https://doi.org/10.2352/issn.2169-2629.2021.29.264


# Reflectance estimation from snapshot multispectral images captured under unknown illumination

***Vlado Kitanovski, Jean-Baptiste Thomas, Jon Yngve Hardeberg***
***Department of Computer Science, Norwegian University of Science and Technology, Gjøvik, Norway***

## Abstract

*Multispectral images contain more spectral information of the scene objects compared to color images. The captured information of the scene reflectance is affected by several capture conditions, of which the scene illuminant is dominant. In this work, we implemented an imaging pipeline for a spectral filter array camera, where the focus is the estimation of the scene reflectances when the scene illuminant is unknown. We simulate three scenarios for reflectance estimation from multispectral images, and we evaluate the estimation accuracy on real captured data. We evaluate two camera model-based reflectance estimation methods that use a Wiener filter, and two other linear regression models for reflectance estimation that do not require an image formation model of the camera. Regarding the model-based approaches, we propose to use an estimate for the illuminant's spectral power distribution. The results show that our proposed approach stabilizes and marginally improves the estimation accuracy over the method that estimates the illuminant in the sensor space only. The results also provide a comparison of reflectance estimation using common approaches that are suited for different realistic scenarios.*

## Introduction

Multispectral cameras based on the spectral filter array (SFA) technology are relatively new on the market [1]. They are snapshot compact devices that jointly sense spatio-spectral scene information and they have high framerates which makes them suitable for real-time machine vision applications. One of the main goals in a multispectral imaging application is recovering accurate spectral information of the objects in the scene. The spectral accuracy of an SFA camera is typically lower than that of a hyperspectral imaging system that provides standardized spectral data (e.g. reflectance factors), however, SFA cameras are faster and compact, which makes them a good choice in various image and video applications, such as: medical imaging [2], computer vision [3], or appearance measurement [4]. It is possible to estimate a standardized data representation from multispectral captures using computational techniques, such as multispectral constancy [5]. The key point of multispectral constancy is estimating the illuminant in the scene, so that it can be discounted from the captured data, i.e., its impact on the captured scene information can be removed. Thus, estimating the scene illuminant is an important part in the estimation of scene reflectances. It can be done explicitly, by estimating the illuminant in the camera sensor (multispectral) space [5-6] or its spectral power distribution (SPD) [7-8], or implicitly, by directly calibrating a transform between the camera captures under the specific illuminant and known scene reflectances.

In this paper, we implement an imaging pipeline for a visible range multispectral SFA camera and evaluate its performance on the estimation of reflectances in scenes where the illuminant is not known. We evaluate four reflectance estimation approaches that are suitable for different usage scenarios: the first two utilize a linear camera model, while the other two do not. The first approach is using a conventional Wiener filter that is trained on a canonical (default) illuminant and that first uses a diagonal transform to transform the camera data from the estimated to the canonical illuminant [9]. The second approach is also a Wiener filter, but it is trained on an estimated spectral power distribution (SPD) of the illuminant in the scene. We propose a linear regression model for estimating the illuminant's SPD from its sensor-space estimate, using camera sensitivities and illuminants database as prior information. The third approach is a linear regression model trained under a canonical illuminant using a set of know reflectances in the scene, which uses a diagonal transform to transform the data from the estimated to the canonical illuminant. The fourth approach is suitable for the best-case scenario where a linear regression model is trained during the capture using a set of known reflectances in the captured scene, such as a color checker.

These four reflectance estimation methods are evaluated using real camera captures in three usage scenarios when the illuminant is unknown: a most challenging scenario where there is no known reflectance in the scene, a scenario where there is a white or gray patch in the scene, and a scenario where we have a set of known reflectances in the scene. The number of evaluated combinations (between the three usage scenarios and the four reflectance estimation methods) is seven – as not all combinations are feasible. The evaluation shows that using the recovered illuminant's SPD stabilizes and slightly improves the accuracy of the model-based reflectance estimation using Wiener filter. The results also show that a linear regression model trained on a canonical illuminant and cascaded with a diagonal transform performs similar to the camera model-based Wiener estimators, and they also confirm that using more known reflectances in the scene significantly improves the reflectance estimation accuracy.

In the next sections, we first describe the proposed method for illuminant's SPD estimation, followed by the reflectance estimation methods we have used. The evaluation framework and the results are then presented, followed by conclusions at the end.

## Illuminant estimation

In the commonly used linear image formation model, the camera response of the $i$-th channel $c_i$, is calculated using the camera channel's spectral sensitivity $S_i(\lambda)$, the SPD of the illuminant $L(\lambda)$, the spectral reflectance of the captured object $R(\lambda)$, a normalization constant $k$, and with added dark current $b_i$ and random noise $\varepsilon_i$:

$$c_i = k \int_\lambda S_i(\lambda)L(\lambda)R(\lambda)d\lambda + b_i + \varepsilon_i \quad (1)$$

In practice, the integral in Eq. (1) is substituted with a finite sum using a fixed size for $\Delta\lambda$ ($\Delta\lambda$ =10nm in this work), and the dark current that is subtracted as the first processing step is excluded from the model. Then, the model can be written in discrete form:

$$\mathbf{c} = k\mathbf{SLr} + \boldsymbol{\varepsilon} \tag{2}$$

In Eq. (2), $\mathbf{c} \in [0,1]^m$ is the vector-column of the $m$ camera channel responses compensated for the dark current, $\mathbf{r} \in [0,1]^n$ is a vector-column of the $n$-bands captured reflectance, $\mathbf{L}$ is a $n \times n$ diagonal matrix containing the SPD of the illuminant, $\mathbf{S}$ is an $m \times n$ matrix containing the $m$ camera system sensitivities as row vectors, $k$ is a constant, and $\boldsymbol{\varepsilon}$ is the additive noise. The constant $k$ is a normalization constant that is used to equate the model values with the real camera sensor values, and it is used e.g., to model the variation of the camera responses with exposure or lens aperture.

For a given camera system with concrete capture settings and for a given object reflectance, the camera responses are mainly affected by the SPD of the illuminant. Since recovering the scene objects properties (e.g. reflectance, color) from the camera responses is a goal in most camera applications, the impact of the illuminant on the camera responses should be minimized, or ideally removed. This can be done by measuring the illuminant and/or calibrating the camera for that particular illuminant. When the illuminant in the scene is not known, it can be estimated from the captured image using algorithms such as max-spectral, spectral gray world, spectral shades of gray, or spectral gray edge [6]. These algorithms estimate the illuminant in the camera sensor space. Alternatively, the SPD of the illuminant can be estimated from the captured image, e.g., by estimating the most likely one from a dataset of known illuminants [10], using a generalized inverse [7], or using convolutional neural networks [8].

In this work, we propose a linear estimator of the illuminant's SPD based on regularized least squares optimization. We first estimate the illuminant in the sensor space using one of the existing methods, and then use the camera sensitivities and illuminants database as prior information to train a linear regression estimator for the illuminant's SPD. If the illuminant in the sensor space $\mathbf{l}$ is defined as the camera response to a perfect white diffuser $\mathbf{r}_\mathrm{w}$, then according to the Eq. (2) it is modelled as:

$$\mathbf{l} = k\mathbf{SL}\mathbf{r}_\mathrm{w} + \boldsymbol{\varepsilon} \tag{3}$$

The regularized least squares estimation for the illuminant's SPD $\hat{\mathbf{L}}$ from the sensor-space estimated illuminant $\hat{\mathbf{l}}$ is:

$$\hat{\mathbf{L}} = \mathbf{Q}_\mathrm{L}\hat{\mathbf{l}} = \mathbf{L}_\mathrm{t}\mathbf{L}_\mathrm{st}{}^\mathrm{T}\left(\mathbf{L}_\mathrm{st}\mathbf{L}_\mathrm{st}{}^\mathrm{T} + \lambda_\mathrm{L}\mathbf{E}\right)^{-1}\hat{\mathbf{l}} \tag{4}$$

In Eq. (4), $\mathbf{Q}_\mathrm{L}$ is the linear regression matrix calculated using column-arranged training illuminant SPDs $\mathbf{L}_\mathrm{t}$ and simulated (using Eq. (3)) sensor-space illuminants $\mathbf{L}_\mathrm{st}$. The parameter $\lambda_\mathrm{L}$ is a regularization constant, $\mathbf{E}$ is the identity matrix. For the calculation of the regression matrix $\mathbf{Q}_\mathrm{L}$, the columns in $\mathbf{L}_\mathrm{t}$ and $\mathbf{L}_\mathrm{st}$ are normalized so that their $\mathrm{L}^1$ norm is one, and consequently, the estimated sensor-space illuminant $\hat{\mathbf{l}}$ is normalized to a $\mathrm{L}^1$ norm of one before it is used for calculating $\hat{\mathbf{L}}$. The $\mathrm{L}^1$ norm of the estimated illuminant's SPD $\hat{\mathbf{L}}$ in general is not equal to one, and it can sometimes be negative. Therefore, the negative values in $\hat{\mathbf{L}}$ are clipped to zero and $\hat{\mathbf{L}}$ is normalized to a unit $\mathrm{L}^1$ norm. The estimation of the sensor-space scene illuminant $\hat{\mathbf{l}}$ can be done using commonly used methods; in this work, we use the spectral gray world [6] for the scenario where there is no reference patch in the scene, and we use the camera sensor values of the white patch for the scenario when a white patch is present in the scene.

## Reflectance estimation

In this work, we evaluate four different methods for estimating reflectance from the camera responses. The first one is based on a Wiener filter estimator $\mathbf{W}_\mathrm{C}$ trained on a canonical illuminant $\mathbf{L}_\mathrm{C}$ and a diagonal transform $\mathbf{D}$ that converts the camera responses $\mathbf{c}$ from the estimated scene illuminant $\hat{\mathbf{l}}$ to the canonical illuminant $\mathbf{l}_\mathrm{C}$ [9]:

$$\mathbf{D} = \mathrm{diag}\left(\mathbf{l}_\mathrm{C} \div \hat{\mathbf{l}}\right) \tag{5}$$

$$\hat{\mathbf{r}} = \mathbf{W}_\mathrm{C}\mathbf{D}\mathbf{C} = k\mathbf{R}\mathbf{R}^\mathrm{T}\mathbf{L}_\mathrm{C}^\mathrm{T}\mathbf{S}^\mathrm{T}\left(k^2\mathbf{S}\mathbf{L}_\mathrm{C}\mathbf{R}\mathbf{R}^\mathrm{T}\mathbf{L}_\mathrm{C}^\mathrm{T}\mathbf{S}^\mathrm{T} + \mathbf{K}_{\boldsymbol{\varepsilon}}\right)^{-1}\mathbf{D}\mathbf{c} \tag{6}$$

In Eq. (5), $\hat{\mathbf{l}}$ is the estimated scene illuminant in the sensor space, $\mathbf{l}_\mathrm{C}$ is canonical illuminant in the sensor space, and the operator "÷" denotes element-wise division. In Eq. (6), $\hat{\mathbf{r}}$ is the estimated reflectance, while $\mathbf{R}$ is a $n \times N$ matrix of $N$ training spectral reflectances. The $m \times m$ diagonal covariance matrix of the noise $\mathbf{K}_{\boldsymbol{\varepsilon}}$ is estimated from the difference between simulated and observed camera responses [11].

The second reflectance estimation method is also based on a Wiener estimator, $\mathbf{W}_\mathrm{E}$, but it is trained using the estimated illuminant's SPD:

$$\hat{\mathbf{r}} = \mathbf{W}_\mathrm{E}\mathbf{C} = k\mathbf{R}\mathbf{R}^\mathrm{T}\mathbf{L}_\mathrm{E}^\mathrm{T}\mathbf{S}^\mathrm{T}\left(k^2\mathbf{S}\mathbf{L}_\mathrm{E}\mathbf{R}\mathbf{R}^\mathrm{T}\mathbf{L}_\mathrm{E}^\mathrm{T}\mathbf{S}^\mathrm{T} + \mathbf{K}_{\boldsymbol{\varepsilon}}\right)^{-1}\mathbf{c} \tag{7}$$

The estimated illuminant $\mathbf{L}_\mathrm{E}$ in Eq. (7) is an $n \times n$ diagonal matrix containing the estimated illuminant's SPD $\hat{\mathbf{L}}$ that is calculated using Eq. (4). Note that these two reflectance estimation methods require the knowledge of the camera sensitivities.

The third reflectance estimation method is also a linear regression method, but it is based on regularized least squares optimization for a canonical illuminant:

$$\hat{\mathbf{r}} = \mathbf{Q}_\mathrm{C}\mathbf{D}\mathbf{C} = \mathbf{R}_\mathrm{t}\mathbf{C}_\mathrm{t}^\mathrm{T}\left(\mathbf{C}_\mathrm{t}\mathbf{C}_\mathrm{t}^\mathrm{T} + \lambda_\mathrm{C}\mathbf{E}\right)^{-1}\mathbf{D}\mathbf{c} \tag{8}$$

In Eq. (8), $\mathbf{R}_\mathrm{t}$ is an $n \times M$ matrix of $M$ training spectral reflectances, $\mathbf{C}_\mathrm{t}$ is an $m \times M$ matrix of $M$ captured sensor responses under the canonical illuminant, $\lambda_\mathrm{C}$ is a regularization constant, and $\mathbf{E}$ is the identity matrix. The regression matrix $\mathbf{Q}_\mathrm{C}$ can be trained using a color checker [12]. The diagonal matrix $\mathbf{D}$ is calculated using Eq. (5). Unlike the previous two methods, this method does not require the knowledge of the camera sensitivities.

The fourth reflectance estimation method is the baseline method which does not explicitly estimate the scene illuminant, but it relies on having a set of known reflectances, $\mathbf{R}_\mathrm{t}$ , in the captured scene. The captured multispectral values of these known reflectances, $\mathbf{C}_\mathrm{t}$, are then used to train a linear regression estimator based on regularized least squares optimization:

$$\hat{\mathbf{r}} = \mathbf{Q}\mathbf{C} = \mathbf{R}_\mathrm{t}\mathbf{C}_\mathrm{t}^\mathrm{T}\left(\mathbf{C}_\mathrm{t}\mathbf{C}_\mathrm{t}^\mathrm{T} + \lambda_\mathrm{C}\mathbf{E}\right)^{-1}\mathbf{c} \tag{9}$$

We used $M$ = 24 patches from the X-Rite Digital SG Color Checker (marked in red in Fig. 1) to train the regression matrices $\mathbf{Q}$ and $\mathbf{Q}_\mathrm{C}$. The regularization constant $\lambda_\mathrm{C}$ in Eqs. (8) and (9) is calculated as the average of the optimal regularization values for several different illuminants using leave-one-out cross-validation.

In this work, we consider three different scenarios for reflectance estimation. In the first scenario, there is no known reflectance in the scene so that the constant $k$ in Eq. (2) cannot be accurately estimated. Therefore, the value of $k$ can be chosen e.g.,

as the value for which the maximal sensor response to a perfect white diffuser equals a certain value (0.8 in our work). For this scenario, the first three reflectance estimation methods (using Eqs. (6)-(8)) are evaluated. The illuminant in the sensor domain $\hat{\mathbf{l}}$ is estimated using the spectral gray world algorithm, and it's SPD $\hat{\mathbf{L}}$ is estimated using Eq. (4). In the second scenario, there is one known reflectance in the scene - a white (or gray) patch. We use the known reflectance to estimate the constant $k$, and also to estimate the illuminant in the sensor space $\hat{\mathbf{l}}$ and afterwards its SPD $\hat{\mathbf{L}}$ using Eq. (4). For this scenario, we also evaluate the first three reflectance estimation methods (using Eqs. (6)-(8)). In the third scenario, there is a set of known reflectances in the scene, e.g., the 24 patches from a classic color checker. This simulates a scenario where the use of a color checker is a viable option, and therefore we evaluate only the fourth reflectance estimation method (using Eq. (9)). With these three scenarios, the number of evaluated reflectance estimations in this work is seven.

## Evaluation

### *Imaging pipeline for SFA camera*

We used the 1.3MP Silios CMS-C multispectral camera [13] with a 3×3 filter array that captures one panchromatic band and eight narrow bands in the visible range, centered at: 440 nm, 473 nm, 511 nm, 550 nm, 587 nm, 623 nm, 665 nm, and 703 nm. The camera sensitivities were measured by the manufacturer. Using the IDS software suite for uEye cameras [14], we implemented an imaging pipeline that follows generally common steps [15-16]. These are briefly described next in the order in which they are performed.

In the first stage of setting up the camera, the lens aperture and focus are adjusted to a satisfying level in terms of image sharpness and depth of field. Subsequently, the exposure time is adjusted so that the dynamic range is well utilized for the target capture objects.

For the given settings and with closed camera lens, a dark current image is estimated as the median of $d$ successive captures. The dark current image is then subtracted from any captured raw image with the same settings. In this work, we used $d$=25.

The sensor is set to a state where it is most linear with incident irradiance. This included adjusting the sensor black level, setting the camera gain as well as the sensor gamma to 1, and switching off other default processing on the sensor chip. The camera response function is then estimated from multiple captures with different exposures, and it used to obtain a look-up table for mapping the sensor values to quantities proportional to the incident irradiance.

In an imaging system, there are many different sources of noise, coming from the light itself (e.g. photon-shot noise or time-varying light) and from the imaging system (e.g. readout noise, hot pixels, or reset noise). Therefore, we average $f$ input raw frames before proceeding further. The high framerate of SFA cameras makes this operation very time efficient. In this work, we set $f$=10.

In this work, we use a computationally efficient demosaicing algorithm – the bilinear interpolation. While it is known that the perceptual quality of the demosaiced image obtained using bilinear interpolation is far from optimal, we find it good enough for surfaces with low spatial variation of the reflectance – such as the captured color checker patches that are used in this work.

To compensate for the spatially varying illumination as well as for the spatially varying sensitivity of the camera (coming from the sensor or the lenses), we flat-field each of the multispectral channels separately. In this work, the flat-field mask for each of the channels is calculated using the X-Rite ColorChecker® White Balance target [12], and it is normalized to a small region in the center of the image. The mask is then used to divide the input multispectral channels. The output of this step is the flat-fielded multispectral image, or as often referred in this paper, the camera responses.

### *Experimental setup*

The evaluations of the illuminant and reflectance estimation methods are performed using multispectral images obtained from the described pipeline. The estimation methods were evaluated on the captured X-Rite Digital SG ColorChecker®, specifically on the 72 patches that exclude the border neutral patches as well as the 24 patches from the classic color checker that were used for training. Figure 1 illustrates one captured scene, together with markings for the training and testing patches. The ground-truth reflectances of the Digital SG color checker were obtained using X-Rite i1 Pro 2 spectrophotometer, which uses 45°/0° measurement geometry. We used six illuminants in our work that cover a variety in terms of shape of their SPD: four illuminants (Daylight, A, Cool White, TL84) from the viewing booth GretagMacbeth Spectralight III, and two LED illuminants using the PIXEL K80 RGB LED lighting system with CCT of 2600K and 6500K, respectively. We measured the six illuminant's SPDs using a Konica-Minolta CS-2000 spectroradiometer; they are shown in Fig. 2. We chose to use the Daylight illuminant as the canonical illuminant in this work because it has the most uniform SPD of all the six. We trained the Wiener filter estimators in Eqs. (6) and (7) using the Munsell chips reflectances [17]. As training illuminants data in Eq. (4), we used the SFU dataset of 102 illuminants [18]. The camera capture geometry was set to be as close as possible to the 45°/0° geometry used for measuring the ground-truth: the camera viewing angle was 0°, while the direct component of the six light sources was set to be around 45° relative to the Digital SG color checker plane.

The accuracy of illuminant and reflectance estimation can be evaluated using different objective metrics. For the estimated illuminants we use the ΔA angle (also known as spectral angle) between the two row-vectors representing the ground-truth $\mathbf{g}$ and the estimation $\hat{\mathbf{g}}$:

$$\Delta A = arccos\left(\frac{\mathbf{g}\hat{\mathbf{g}}^{\mathrm{T}}}{\sqrt{(\mathbf{g}\mathbf{g}^{\mathrm{T}})(\hat{\mathbf{g}}\hat{\mathbf{g}}^{\mathrm{T}})}}\right) \tag{10}$$

For the estimated reflectances we use three different metrics: the root mean square error (RMSE), the cosine distance (calculated as 1-cos(ΔA)), and the $\Delta E_{00}$ color difference metric.

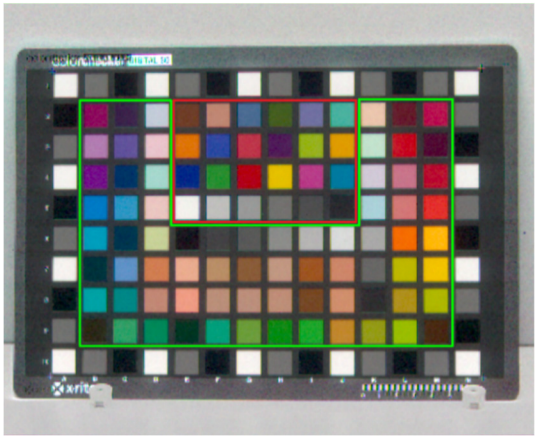

*Figure 1. sRGB rendering of the estimated scene reflectances from the multispectral image captured under the Cool White illuminant. Additionally, the training 24 patches used in this work are marked with the red box, while the testing 72 patches are inside the marked green area.*

## *Results*

We first provide results for the illuminant estimation using the case of no reference patch in the scene. In this case, the illuminant's SPD is estimated using the proposed method from the sensor domain illuminant estimate. The estimation errors of all the six illuminants, in terms of spectral angular error are given in Table 1. It can be seen that the spectral gray world algorithm estimates the illuminants in the sensor domain with relatively high accuracy – this is expected since the scene is dominated by the Digial SG color checker and relatively large areas of spectrally flat reflectance. The average angular error for the estimated SPDs, excluding the LED illuminants that are not included in the training SFU dataset, is around 9° - which can be considered as a reasonably good estimation. The results point out that the SPD estimation can be improved by only using training illuminants in Eq. (4) similar to the ones that are expected during capture. The estimated SPDs, from the spectral gray-world sensor estimates (no reference patch scenario), are shown in Fig. 2.

Regarding the reflectance estimation, the provided results are for five illuminants – the Daylight illuminant is excluded because it is used as a canonical illuminant for some of the evaluated methods. Furthermore, we evaluate seven reflectance estimations: three using Eqs. (6-8) in a 'no reference patch' scenario, the same three but in a 'white patch in the scene' scenario, and one estimation using Eq. (9) that uses a color checker (CC) in the scene, which corresponds to the best-case 'set of known reflectances in the scene' scenario. The average estimation errors under all five illuminants is given in Table 2, while the box plots for the per-illuminant estimations are shown in Fig. 3. An example showing the sRGB visualization of the estimated scene reflectances in the best-case scenario, under the Cool White illuminant, is shown in Fig. 1.

**Table 1: Angular error (°) for the estimated illuminants**

| | Daylight | A | LED 2600K | LED 6500K | Cool White | TL84 |
|---|---|---|---|---|---|---|
| Sensor domain | 1.83 | 1.51 | 1.96 | 2.16 | 2.13 | 2.40 |
| SPD | 8.20 | 9.87 | 15.49 | 21.52 | 8.88 | 8.93 |

**Table 2: Overall average reflectance estimation error for the different methods (referred by their estimation matrix)**

| | No reference patch scenario | | | White patch in the scene scenario | | | CC |
|---|---|---|---|---|---|---|---|
| estimation matrix | $W_cD$ | $W_e$ | $Q_cD$ | $W_cD$ | $W_e$ | $Q_cD$ | Q |
| RMSE | 0.096 | 0.079 | 0.072 | 0.063 | 0.051 | 0.056 | 0.021 |
| Cosine distance | 0.025 | 0.017 | 0.023 | 0.022 | 0.017 | 0.021 | 0.006 |
| $\Delta E_{00}$ | 8.95 | 7.28 | 7.26 | 7.32 | 5.94 | 6.79 | 2.49 |

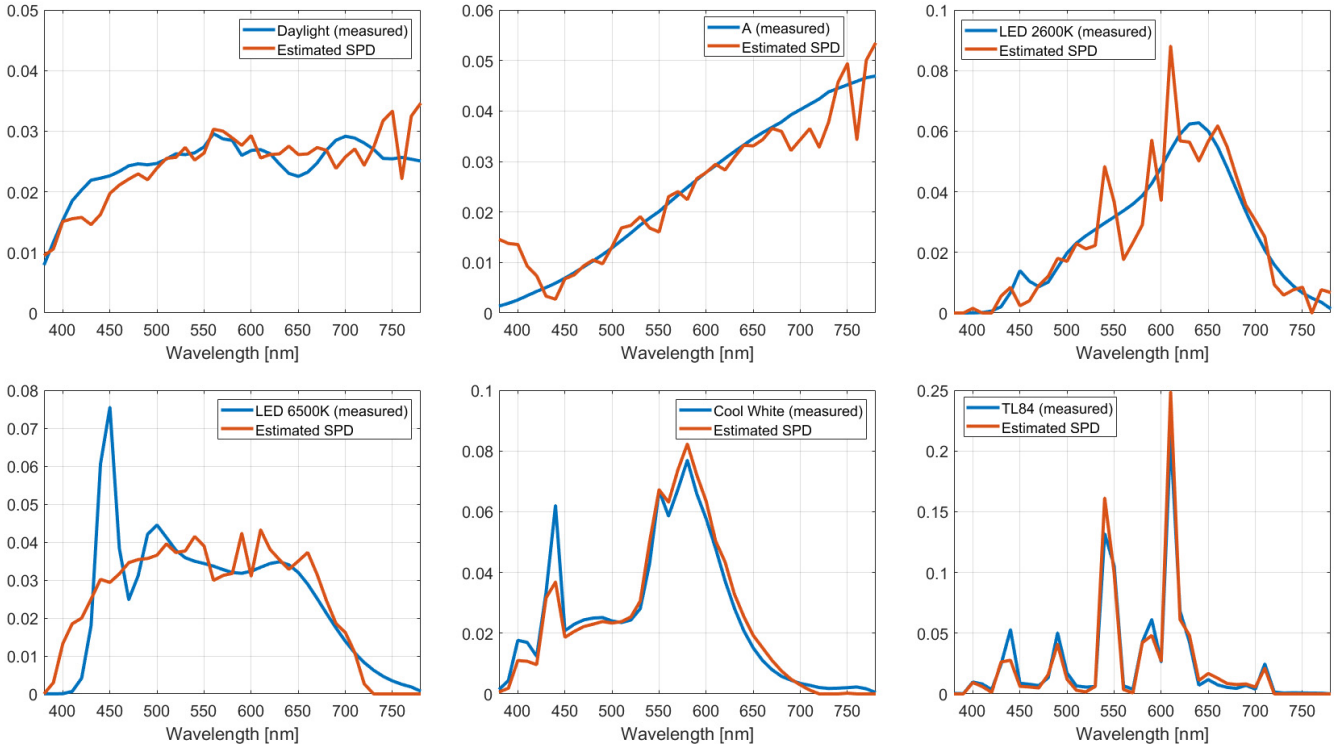


*Figure 2. Measured and estimated SPD of all the six illuminants used on this work, normalized to a unit $L^1$ norm.*

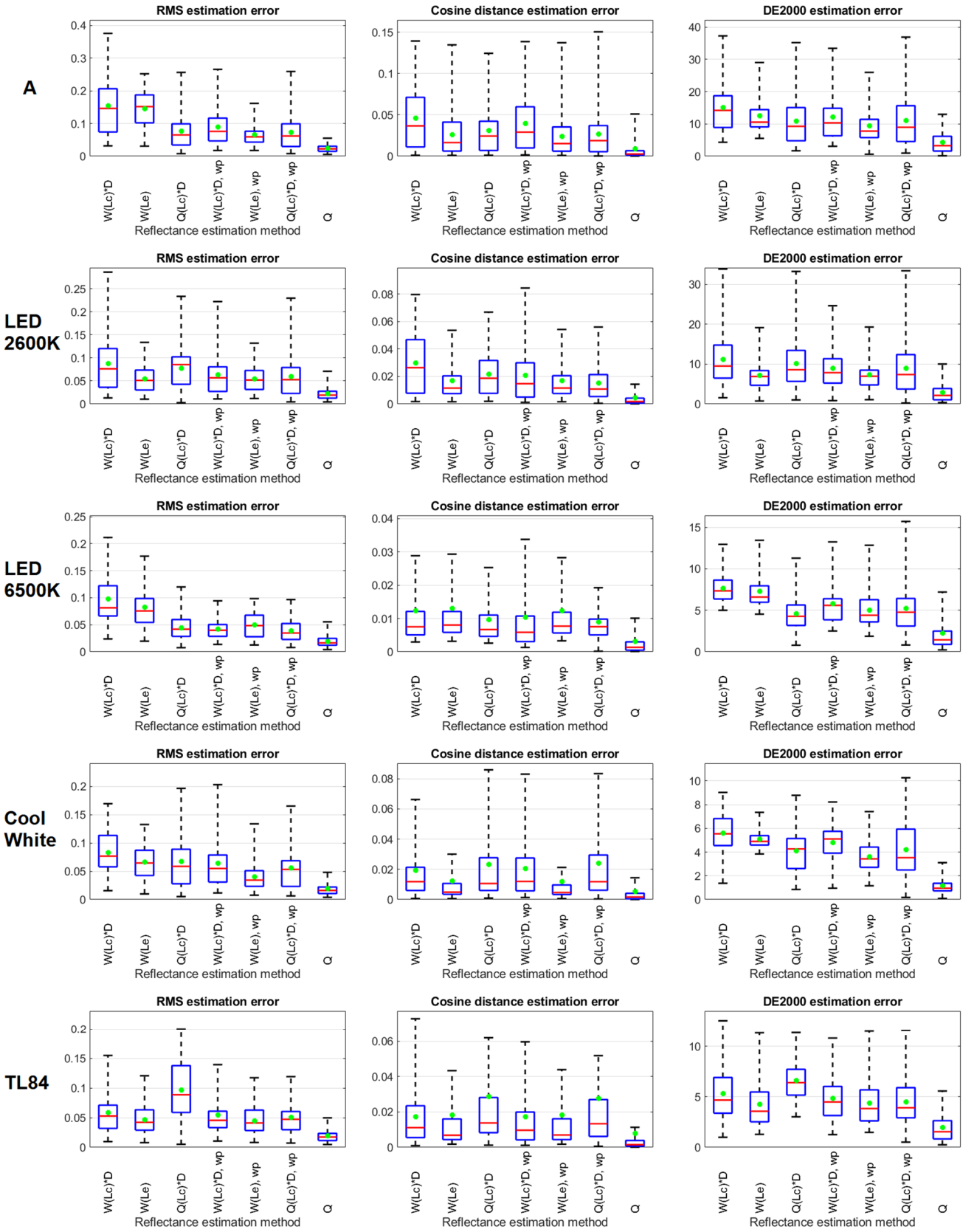


*Figure 3. Box plots of the reflectance estimation for the seven evaluated cases under different scene illuminants. The median values are shown with red lines, the mean values with green dots, and the 25%-75% percentile range with blue rectangles. From top to bottom: A, LED 2600K, LED 6500K, Cool White, and TL84.*

Several observations can be made from Table 2. First, from the comparison between the 'no reference patch' and the 'white reference patch' scenarios, it can be seen that the estimated reflectances by utilizing the white patch are more accurate in terms of RMSE, but not so much in terms of cosine distance. This can be attributed to the more accurate estimation of the constant $k$ in the presence of a white patch in the scene, which decreases the absolute RMS error, but it does not have significant impact on the estimated reflectance shapes – to which the cosine distance is more sensitive. It also should be noted that the used value for $k$ in the 'no reference patch' scenarios was close to the optimal one for some of the five captured scenes; using a value that is very different from the optimal will lead to worse estimations, especially in terms of RMSE. Second, it can be seen that in both of these scenarios, the Wiener estimator based on the estimated illuminant's SPD on average performs better than the Wiener estimator based on the canonical illuminant and a diagonal transform for compensating the scene illuminant; this is the case for all three objective metrics. Third, the regression-based estimator, which does not use a model of the camera and which is trained on a canonical illuminant and therefore utilizes a diagonal transform, on average is ranked in the middle between the two Wiener estimators. This shows that a competitive reflectance estimation can be performed without using the camera sensitivities, which can be beneficial in certain scenarios. As last, the regression-based estimator trained on the color checker present in each scene performed significantly better than the others, and this is the preferred method for reflectance estimation whenever a set of known reflectance targets can be present in the scene.

The above points can be made from observing the reflectance estimation results in Fig. 3, where the performance of each of the three metrics on the seven estimations is presented separately for each of the five illuminants (not including the canonical 'Daylight' illuminant), in the form of box plots. The 25%-75% percentile range is shown with blue rectangles, with the median and mean value shown with a red line and a green dot, respectively. As an additional insight from Fig. 3, it can be seen that in most of the cases and for all metrics, the Wiener estimator trained using the estimated illuminant's SPD results in more stable (less varying) estimations when compared to the four estimators that utilize a canonical illuminant and a diagonal transform. Given that the spectral gray world algorithm resulted in fairly accurate estimation of the illuminants in the sensor space – nearly the same as in the case when a white patch was used, an interesting issue to be investigated in a future work is the impact of the accuracy of the sensor-space estimated illuminant on the estimated SPD, and furthermore, on the estimated reflectance.

## Conclusions

In this paper, we presented the imaging pipeline we have implemented for a visible-range multispectral SFA camera. The focus was put on evaluating the pipeline for estimating the captured scene reflectances. We proposed and evaluated a linear estimator for the illuminant's SPD from the estimated or white-patch measured illuminant in the sensor-space. This estimator is trained using the camera sensitivities and a database of illuminants as prior information We considered different linear reflectance estimation methods that are suitable for use in different scenarios regarding the information or equipment that is available in the modelling stage and during the scene capture. The evaluation of the different reflectance estimators showed that estimating the illuminant's SPD brings an improved accuracy and more stable estimations compared to the methods relying on a diagonal transform for discounting the scene illuminant. The results also showed that a trained linear regression model is competitive to the camera model -based Wiener estimators. Confirming these results using real scene objects under more illuminants is a potential direction for future work. Another direction would be introduction of an HDR workflow in the imaging pipeline, which could extend the scope of use of SFA cameras as reflectance measurement devices.

## Acknowledgements

We thank Spektralion AS for providing the software tool used for capturing the multispectral images.

## References

[1] P.J., Lapray, X. Wang, J.B. Thomas, P. Gouton, "Multispectral Filter Arrays: Recent Advances and Practical Implementation", MDPI Sensors, 14, pp. 21626–21659, 2014.

[2] T. Uemori, A. Ito, Y. Moriuchi, A. Gatto, J. Murayama, "Skin-based identification from multispectral image data using CNNs", in Conf. on Comp. Vision and Pattern Recog. (CVPR), pp. 12342-12350, 2019.

[3] A. Nourizonoz, R. Zimmermann, C.L.A. Ho, S. Pellat, Y. Ormen, C. Prevost-Solie, G. Reymond, F. Pifferi, F. Aujard, A. Herrel, D. Hubber "EthoLoop: automated closed-loop neuroethology in naturalistic environments", Nature Methods 17, pp. 1052-1059, 2020.

[4] V. Kitanovski, J.Y. Hardeberg, "Objective evaluation of relighting models on translucent materials from multispectral RTI images", Proc. IS&T Electronic Imaging, Material Appearance, pp. 133-1 - 7, 2021.

[5] M. Mosny and B. Funt, "Multispectral colour constancy," in Color and Imaging Conference, Society for Imaging Science and Technology, Vol. 2006, pp. 309–313, 2006.

[6] H.A. Khan, J.B. Thomas, J.Y. Hardeberg, and O. Laligant, "Illuminant estimation in multispectral imaging", J. Opt. Soc. Am. A 34, pp. 1085–1098, 2017.

[7] S. Ratnasingam, J. Hernandez-Andres, "Illuminant spectrum estimation at a pixel", J. Opt. Soc. Am. A 28, pp. 696-703, 2011.

[8] R. Deeb, J. Van de Weijer, D. Muselet, M. Hebert, A. Tremeau, "Deep spectral reflectance and illuminant estimation from self-interreflections", J. Opt. Soc. Am. A 36, pp. 105-114, 2019.

[9] H.A. Khan, J.B. Thomas, J.Y. Hardeberg, and O. Laligant, "Multispectral camera as spatio-spectrophotometer under uncontrolled illumination", Opt. Express, 27, pp. 1051-1070, 2019.

[10] R. Shrestha and J.Y. Hardeberg, "Spectrogenic imaging: A novel approach to multispectral imaging in an uncontrolled environment," Opt. Express 22, pp. 9123–9133, 2014.

[11] H.L. Shen, P.Q. Cai, S.J. Shao, and J.H. Xin, "Reflectance reconstruction for multispectral imaging by adaptive wiener estimation," Opt. Express 15, pp. 15545–15554, 2007.

[12] Color checker targets, available online. Accessed on 08.06.2021. https://www.xrite.com/categories/calibration-profiling/colorchecker-targets

[13] Silios multispectral cameras, available online. Accessed on 08.06.2021. https://www.silios.com/cms-series

[14] IDS Software suite, available online. Accessed on 08.06.2021. https://en.ids-imaging.com/ids-software-suite.html

[15] R. Ramanath, W.E. Snyder, Y. Yoo, M.S. Drew, "Color image processing pipeline", IEEE Signal Process. Mag., 22, pp. 34–43, 2005.

[16] P.J. Lapray, J.B. Thomas, P. Gouton, "High Dynamic Range Spectral Imaging Pipeline for Multispectral Filter Array Cameras", Sensors 17, 1281, 2017.

[17] Munsell colors matt, available online. Accessed on 08.06.2021. https://sites.uef.fi/spectral/munsell-colors-matt-spectrofotometer-measured/

[18] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research", Color Res. Appl. 27, pp. 147-151, 2002.