (20 ms per image). Computing the colour-corrected images therefore takes less than 1 s.

Sensor spatial non-uniformity

Using the colour-corrected linearized image of the irradiance scene, the method computes a LUT (Fig. 3) (0.13 s) for the sensor spatial non-uniformity correction (Fig. 4) (20 ms per image). Computing the colour-corrected images therefore takes less than 1 s.
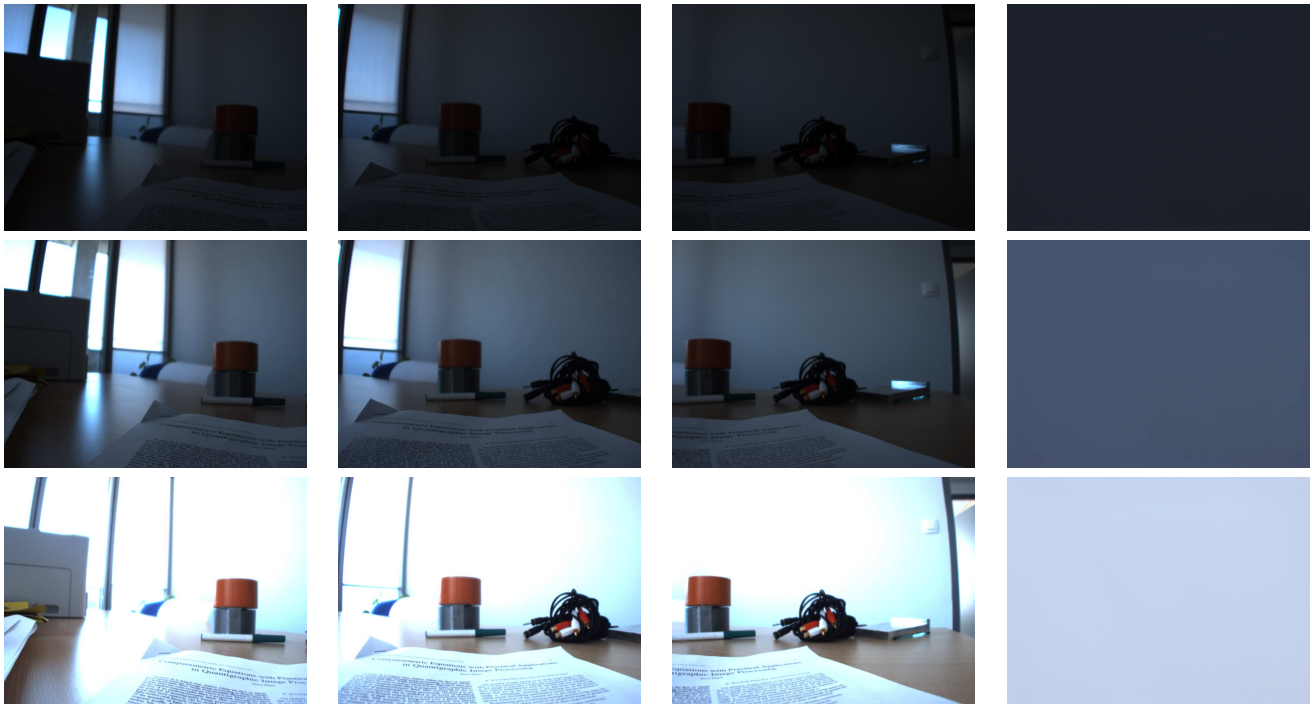

Figure 2: The calibration scene under different view directions (columns 1 to 3) and the irradiance scene (column 4), under different exposure times (first row: 2 ms; second row: 5 ms; third row: 20 ms).
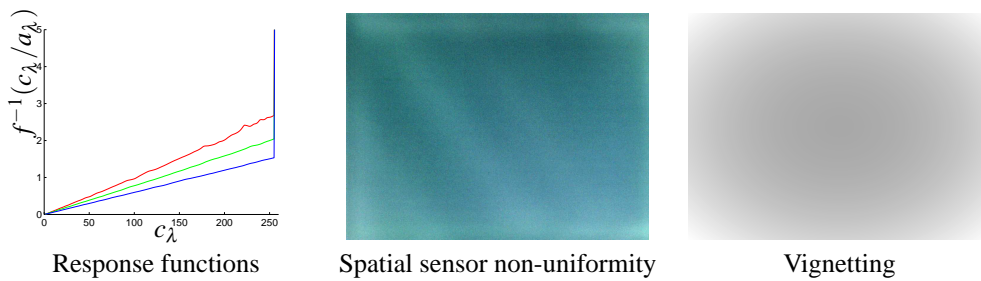

Response functions   Spatial sensor non-uniformity   Vignetting

Figure 3: Calibration of the mvBlueFox-121C camera.

HDR acquisition   Colour correction   Sensor uniformization   Vignetting correction
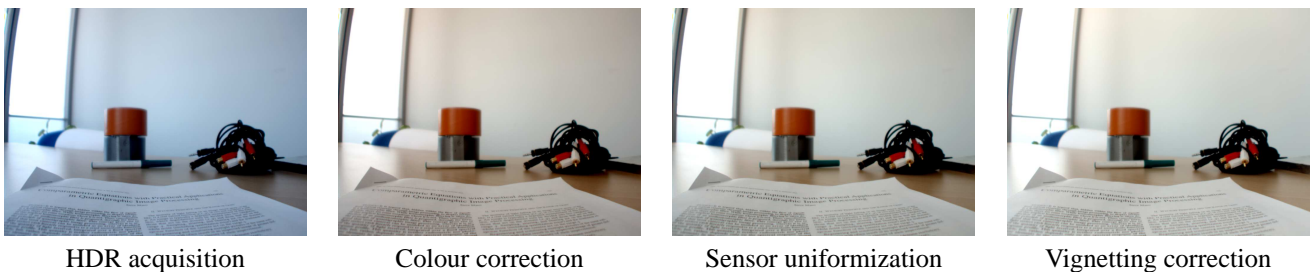
Figure 4: Correction of the different components of the model.

Vignetting

Finally, using the different view directions of the calibration scene, the method estimates the vignetting model for the camera and computes a correction LUT (Fig. 3) which is combined with the sensor spatial non-uniformity LUT (40 s with 10 reference points). This enables us to compute the photoquantity of the calibration scene (Fig. 4).

Figure 5: Images of a test scene under different exposure times (columns 1 to 3). Linearly tone-mapped photoquantity of the test scene obtained with our method (column 4).

6.2 Correction

Image acquisition

First, we acquire some images of the scene. Since we are considering a single view direction, we need NK images only (Fig. 5). We may notice that N, K and the exposure times can be different than the ones used during the calibration step. We may even state $N = 1$ while having good results because the correction method inherently performs noise reduction (this is detailed below).

Image correction

The previously estimated response functions enable us to compute a linearized colour-corrected HDR image. Then, using the estimated LUT, we can uniformize this image. This results in the expected photoquantity (Fig. 5).

6.3 Analysis

In this section, we analyze the method from several standpoints: the accuracy of the linearization and colour-correction processes, the influence of noise and the accuracy of the vignetting modeling.

We call e_p the *relative error* for a pixel p of an image P:

$$e_p = \frac{||c_p - \hat{c}_p||_2}{||c_p||_2} \tag{8}$$

where c_p is the reference colour and $\hat{c}_p$ the computed colour. Let $e_{\max}$ be the maximal value of the relative error, e_{mean} the mean value of the relative error and e_{std} the standard deviation of the relative error:

$$e_{\max} = \max_{p \in P} (e_p) \ , \quad e_{\mathrm{mean}} = \frac{1}{\#P} \sum_{p \in P} e_p \ , \quad e_{\mathrm{std}} = \frac{1}{\#P - 1} \sqrt{\sum_{p \in P} (e_p - e_{\mathrm{mean}})^2}$$

Linearization and colour-correction

In order to analyze the linearization and colour-correction accuracy of the correction method, we use a scene containing a ColorChecker. A ColorChecker has a set of patches, reflectances of which are calibrated. The manufacturer gives the CIE-RGB values for the reflected light under the D65 CIE Standard Illuminant. Our scene is exposed to daylight conditions, which corresponds to this illuminant. Thus, the measured values should match the given values (ground-truth). However, the values measured by a camera rely on exposure time, sensor response... so that we cannot expect absolute equality but an affine relationship.

Thus, for each channel of each ColorChecker patch, we have the measured value and the ground-truth (i.e. a 2D point). We report all these points on a graph and compute the best affine function between the measured values and ground-truth (by linear regression). The measured values are accurate if the points are aligned (linear correlation).

Fig. 6 (a) shows the uncorrected LDR image, directly output by the camera (exposure time: 20 ms), and Fig. 7 (a), its correlation to ground-truth. We may notice that below 230, the measured values are quite well correlated to ground-truth. This can be explained by the linear response of CCD sensors (CMOS sensors give less correlated values because of their non-linear response). However, above 230, the sensor is over-exposed and the measured values do not match ground-truth (outliers). The graph also reveals colour shifting of the measured values (blue channel data points are below the regression line).

Fig. 6 (b) shows the corrected HDR image (linearly tone-mapped in a displayable range), given by our correction method, and Fig. 7 (a), its correlation to ground-truth. Here, the measured values are quite well aligned with no significant outliers (more generaly, the method also linearizes data from CMOS sensors). Moreover, the points of each colour channel are distributed on both sides of the regression line, which means that the method corrects colour shifting.

Noise

We first investigate how many images (i.e. the value of N) are required to significantly reduce the noise. Theoritically, the standard deviation of $\bar{c}_{pk\lambda}$ drops with $1/\sqrt{N}$ (see Appendix A). In our experiments, setting $N = 4$ was sufficient to obtain $e_{\mathrm{mean}} < 0.5\%$ ($e_{\max} = 2.4\%$, $e_{\mathrm{mean}} = 0.4\%$ and $e_{\mathrm{std}} = 0.2\%$).

We also consider another question about noise: is the correction method inherently robust enough to use $N = 1$? We compared correction results obtained with $N = 10$ (c_p in Equation 8) and with $N = 1$ ($\hat{c}_p$). The relative error e_p is shown Fig. 8 ($e_{\max} = 4.9\%$, $e_{\mathrm{mean}} = 0.8\%$ and $e_{\mathrm{std}} = 0.4\%$).

Table 1: Correlation between output values and ground-truth (the higher the better).

	CCD outdoor	CCD indoor	CMOS outdoor	CMOS indoor
uncorrected values	0.85	0.63	0.84	0.67
corrected values	0.87	0.84	0.9	0.86
Hugin values	0.84	0.76	0.9	0.75

Vignetting

Finally, we evaluate the accuracy of the vignetting modeling. Fig. 9 (a) shows the vignetting model that we find in our experiment ($m_1 = 0.48$, $m_2 = 0.66$, $m_3 = -1.09$, $m_4 = 1.48$, $m_5 = 0.5$ and $m_6 = 0.69$). This figure gives the quantity of light u_p that reaches the pixel (p_x, p_y) of the sensor after vignetting. We compare this result with the vignetting ground-truth (Fig. 9 (b)), obtained in experimental conditions (using an industrial uniform 20×30 cm light source). We find $e_{max} = 9\%$, $e_{mean} = 1.6\%$ and $e_{std} = 1\%$ (Fig. 9 (c)).

6.4 Other results

We apply the proposed method using different kinds of camera (CCD, CMOS) and lighting conditions (outdoor scene lit by the sun, indoor scene lit by a halogen lamp). To evaluate the method, we consider some reference points and compare the uncorrected values, the values corrected by the proposed method and the values corrected using Hugin[2] [1]. The corrected values are defined in a high dynamic range therefore we compute correlation between output values and ground-truth (Table 1).

We may notice that our test scenes are not very constrasted which makes the uncorrected values quite representative of ground-truth, excepted when lighting conditions are not neutral (indoor scene). Hugin has many parameters and should give better results using a finer tuning. Finally, the method proposed in this paper gives good results, independently of lighting conditions, while requiring only minimal supervision.

(a)   (b)

Figure 6: Study of the linearization and colour-correction accuracy of the method: uncorrected LDR image (a) and corrected HDR image (b).

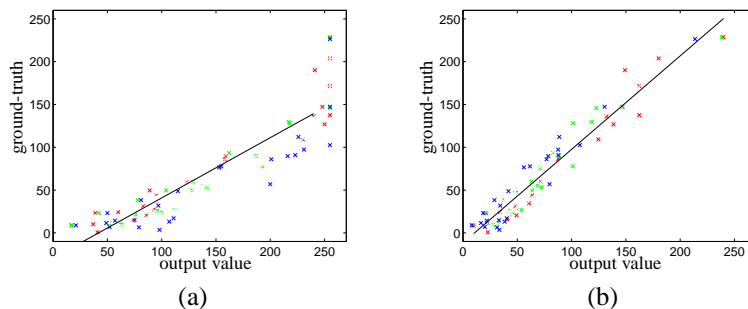(a)   (b)

Figure 7: Correlation to ground-truth: uncorrected LDR image (linear regression: $y = 0.7x - 29.47$, residual norm: 215.35) (a) and corrected (linearly tone-mapped) HDR image (linear regression: $y = 1.09x - 11.68$, residual norm: 126.34) (b). Each point corresponds to one colour channel (indicated by the colour of the point) of one patch of the ColorChecker.

---

[2]Image stitching software accounting for sensor response, vignetting and colour balancing.

Figure 8: Relative error resulting from ignoring noise during photoquantity computation (same scene as Fig. 6).
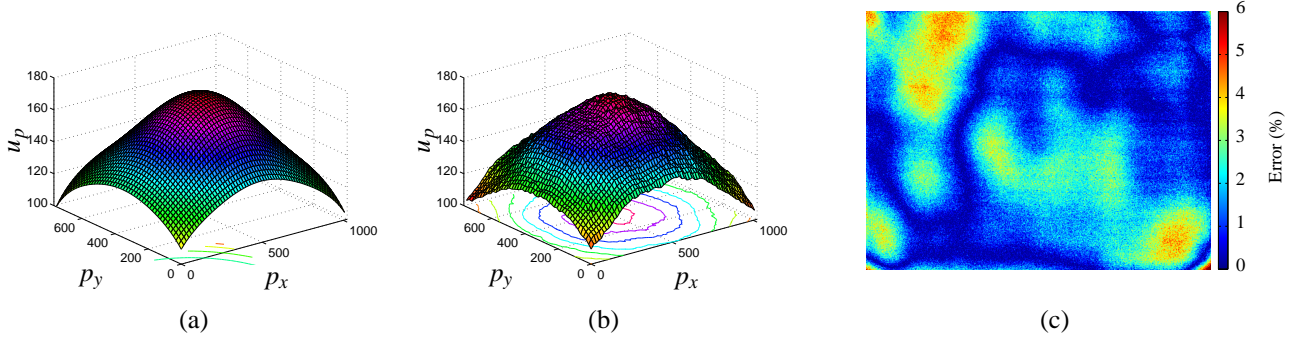

(a) (b) (c)

Figure 9: Vignetting modeling (a) vs measured ground-truth (b) (p_x and p_y are pixel coordinates). Relative error between modeling and ground-truth (c).

7   Conclusion

This paper described a practical method for measuring photoquantity using a camera. First, we proposed a photoquantity model of the imaging process. This model accounts for non-linear response functions, sensor exposure, colour balancing, sensor spatial non-uniformity and vignetting. Then, by inverting this model, we gave a photoquantity correction method which uses a small set of images together with the model parameters of the camera. We then gave a method to estimate these parameters using the camera. Finally, we presented and discussed some of our results.

Our method is general as well as practical. It accounts for the main radiometric distortions occurring during the end-to-end camera imaging process. Furthermore, the method is not limited to a specific kind of camera or application. Finally, it is practical in the sense that it requires no specific calibration apparatus and only minimal supervision.

Acknowledgements

A   Analysis of the denoising technique

We model the noise terms $b_{pk\lambda n}$ (Equation 1) as independent random variables with expected value 0 and standard deviation $\sigma_{pk\lambda}$. Therefore, the $c_{pk\lambda n}$ are also independent random variables, with expected value $c_{pk\lambda}$ and standard deviation $\sigma_{pk\lambda}$.

In Equation 2, we define $\bar{c}_{pk\lambda}$ as:

$$\bar{c}_{pk\lambda} = \frac{1}{N} \sum_{n=1}^{N} c_{pk\lambda n}$$

Since the expected value is linear, we have:

$$\mathbb{E}(\bar{c}_{pk\lambda}) = \mathbb{E}\left(\frac{1}{N} \sum_{n=1}^{N} c_{pk\lambda n}\right) = \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}(c_{pk\lambda n}) = \frac{1}{N} \sum_{n=1}^{N} c_{pk\lambda} = \frac{1}{N} N c_{pk\lambda} = c_{pk\lambda}$$

This means that the denoised value effectively converges to the desired value.

Finally, the variance of $\bar{c}_{pk\lambda}$ is:

$$\text{Var}(\bar{c}_{pk\lambda}) = \text{Var}\left(\frac{1}{N} \sum_{n=1}^{N} c_{pk\lambda n}\right) = \frac{1}{N^2} \sum_{n=1}^{N} \text{Var}(c_{pk\lambda n}) = \frac{1}{N^2} \sum_{n=1}^{N} \sigma_{pk\lambda}^2 = \frac{1}{N^2} N \sigma_{pk\lambda}^2 = \frac{\sigma_{pk\lambda}^2}{N}$$

This means that standard deviation of the denoised value is $\sigma_{pk\lambda}/\sqrt{N}$ and therefore drops with $1/\sqrt{N}$.

References

[1] d'Angelo, P. : Radiometric alignment and vignetting calibration, The 5th International Conference on Computer Vision Systems (ICVS), (2007)

[2] Debevec, P. E., Malik, J.: Recovering high dynamic range radiance maps from photographs, Proceedings of the 24th annual conference on Computer graphics and interactive techniques (SIGGRAPH), pp. 369-378 (1997).

[3] Goldman, D.B., Chen, J.-H.: Vignette and Exposure Calibration and Compensation, Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV), vol. 1, pp. 899-906 (2005)

[4] Grossberg, M.D., Nayar, S.K.: What is the Space of Camera Response Functions?, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 602-609 (2003)

[5] Horn, B. K. P., Brooks, M. J.: Shape From Shading, MIT Press (1989)

[6] Litvinov A, Schechner Y. Y.: Addressing Radiometric Nonidealities: A Unified Framework, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 52-59 (2005)

[7] Litvinov, A., Schechner, Y. Y.: Radiometric framework for image mosaicking, Journal of the Optical Society of America A, vol. 22, pp. 839-848 (2005)

[8] Kang, S. B., Weiss, R. S.: Can We Calibrate a Camera Using an Image of a Flat, Textureless Lambertian Surface?, Proceedings of the 6th European Conference on Computer Vision (ECCV), vol. 2, pp. 640-653 (2000)

[9] Kim, S. J., Pollefeys, M.: Robust Radiometric Calibration and Vignetting Correction, IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, pp. 562-576 (2008)

[10] Mann, S., Picard, R. W.: On Being 'undigital' With Digital Cameras: Extending Dynamic Range By Combining Differently Exposed Pictures, Proceedings of IS&T, pp. 442-448 (1995)

[11] Mann, S., Picard, R. W.: Video Orbit of the Projective group: A New Perspective on Image Mosaicing, MIT Media Laboratory Perceptual computing Section Technical Report No. 338 (1995)

[12] Mann, S.: Comparametric equations with practical applications in quantigraphic image processing, IEEE Transactions on Image Processing, vol. 9, pp. 1389-1406 (2000)

[13] Marquardt, D. W.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters, SIAM Journal on Applied Mathematics, vol. 11, pp. 431-441 (1963)

[14] Mitsunaga, T., Nayar, S. K.: Radiometric Self Calibration IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 374-380 (1999)

[15] Motwani, M., Gadiya, M., Motwani, R., Harris, F. C. Jr.: A Survey of Image Denoising Techniques, Proceedings of GSPx (2004)

[16] Robertson, M. A., Borman, S., Stevenson, R. L.: Estimation-theoretic approach to dynamic range enhancement using multiple exposures, Journal of Electronic Imaging, vol. 12, pp. 219-228 (2003)

[17] Tarel, J.-P.: Calibration Radiométrique de Caméra, INRIA Research Report No. 2509 (1995)

[18] Yu, W., Chung, Y., Soh, J.: Vignetting Distortion Correction Method for High Quality Digital Imaging, Proceedings of the 17th International Conference on Pattern Recognition (ICPR), vol. 3, pp. 666-669 (2004)

[19] Yu, Y., Debevec, P., Malik, J., Hawkins, T. : Inverse global illumination: recovering reflectance models of real scenes from photographs, Proceedings of the 26th annual conference on Computer graphics and interactive techniques (SIGGRAPH), pp. 215-224 (1999)

[20] Tu, W.: Practical anti-vignetting methods for digital cameras, IEEE Transactions on Consumer Electronics, vol. 50, pp. 975- 983 (2004)

[21] Zheng, Y., Lin, S., Kambhamettu, C., Yu, J., Kang, S. B.: Single-Image Vignetting Correction, IEEE Trans. Pattern Anal. Mach. Intell., vol. 99, pp. 461-468 (2008)

[22] Zhang, R., Tsai, P.-S., Cryer, J. E., Shah, M.: Shape from Shading: A Survey, IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, pp. 690-706 (1999)