Single-Image Estimation of the Camera Response Function in Near-Lighting

Pedro Rodrigues[1], João P. Barreto[1,2]
[1]Institute of Systems and Robotics, University of Coimbra.
[2]Department of Electrical and Computer Engineering, Faculty of Sciences and Technology, University of Coimbra.

Radiometric calibration is a classical problem in computer vision. However, the estimation of the CRF of medical endoscopes poses challenges that are not properly addressed by the current state-of-the-art. Assumptions on the position and/or isotropy of the illuminant fall short on accuracy, making the vast majority of methods not applicable. Moreover, the calibration procedure must be fast, robust, and require minimal user intervention to be carried by the surgeon without disturbing the existing clinical routine.

With this work, we describe a method for estimating the CRF from a single image of a surface with regions of two colors for which the albedo ratio is known in advance. The algorithm makes no assumptions about the illumination conditions or the vignetting. Although the shape of the calibration surface is irrelevant for the CRF estimation, the experiments were carried using a planar checkerboard target similar to the one described in [1]. The reasons are twofold: the automatic segmentation and identification of the two color regions is straightforward, and by combining our radiometric calibration method with the geometric calibration approach proposed in [2] it is possible to fully model the camera from a single calibration frame.

Other applications include indoor calibration and the calibration of smartphone cameras and generic cameras with strong vignetting.

The radiometric image formation model can be written in a generic form as

$$f^{-1}(d(\mathbf{x})) = \alpha\rho(\mathbf{x})m(\mathbf{x})q(\mathbf{x}) \tag{1}$$

where f is the CRF, d is the image, $\mathbf{x}$ is a scene point, α is the exposure, m is the vignetting, ρ is the albedo, and q is the light intensity on the scene.

Following a similar reasoning to the work of Wu *et al.* [3], we will look for pairs of pixels with different albedos for which we know that the vignetting and the light effect can be cancelled. In fact, we do not need to be invariant to both effects, as in [3], only to their joint effect.

Let $u(\mathbf{x}) = \alpha m(\mathbf{x})q(\mathbf{x})$ be the irradiance normalized by albedo (INA). On the ith isovalue line of the INA

$$u(\mathbf{x}_j) = \kappa_i = \frac{f^{-1}(d(\mathbf{x}_j))}{\rho(\mathbf{x}_j)} \tag{2}$$

$\forall j \in \mathcal{L}_i$, where $\mathcal{L}_i$ is the set of pixels crossed by isovalue line i and κ_i is a constant and the value of $u(\mathbf{x})$ on the isoline. Thus, if a line i passes through multiple albedos (ρ_1 and ρ_2 as an example), from (2), one will have

$$f^{-1}(d(\mathbf{x}_j)) - \frac{\rho_1}{\rho_2} f^{-1}(d(\mathbf{x}_k)) = 0 \tag{3}$$

$\forall j \in \mathcal{L}_i \cap \mathcal{A}_{\rho_1}, \forall k \in \mathcal{L}_i \cap \mathcal{A}_{\rho_2}$, with an equation for each pair of albedos on each isoline. $\mathcal{A}_{\rho_n}$ is the set of points with a specific albedo $\rho(\mathbf{x}) = \rho_n$. This equation will be used for the single-image CRF estimation.

However, $u(\mathbf{x})$ is not known, and the isolines must be found on the image space $d(\mathbf{x})$. From (2) it is known that, for a given albedo, an isoline in the sensor irradiance is also an isoline in the image $d(\mathbf{x})$. In addition, along an isoline of $u(\mathbf{x})$, $\mathcal{L}_i$, the image values form a piecewise constant function (with a different constant value for each albedo). Figure 1 shows each image component individually.

In the image space we have a set of isolines for each albedo. However, the isolines of $d(\mathbf{x})$ for each albedo are the same, and equal to the isolines of $u(\mathbf{x})$, except for its value (figure 1).

To find the isolines of $u(\mathbf{x})$, let us model the image along each albedo ρ_n as a generic model h_n where the isolines are known

$$d(\mathbf{x}_j) \approx h_1(\mathbf{x}_j) \tag{4a}$$
$$d(\mathbf{x}_k) \approx h_2(\mathbf{x}_k) \tag{4b}$$

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage. The authors want to thank QREN-Mais Centro for funding under the research project "Novas Tecnologias para Apoio à Saúde e Qualidade de Vida, Projecto Cirurgia e Diagnóstico Assistido por Computador Usando Imagem".

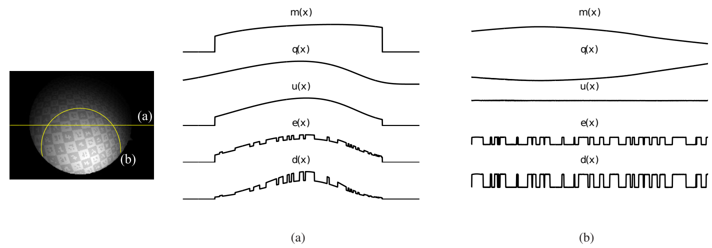

Figure 1: Values (arbitrary units) of each image component showing the formation process of an image of a checkerboard along (a) a horizontal line and (b) an isoline of the irradiance normalized by albedo.

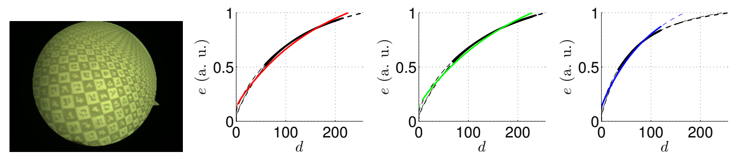

Figure 2: Example of a CALTag grid image acquired with an endoscopic rig, and the resulting CRFs for the red, green, and blue channels (from left to right). Each plot shows the result for the Wu *et al.* approach in color, and the our results for five CALTag images.

$\forall j \in \mathcal{A}_{\rho_1}, \forall k \in \mathcal{A}_{\rho_2}$. We know that the isolines of $h_1(\mathbf{x})$ and $h_2(\mathbf{x})$ will have the same shape as the ones in $u(\mathbf{x})$ but with different values. The shape of the surfaces represented by the models are different, since the step between isolines varies from one formulation to the other, but the isolines are the same. If we consider one of the albedos, let us say ρ_1, as a reference we can show that the two models are related by

$$h_1(\mathbf{x}_k) = f\left(\frac{\rho_1}{\rho_2} f^{-1}(h_2(\mathbf{x}_k))\right) = g(h_2(\mathbf{x}_k)) \tag{5}$$

$\forall k \in \mathcal{A}_{\rho_2}$, and thus

$$g(d(\mathbf{x}_k)) \approx h_1(\mathbf{x}_k). \tag{6}$$

where g is a positive and monotonically increasing function that is used to transform the model h_2 into the model h_1. This is the equivalent of having a gain for each isoline for the points of the albedo ρ_2, to be able to use only the model h_1 (relative to the albedo ρ_1) for both albedos. The isovalue lines will then be extracted as the level sets of the model h_1, and, having the isolines, equations of the same form as (3) can be used to find the CRF. Results are shown in figure 2.

The implementation of the estimation and the white balance procedure (for RGB images) is detailed in the paper. Unlike other single-image calibration algorithms that require the detection of regions where light is constant on the scene and no vignetting is present, our approach benefits from the effects of near-lighting and vignetting to perform the estimation.

This approach is scalable. Additional images would provide additional equations that can be grouped together to improve the estimation. Even with different exposures, different poses, and changes in the vignetting.

[1] B. Atcheson, F. Heide, and W. Heidrich. CALTag: High precision fiducial markers for camera calibration. *Vision, Modeling, and Visualization*, pages 41–48, 2010.

[2] R. Melo, J. P. Barreto, and G. Falcão. A new solution for camera calibration and real-time image distortion correction in medical endoscopy-initial technical evaluation. *IEEE TBME*, 59(3):634–44, 2012.

[3] C. Wu, S. G. Narasimhan, and B. Jaramaz. A multi-image shape-from-shading framework for near-lighting perspective endoscopes. *IJCV*, 86 (2-3):211–228, 2010.