Combining transverse field detectors and color filter arrays to improve multispectral imaging systems

Miguel A. Martínez,[1,*] Eva M. Valero,[1] Javier Hernández-Andrés,[1] Javier Romero,[1] and Giacomo Langfelder[2]

[1]Color Imaging Laboratory, Department of Optics, Faculty of Sciences, University of Granada, Spain

[2]Politechnic University of Milano, Italy

*Corresponding author: m.martinez.domingo@gmx.com

Received 3 December 2013; revised 19 March 2014; accepted 19 March 2014; posted 19 March 2014 (Doc. ID 201775); published 16 April 2014

This work focuses on the improvement of a multispectral imaging sensor based on transverse field detectors (TFDs). We aimed to achieve a higher color and spectral accuracy in the estimation of spectral reflectances from sensor responses. Such an improvement was done by combining these recently developed silicon-based sensors with color filter arrays (CFAs). Consequently, we sacrificed the filter-less full spatial resolution property of TFDs to narrow down the spectrally broad sensitivities of these sensors. We designed and performed several experiments to test the influence of different design features on the estimation quality (type of sensor, tunability, interleaved polarization, use of CFAs, type of CFAs, number of shots), some of which are exclusive to TFDs. We compared systems that use a TFD with systems that use normal monochrome sensors, both combined with multispectral CFAs as well as common RGB filters present in commercial digital color cameras. Results showed that a system that combines TFDs and CFAs performs better than systems with the same type of multispectral CFA and other sensors, or even the same TFDs combined with different kinds of filters used in common imaging systems. We propose CFA+TFD-based systems with one or two shots, depending on the possibility of using longer capturing times or not. Improved TFD systems thus emerge as an interesting possibility for multispectral acquisition, which overcomes the limited accuracy found in previous studies.

OCIS codes: (110.4234) Multispectral and hyperspectral imaging; (110.0110) Imaging systems; (330.1690) Color; (330.1730) Colorimetry.

http://dx.doi.org/10.1364/AO.53.000C14

1. Introduction

Spectral science has been receiving gradually increased attention in the last few decades. New applications come up on a daily basis, offering an interesting range of possibilities [1–3]. Industrial, medical, military, remote sensing, and many more fields of research, focus nowadays on the use of spectral sciences for their new technological advances. One of the outstanding disciplines in the spectral science field is spectral imaging. It is mainly concerned with the problem of obtaining pixel-wise spectral information in an image. Many types of systems are found in the literature for this purpose. We could divide the spectral imaging devices into two main groups: measuring devices and estimating devices. The first group is made up mainly of hyperspectral devices, which acquire full spectral information during the capturing process. In this group, we can find systems formed by monochrome cameras attached to tunable filters [4], such as acousto-optic tunable filters (AOTFs) [5,6] or liquid-crystal tunable filters (LCTFs) [7]. We also find Bragg-grating-based systems [8], or snapshot systems [9] used in spectral microscopy [10]. The clear disadvantage of this group

1559-128X/14/130C14-11$15.00/0

of systems is either their bulky size (usually coupled to long capture times, since they require as many shots as spectral bands), or else a reduced spatial resolution, which is clearly not enough for most applications. In the second group, we find numerous systems that can acquire three or more spectral bands (commonly termed multispectral imaging devices), which are used for estimation. These systems usually have a lower number of bands than those in the first group and their spectral sensitivities are not so spectrally narrow-banded. Such devices acquire the sensor responses and then require different techniques to map these responses onto spectra [11]. The quality of the estimated spectra depends critically on the number of channels, their spectral shape, and the algorithms used to obtain the mapping from sensor responses to full spectra. Within estimation capturing systems, we find as representative instances: monochrome cameras with a color filter wheel with few filters [12], multi-sensor cameras [13–15], color filter array (CFA)-based cameras [16,17], and hybrid systems [18]. Some of these systems suffer from relatively long capture times and need the mechanical movement of parts that might cause misalignment problems (e.g., filter wheel devices). Others have expensive hardware, including many sensors and light splitting optical components, which suffer from low-light throughput, increased exposure times, limited portability, etc. Given all these problems, it is clear that the question of how to design a capturing system that is portable, nonexpensive, easy to use, and capable of acquiring full spectral information, pixel-wise, in real-time is still far from solved. In this study, we present a proposal for a multispectral imaging system based on a recently developed technology as a potential answer to this question. The transverse field detector (TFD) technology definitely has the advantage of allowing for the capture of multiple channels at a pixel with full spatial resolution. However, the TFD by itself does not provide enough accuracy in the estimated spectra [19–21]. In [22] we pointed out the hypothesis that one of the main causes of this decrease in accuracy is the width of the spectral response functions typical of TFDs. This width is such that we deem that conventional sharpening techniques are not enough to solve the problem [23]. Instead, in this paper, we propose combining the TFD sensors with a specifically designed CFA [17] to obtain the required accuracy in the spectral estimation. The filters narrow down the spectral sensitivities of TFDs and improve their color and spectral estimation performance, trading off their full spatial resolution property. We carried out different experiments to check if the improvement of performance is due to combining TFDs and CFAs. We also checked whether our method for selecting the filters to be used was effective for selecting the combinations of filter time sensitivities (what we call channels).

Using simulated sensor responses, we compared this system with several others. We present convincing evidence in Section 5 of the advantages of such a design, even allowing for the inevitable need to sacrifice the full spatial resolution feature of conventional TFD-based systems over the ones presented in previous works [21,22]. We offer some alternative designs depending on the requisites that each individual application may require, regarding capturing times, colorimetric, and spectral accuracy.

The remainder of this paper is organized as follows: in Section 2 we present the TFD sensors. Section 3 describes all the systems simulated and compared in this study. In the methods section (4), we clarify the filter selection procedure, the computations of sensor responses, the estimation procedure and the experiments performed. In Section 5 we provide the results found in the different experiments. Finally, Section 6 draws relevant conclusions from these results.

2. TFD Sensors

In silicon sensors, light is absorbed following an exponential intensity decrease, described by the Beer–Lambert law. Electron-hole pairs are photogenerated with different depth profiles, according to the material and the radiation energy (or wavelength). In a typical pn-junction-based CMOS pixel, photogenerated electrons are collected by the same well, regardless of the generation depth [24]. Thus, for each pixel, only one value is read in a single exposure. In contrast to these normal sensors, others, such as filter-less sensors like Foveon X3 [25] or TFDs [19], account for the fact that photons are able to penetrate deeper into the silicon as their wavelength increases. Therefore, electrons generated at different depths can be separated by junctions built at different depths (as in the case of Foveon sensors), or by suitable transverse electric fields (as in the case of TFDs). Understanding this principle makes it possible to set the collection of generated electrons at different depths within the silicon, so that in a single pixel, and in a single exposure, we can retrieve information from three (Foveon) and up to five (TFDs) spectral bands or channels. Usually, in a 3-channel configuration, the responsivities have peaks in the short, middle, and long wavelength ranges, similar to standard RGB sensors; however, they are spectrally wider since there is no sharpening filter in front of the sensor. Furthermore, TFDs also offer the capability of modifying the collection depth of photons by modifying the applied transverse electric field via a tunable biasing voltage. Thus, using a given biasing voltage, the sensor presents a set of three sensitivities with certain spectral shape and maxima positions, and setting a different value for this voltage, the spectral shapes and positions of the sensitivities change. We can see this in system number 8 (Fig. 2). In this figure, we present eight different sets of RGB sensitivities corresponding to eight different values of biasing voltage. The difference in the spectral shapes and peaks is quite clear. The possibilities offered by this technology are many,

and few studies have so far been done to examine the capability of TFDs for multispectral imaging [21,22]. As pointed out before, TFD sensitivities are too spectrally broad to provide satisfactory accuracy in the estimated spectra within the context of 3-channel noisy systems. This is a drawback when trying to use them for spectral imaging [26], because even when adding more shots corresponding to different biasing voltages (up to 8, with a total of 24 channels), the results were equaled or outperformed by a simple RGB system based on CFAs [22], where spectral sensitivities were much narrower in comparison.

In previous studies, many authors have reported good estimation accuracies obtained with both simulated and real camera responses from narrow spectral responsivities [27] and, more recently, some have proposed the possibility of increasing this accuracy by using tunability of the channels' peak response positions and widths with simulated camera responses of Gaussian-shaped channels [28]. Nevertheless, narrow-band tunable responses are not so far achievable in practice without the use of external filters. Nonetheless, TFDs have the advantage of not needing any demosaicking process to get the full spatial information for every spectral band. This prompted an idea. Does adding filters embedded in the TFD sensor matrix (and so combining CFAs with TFDs) boost the estimation accuracy of the system? We present in the next sections a set of experiments comparing the color and spectral accuracy of different capturing system designs, to provide clear evidence on these points. We have assessed the color and spectral reconstruction accuracy of a set of reflectance samples with no spatial information present. Therefore, we did not study the effect of varying the spatial arrangement of the different transmittances present in the CFA pattern. This factor has been shown to influence image quality [29]. Other studies have been made using similar multispectral CFAs with up to seven channels, providing methods to design their spatial arrangement, as well as the demosaicking techniques used to retrieve multispectral information with good image quality [30,31].

TFDs offer the possibility of either being full polarized (all pixels with the same state of polarization), or bi-polarized (the sensor is biased in an interleaved configuration [19] with different bias voltage). Hence, some pixels can have an RGB set of sensitivities, and others can have different ones being biased by a different voltage. We also checked the effect of using this property. We compared a system taking advantage of this property to reduce the number of shots against a system taking more shots with different full-polarized biasing conditions. This second system had a larger number of channels, but also needed a longer capturing time.

As pointed out before, TFDs offer the capability of retrieving up to five channels out of a pixel in a single shot [32], but two of them would be in the near-infrared range (NIR) of the spectrum. Getting NIR information is an advantage in many applications. However, as a starting point for our study, we have decided to use the three channels within the visible range. This matches a general-purpose application. We have left the remaining two channels for future studies that target a more specific field of research, for which the NIR content might be critical. Including data beyond the limits of the visible range of the spectrum would make the simulations closer to the real behavior of the imaging systems in those channels close to the red and blue limits of the visible range. However, no spectral data was available for the set of spectral reflectances used (explained in Subsection 4.A) out of the visible range, so we restricted our camera response simulation to the limits of the visible range. Moreover, common imaging systems use a hot mirror to get rid of crosstalk effects between visible and NIR ranges of the spectrum [33], and most common optics components block naturally the UV components of radiation, to a certain extent.

3. Systems Simulated for Study and Comparison

TFDs are still under development and only prototype sensors exist so far. Nonetheless, their physical properties are very well characterized. The simulations of sensor response calculations are very realistic with the physical model provided by the developers (explained in Subsection 4.B). All systems studied in this work are silicon-based systems. Despite the architecture of each system in particular, the light acquisition process of them all can be realistically described with the CMOS capture model used for simulating camera responses. Therefore, sensor responses for every system studied in this work have been calculated using the same model for signal and noise generation. We can thus be sure that a higher or lower noise level is not the reason for better or worse performance. The average signal-to-noise ratio (SNR) of sensor responses is shown for each system in the last column of Table 1. As it will be explained in Subsection 4.B, the noise model accounts for exposure time and quantum efficiency. Therefore, even though the values of exposure time were different for different systems, this difference was taken into account when computing the noise. The calculation of SNR is also explained in Subsection 4.B. All systems result in similar SNR values. To demonstrate the performance of the designed system, we have created four experiments (explained in Subsection 4.D), in which its features are tested across other imaging systems. In Fig. 1, we can see a representation of the 11 different systems simulated in this study and used to complete the four different experiments.

In each system, there is a filter layer and a sensing layer. The number of different transmittances in the filter layer determines the spatial resolution. A CFA with 3 filters provides 1/3 of spatial resolution and, with 6 filters, 1/6. If the TFDs are bi-polarized (as explained in Section 2) and use no filters, then the spatial resolution reduces to 1/2. The characteristics of each system are explained and summarized in

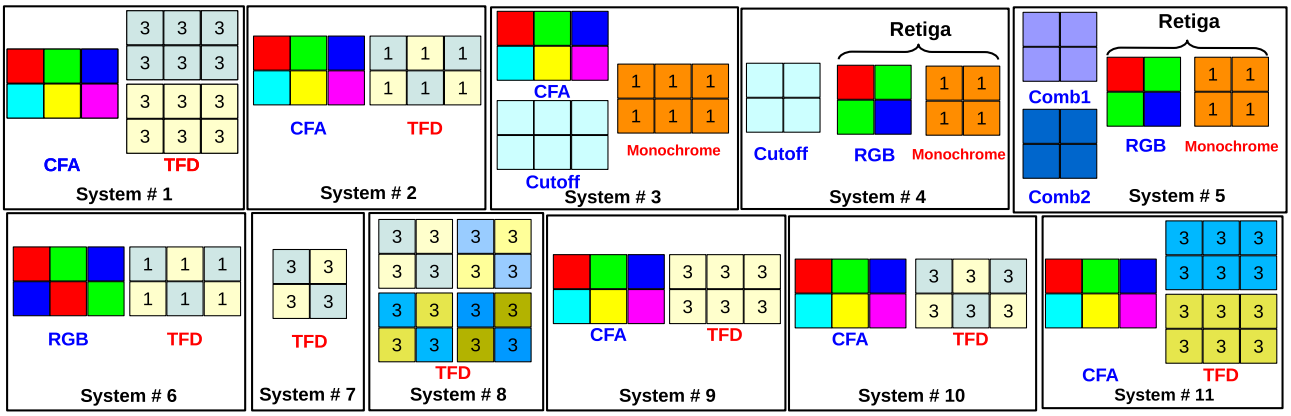

Fig. 1. Schemes for the 11 system configurations studied. Filter layer with blue caption, sensing layer with red caption. The number in each pixel determines the number of channels retrieved in one shot out of it.

Table 1. Figure 2 shows the spectral sensitivities of all systems, which are described in detail in the following paragraphs.

1. This is the proposed system. The filter layer is a CFA made up of six bandpass optical filters selected from a real database (see Subsection 4.A). The sensing layer is a TFD sensor full-polarized where two different biasing conditions were used in two shots. Three channels are retrieved per pixel; thus, without the need of any moving mechanical optical component, we get information from 36 channels in two shots.

2. Same CFA used as in system 1. The sensing layer, however, was a bi-polarized TFD from which we only take the information from one channel per pixel. The channel that was operative for each pixel was selected as explained in Subsection 4.A. This system is similar to the proposed one, but it has been designed to compare fairly its performance with other systems under the same conditions, in terms of number of channels.

3. The sensing layer is a normal monochrome silicon sensor like the one from a commercial scientific camera model Retiga SRV (QImaging Corp., Canada), equipped with a NIR cut-off filter. In the filter layer we set a CFA that has been optimized for the monochrome sensor, using the same technique as in our proposed system (see Subsection 4.A). This system has been designed to test if including a TFD sensor with its tunability property in a CFA-based system helps to improve its performance.

4. In the sensing layer, we used the spectral responsivities of the RGB scientific camera Retiga 1300C (QImaging Corp., Canada). In the first shot, we set no filter in the filter layer, and in the second shot we added an IR–UV cut-off filter (C_{off}) in front of the lens, which sharpened down the sensitivities in the extremes of the spectrum. This system has been designed to test the performance of TFDs plus CFAs against other types of systems and was proven to work well in a previous work [22].

5. In the sensing layer, we used the RGB camera again. This is also a 2-shot system. In each shot, a different ideally custom-made comb-shaped optical filter (Comb_1 and Comb_2) was placed in the filter layer. These filters divided by half the 3 sensitivities into 6. This idea of splitting the spectral sensitivities of single channels into two by using a comb-shaped filter has been used by other authors in the literature, as well [13,34,35].

6. The sensing layer used here was the same as in system 2, but the filter layer used was composed of an RGB Bayer filter like the one from the scientific camera Retiga 1300C (QImaging Corp., Canada)

Table 1. System Characteristics

System #	System Name	Filter Layer	Sensing Layer	# Shots	# Channels	Spatial Resolution	Average SNR (dB)
1	CFA+2TFD	CFA	TFD full	2	36	1/6	40.5
2	CFA+1TFDbi1	CFA	TFD bi	1	6	1/6	42.1
3	CFA+Monochrome	CFA	Silicon	1	6	1/6	42.5
4	Cut+Re	RGB RGB + C_{off}	Silicon	2	6	1/3	42.3
5	Comb+Re	RGB + Comb_1 RGB + Comb_2	Silicon	2	6	1/3	44.6
6	RGB+TFDbi	RGB	TFD bi	1	6	1/6	44.4
7	TFDbi	Empty	TFD bi	1	6	1/2	42.8
8	TFDbi4	Empty	TFD bi	4	24	1/2	47.7
9	CFA+1TFDfull	CFA	TFD full	1	18	1/6	47.4
10	CFA+1TFDbi3	CFA	TFD bi	1	18	1/6	40.9
11	CFA+2TFDalt	CFA	TFD full	2	36	1/6	45.1

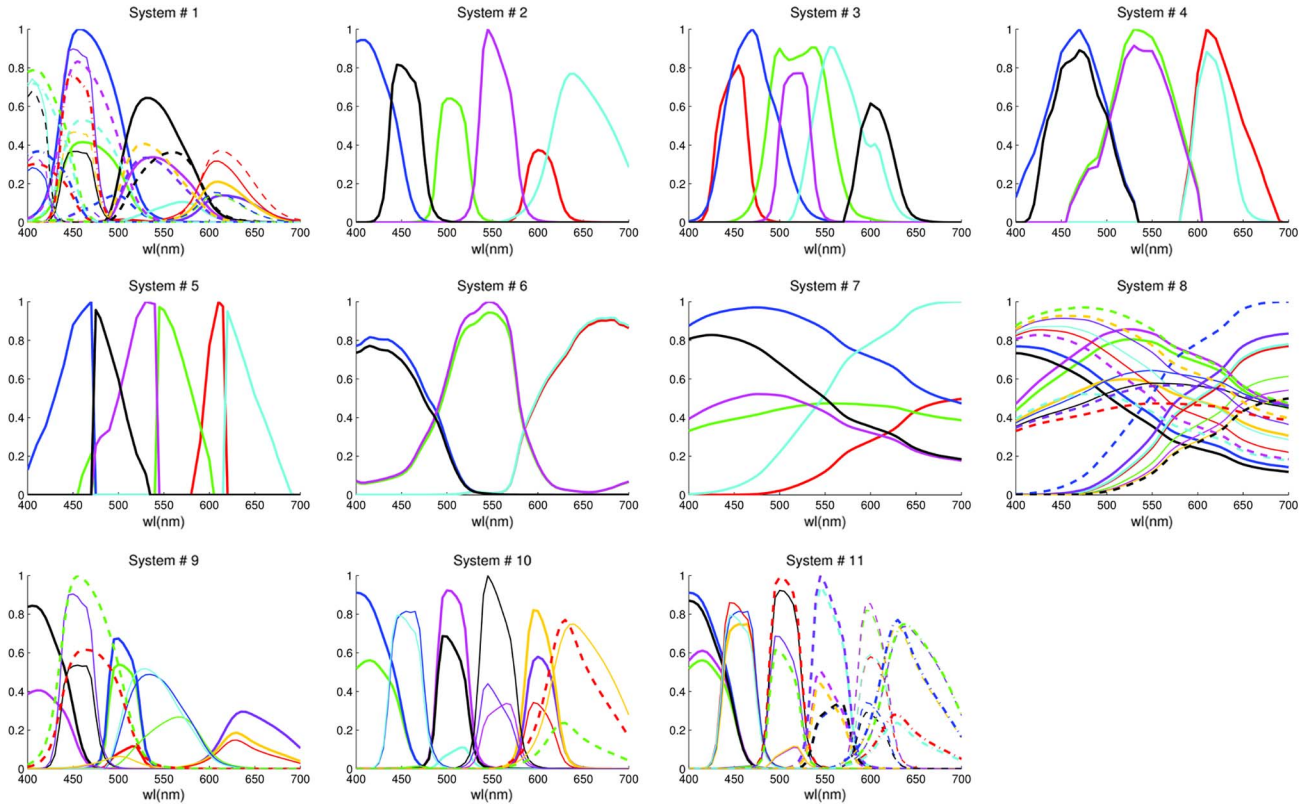

Fig. 2. Normalized spectral sensitivities of all systems versus wavelength in nanometers. Normalization was done just for displaying.

used in this study as representative of a standard RGB CCD camera. This system has been designed to test the effect of the selected CFA versus other kinds of narrow bandpass filters present in common color imaging systems.

7. The filter layer was left empty in this system. The sensing layer used was a bi-polarized TFD sensor. Only one shot was taken. This system and the next one were designed to test if adding a CFA to a TFD sensor improves its performance.

8. The same as the previous configuration but we took 4 shots with 4 different bi-polarization conditions. We wanted to push the number of channels to the extreme using all the biasing conditions that the developers provided.

9. The filter layer was the same CFA used in system number 1, but the sensing layer was a single shot full-polarized TFD sensor. This system has been designed to test the effect of bi-polarization property of the TFD sensor versus full-polarization state.

10. The filter layer was the same CFA used in system number 1. The sensing layer in this case was a one-shot bi-polarized TFD. This system is also proposed as an alternative to system number 1 if the application requires shorter capturing times (since it only uses one shot instead of two). In addition, this system has been compared with system number 9 to check the effect of using the bi-polarization property of TFD sensors.

11. This system has been designed exactly the same as the proposed system. However, its channels were randomly selected from the available ones instead of using voting principal feature analysis (VPFA) [36,37], as explained in Subsection 4.A. Ten reasonable combinations were simulated, in which the channels were selected to cover the whole visible range with some overlap. We selected the one with best results. This system was simulated to check whether selecting the channels using VPFA effectively increases the performance of our system or not.

We have specifically selected these configurations to better isolate the effect of the feature that we wanted to test in our proposed system. To sum up: system numbers 2, 3, 4, and 5 were designed to prove, under the same conditions in terms of number of channels, that using a TFD sensor improves the accuracy of the results over common monochrome sensor systems. System numbers 2, 6, 7, and 8 were designed and compared to test whether combining the CFA with the TFD also helps to improve its performance. System numbers 9 and 10 were designed to test if using the bi-polarization property of TFDs is better than not using it. System number 1 evolved from system number 10, pushing up the number of channels retrieved with only two biasing conditions adding one shot to the capturing process. System number 11 was designed to test the advantages of using VPFA to select the spectral channels.

4. Data, Methods, and Experiments

In this section, we explain the methods used to calculate the simulated sensor responses, select the set of sensors and filters, estimate the reflectances from

sensor responses, and evaluate the performance of each system.

A. Spectral Data, Sensors and Filters Selection

The illuminant used for the simulations was the CIE standard D65 illuminant, and the spectral reflectances were 1700 samples from the Natural Color System (NCS) [38,39], since they represent natural colors found in common scenes and are well-used for general-purpose color correction.

The TFD developers [40] provided a set of 8 different RGB sensitivities corresponding with 8 different biasing voltages (24 sensitivities in total). We can see them in Fig. 2 (system number 8). We used a set of 13 real filter transmittances from the Andover Corp. website [41]. These transmittances were selected so that they all covered the whole visible range of the spectrum, with certain overlap between them. We carefully chose different bandwidths and maxima positions. The commercial references of these 13 selected filters are: *400FS40, 400FS70, 450FS40, 450FS80, 500FS40, 500FS80, 550FS20, 550FS40, 550FS80, 600FS40, 600FS80, 650FS40*, and *650FS80*.

Combining all sensitivities with all filters, we obtained a total of 312 channels. Out of them we wanted to select six, corresponding to only two different biasing conditions (which are the limits of bi-polarization, as mentioned in Section 2) that could best recover the spectral information of the samples imaged in the visible range.

Basically, the VPFA method employed for this selection consisted of taking the sensor responses from all 312 channels of the 1700 NCS [39] color samples used, and perform principal component analysis on them. Then, we selected a reduced number of projections of the data onto the principal component vectors (half of them) and clustered them into 6 clusters (which was the number of channels we aimed to find) using the k-means clustering method [42]. We calculated the mean point (center) of each cluster, and then selected the filters corresponding to the vectors closest to each of the cluster centers. Due to the random initialization of k-means clustering, the clustering step is repeated over several iterations (50) and the selected vectors are voted. Finally, the most voted ones that correspond to only two biasing conditions are selected. Since each pixel of the TFD still gives us 3 channels, even though we only accounted for the 6 best performing ones (system number 2), we also studied the possibility of using the other two channels per pixel (system number 10), since results showed that there is a slight improvement when adding all of them. Furthermore, a two-shot system was also designed using the two full polarizations of the TFD resulting from VPFA and retrieving the information from all resulting channels (system number 1). After this process, the spectral sensitivities of the channels selected were the ones shown in Fig. 2 (upper left). As this figure shows, they span the whole visible range with some overlap. Overlap is important, so as not to leave a gap in the spectrum without retrieved information. However, too much overlap could cause bad performance as well, due to the consequent stronger effect of cross talk between channels. We also simulated other systems with the same characteristics, but selecting randomly the combination of filters and sensors to get a set of sensitivities that visually looks good (covering the spectrum, relatively narrow, and with some spectral overlap). The one with the best results was chosen (system number 11) and compared with the one resulting from the VPFA. The metrics used for evaluation of the performance were spectral (goodness-of-fit coefficient GFC [22,43], also known as complementary Pearson distance [44], and root mean square error, RMSE [22,45]), and colorimetric (ΔE^*_{00} also called CIEDE2000 [46]). If GFC > 0.999, then the estimation was considered quite good, and if GFC > 0.9999, then the estimation was almost an exact fit [47]. Regarding CIEDE2000 colorimetrics, differences of less than $1\Delta E^*_{00}$ units were considered acceptable.

B. Sensor Responses Simulation

As mentioned before, TFDs are still a technology under development, from which only prototype sensors exist and there is no implementation of a complete imaging system available yet. We were provided by the developers [40] with a physical model that well describes the opto-electronic behavior of such sensors. They are based on CMOS technology. The first step was to calculate the power spectral density ($\mathrm{PSD}_i(\lambda)$) of the color signal incident on a pixel, composed by the illuminant ($D65$) spectral photon flux per unit time and area ($\mathrm{SPF}_{D65}(\lambda)$), and the sample spectral reflectance of the ith color sample ($R_i(\lambda)$), as Eq. (1) shows:

$$\mathrm{PSD}_i(\lambda) = R_i(\lambda) \cdot \mathrm{SPF}_{D65}(\lambda). \tag{1}$$

Afterwards, we calculated the photocurrent I_{ik} and the output voltage V_{oki} generated in each channel f for each sample i, as Eqs. (2) and (3) show:

$$I_{ik} = q \cdot A_{\mathrm{light}} \cdot \int_{\lambda} \mathrm{PSD}_i(\lambda) \cdot Q_{ek}(\lambda)\mathrm{d}\lambda, \tag{2}$$

$$V_{oki} = I_{ik} \cdot \frac{T_{\mathrm{int}}}{C_f}, \tag{3}$$

where q is the elementary charge, A_{light} is the illuminated area on the TFD surface, $Q_{ek}(\lambda)$ is the spectral quantum efficiency of channel k, T_{int} is the exposure time, and C_f is the pixel feedback capacitance. After calculating the noiseless output voltage (V_{oki}) and before quantifying the signal into digital counts, the additive noise component (η_σ) is added to get the noisy output voltage ($V_{\eta oki}$). This noise was Gaussian and its total variance σ_{tot} was composed by kTC noise (σ_T), and dark current and shot

noise (σ_{DCS}). These calculations are shown in Eqs. (4)–(7):

$$\sigma_T = \sqrt{\frac{K \cdot T}{C_f}}, \tag{4}$$

$$\sigma_{\mathrm{DCS}} = \sqrt{\frac{q \cdot T_{\mathrm{int}}}{C_f^2} \cdot (I_{ik} + J_d \cdot A_{\mathrm{tot}})}, \tag{5}$$

$$\sigma_{\mathrm{tot}} = \sqrt{\sigma_T^2 + \sigma_{\mathrm{DCS}}^2}, \tag{6}$$

$$V_{\eta oki} = V_{oki} + \eta_\sigma, \tag{7}$$

where K is the Boltzmann constant, T is the absolute temperature (300°K = 26.85°C), J_d is the dark current density, and A_{tot} is the overall sensor area. As a last step, we quantify the noisy voltage signal to get the noisy sensor response in digital counts (ρ_{ki}) using B bits, as shown in Eq. (8):

$$\rho_{ki} = \mathrm{round}\left[\frac{V_{\eta oki \cdot (2^B - 1)}}{V_{dd}}\right], \tag{8}$$

where V_{dd} is the pixel maximum output range. The values of all TFD-specific parameters were extracted from Finite Elements Simulation using the Dessis software, as described in [19]. Typical biasing voltages of the TFD anodes range between 0.5 V and 8 V. The simulations of all systems compared in this study were done using the same model for sensor responses calculations. As it would happen in reality, including filters on top of the sensor reduces their quantum efficiency ($Q_{ek}(\lambda)$). Moreover, in some systems, the difference in magnitude between $Q_{ek}(\lambda)$ of different channels would cause low dynamic range problems. The exposure time selected for each system was the one exposing most of the samples within the sensor range between noise floor and saturation level. This means that in all cases, some samples were underexposed and some overexposed, but in all cases this is taken into account for computing the results, which reflect all these realistic limitations of imaging systems. Due to the fact that the noise model used accounts for different $Q_{ek}(\lambda)$ levels and exposure time values, the resulting levels of noise in all systems were similar, as mentioned in Section 3. The calculation of SNR was done by calculating the ratio in logarithmic units between the noiseless output voltage signal (V_{oki}) calculated in Eq. (3) over the total noise variance (σ_{tot}) calculated in Eq. (6).

C. Spectral Reflectance Estimation and Evaluation

The estimation method used was regularized inhomogeneous polynomial kernel regression [48,49], which was the one found to perform best in previous studies [22]. The model used for the estimation needs to set two free parameters, the degree of the polynomial (d) and the regularization term (λ_r), and these values need to be optimized. Then, a double 10-fold cross validation method was used, one nested inside the other. In the outer loop, we randomly separated the 1700 NCS samples into optimization and evaluation sets. The former was used to find the optimal values of the two parameters of the mathematical model, and the latter to evaluate the estimation of reflectances. In each iteration of the outer 10-fold, the optimization set was used in the inner 10-fold loop. A grid of parameters was built and every possible combination of both parameters was tested with the testing set corresponding to each loop of the inner 10-fold. For the parameter optimization, we used the CIEDE2000 color difference formula as a cost function (ΔE_{00}^*). Using a color difference formula for parameter optimization gives to colorimetric performance of the results a slight advantage over spectral performance, compared with using a spectral metric. The opposite happens using a spectral metric as a cost function. In this work, we aimed for a general color correction application via spectral estimation. Therefore, the metric we chose for our application domain was CIEDE2000. After the 10 iterations of the inner loop, the average of the parameters found was calculated, and those were used for evaluation in the outer loop. The standard deviation was also calculated to check if the distributions of the best parameters found were stable. Table 2 shows the means and standard deviations (in parenthesis) of both kernel parameters found in every system out of each of the 100 folds in total (10 times 10 folds). Since the regularization parameter was optimized, looking for the value that performed best in different orders of magnitude, the mean and standard deviation calculated for it are shown in logarithmic scale. After the 10 iterations of the outer loop, the average error metrics were calculated, and these were the values with which we compared the different systems. By doing things this way, we ensured that samples used for training would never be used for testing, which

Table 2. Means and Standard Deviations of Polynomial Degree (d) and Regularization Term (λ_r) Kernel Parameters Values [a]

System #	1	2	3	4	5	6	7	8	9	10	11
d	2.68	5.70	5.9	7.08	3.47	7.00	5.41	4.13	4.46	4.30	4.60
	(0.58)	(0.48)	(1.37)	(0.80)	(1.14)	(0.00)	(1.31)	(0.83)	(0.59)	(0.48)	(0.52)
$\log_{10}(\lambda_r)$	−2.85	−6.06	−4.29	−5.81	−4.15	−4.93	−8.19	−4.39	−3.17	−4.23	−4.43
	(0.43)	(0.13)	(1.06)	(0.87)	(1.94)	(0.22)	(1.19)	(0.64)	(0.68)	(0.42)	(0.50)

[a]The regularization term values are shown in logarithmic scale. Values of d are not integers because they are the computed mean.

could lead to over-fitting and, thus, to overestimating the quality of the recovered samples obtained from the camera responses of our system.

D. Experiments

In this study, four experiments were conducted to test the following four hypotheses:

1. The proposed system performs better than systems based on other types of sensors, such as monochrome sensors or multi-shot systems based on RGB scientific cameras, plus some filters.
2. The proposed system performs better than systems based on TFDs that do not use the designed CFA. Therefore, this CFA helps to improve the performance of TFDs.
3. Using the bi-polarization property of a TFD sensor, the performance can be improved without the need to increase the number of shots taken.
4. The proposed system performs better than a similar system in which the channels selection was not done by VPFA.

Hypothesis 1 demonstrates the capabilities of the proposed design. Hypotheses 2 and 3 try to go deeper into the reasons justifying the superiority of our TFD-based proposed system. Hypothesis 4 demonstrates that using VPFA for selecting candidate filters and channels is worthy. The results for these four experiments are shown in the following Section 5.

5. Results and Discussion

In this section, results from the four experiments are shown in tabular format in Table 3, and then commented on.

A. Experiment 1: Superiority of TFDs

This experiment was performed to check whether our proposed system yields better accuracy in spectral reflectance estimation than systems based on other types of sensors using the same kind of CFAs, and other different ones. We included the 2-shot system made up of a RGB Retiga camera and a cut-off filter, studied in previous work [50], and we also designed a new system, in which we divided each of the R, G, and B sensitivities of the Retiga camera into two halves (see Fig. 2, system 5) with theoretical custom-designed comb-shaped optical filters. The results are shown in Table 3 for system numbers 2, 3, 4, and 5.

We see that system number 2 performed better both for spectral and color metrics than system numbers 3 and 4. The system formed by a monochrome silicon sensor plus the CFA performed worst. This means that the good results obtained by system 2 were not only due to the CFA, but also that the tunability and bi-polarization property of TFD sensor helped to improve the performance. The systems formed by RGB Retiga camera plus filters yielded good results, especially the one with the two comb-shaped filters which worked best. This is in agreement with previous studies [22,50]. This comparison was made in the same conditions as far as the number of channels is concerned. However, when we simulated system number 2, we were neglecting the information from 2 channels from each pixel. This meant that we were not using 12 channels, which we achieved anyway in the capturing process. If we add this available information to the recovery (system number 10), then the results improved, and were closer to system number 5 colorimetrically and better spectrally, and we would not need to add a second shot. Besides, there is no need to mechanically move any filter in and out. Moreover, we went further and designed a two-shot fully-polarized system, and we retrieved the information from all channels under each pixel, increasing the number of channels to up to 36 in just two shots. This was the strategy of system number 1. In this case, we used two shots, which were sequentially captured just by switching the biasing voltage of the TFD. These results outperformed the ones from system 5. Practice systems 1 and 5 have the same number of shots. For system number 5, the capturing process cannot be real-time since the filters need to be switching all the time. In the case of the proposed system, even though it is also a 2-shot system, the switching of the biasing voltage occurs so fast (electronically controlled) that it can be considered as a snapshot system in practice, as long as the lighting conditions in the scene being imaged allow short exposure times. We thus propose system number 10 as a good and convenient option to be used in applications requiring fast capturing times. If we could afford longer capturing times, then system number 1 is our proposal to enhance performance by increasing the number of shots.

B. Experiment 2: Superiority of a CFA-Based System

In this experiment, we wanted to assess whether including the CFA in the TFD-based system is the key factor leading to the superior performance of

Table 3. Results for All Systems Studied (Mean and STD)

System #	1	2	3	4	5	6	7	8	9	10	11
ΔE^*_{00}	**0.23** (0.16)	0.51 (0.35)	1.26 (1.3)	0.66 (0.47)	0.27 (0.18)	0.96 (0.73)	3.15 (2.27)	1.89 (1.32)	0.41 (0.34)	0.38 (0.26)	0.35 (0.24)
GFC	**0.9997** (0.0028)	0.9987 (0.0043)	0.9992 (0.0022)	0.9982 (0.0030)	0.9992 (0.0030)	0.9986 (0.0079)	0.9976 (0.0038)	0.9982 (0.0033)	0.9991 (0.0010)	0.9996 (0.0012)	0.9994 (0.0021)
RMSE	**0.0064** (0.0063)	0.0101 (0.0087)	0.0094 (0.0082)	0.0103 (0.062)	0.0083 (0.0099)	0.0113 (0.034)	0.0169 (0.0092)	0.0143 (0.008)	0.0082 (0.004)	0.0064 (0.0043)	0.0067 (0.0053)

the system shown in the first experiment. For this, we simulated different systems using TFD sensors. One of them used no filters in front (system number 7). The other used the filters present in a common RGB scientific color camera (system number 6). This was done as we had found out that the RGB Retiga camera was working very well and that these filters drastically narrowed down raw sensitivities. The results of this experiment are shown in Table 3 for system numbers 2, 6, and 7. We see that the system we propose performs best for both spectral and color metrics. Including other kinds of filters, such as RGB color filters (system number 6), yielded reasonable results, but still far from the ones reached by system number 2. The system using only two TFD polarizations, such as the proposed one, but with no CFA (system number 7), yielded quite bad results, in comparison. Even pushing this strategy to the limit of taking many shots with different polarizations (up to 4 shots with bi-polarized TFD, as in system number 8), the performance was far from our proposed system. Therefore, it did not outperform a system with only one shot, two TFD polarizations, and a CFA. If the difference is clear when the experiment is performed for systems with the same number of channels, then the difference is still more evident if we use the system number 10 or 1, the advantages of which were pointed out in Subsection 5.A. Therefore, we can conclude that adding in a TFD-based system and a CFA that has been calculated to be the best choice out of a set of available color filters, helps improve the performance for spectral reflectance estimation beyond systems without CFA, or with other filters that are not specially selected.

C. Experiment 3: Bi-Polarization versus Full-Polarization

In this experiment, we wanted to assess whether using the capability of bi-polarizing the TFD sensor helps with its performance. The easiest way to test this was by comparing system numbers 9 and 10. The results are shown in Table 3. We see how using two polarization states of the TFD improved the system performance slightly for both color and spectral metrics. Therefore, it is worth adding a second biasing voltage since, in practice, the acquisition process is the same. Apparently, the use of sensitivities from different biasing voltages perform better than the use of all sensitivities from the same biasing condition.

D. Experiment 4: VPFA Performance

This last experiment was carried out to check whether it makes sense to invest time in selecting the filters using VPFA, or to simply choose any combination of them randomly to cover the visible range of the spectrum and having some little overlap between bands, to yield better results. Out of ten random combinations, we selected the one that produced the best results (system 11) and compared it with our proposed system (number 1). Results show that the VPFA procedure improves the filter selection process. As a final summary of results, we present the quality indexes for all systems compared in this study in Fig. 3.

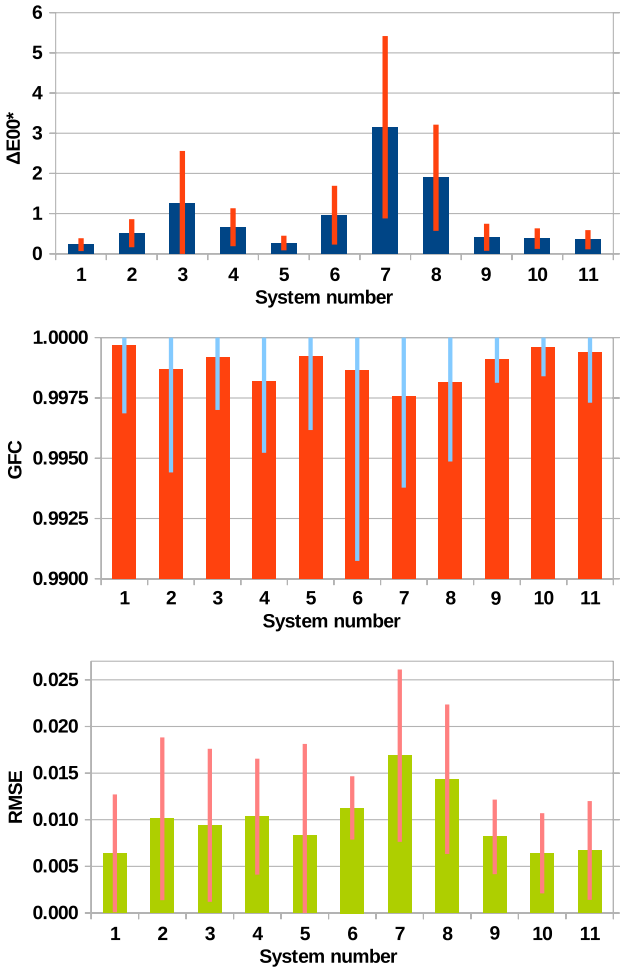

Fig. 3. Color and spectral error metrics for all systems. The bars show the mean value and the lines the standard deviation centered on the mean.

6. Conclusions

Two different approaches for a multispectral imaging system have been proposed, based on a novel technology still under development. TFD-based systems, which exploit the tunability and bi-polarization properties of this new type of sensor, are coupled with a CFA whose filter transmittances have been selected via VPFA method out of a set of real available color filters from a commercial database. Four simulation experiments have been conducted to demonstrate whether the proposed systems outperform others with similar characteristics, but using other design strategies. We aimed to assess whether the combination TFD plus CFA works better than any of them separately. We also studied whether we could use different architectures, depending on the requirements of particular applications. We found that the proposed systems perform both colorimetrically and spectrally as well as or better than the other systems while offering an easier

and more elegant solution to the problem of spectral imaging. Some systems give close spectral results, and only one of them comes significantly close in the color metrics. However, this system is made up of a scientific RGB color camera plus two ideal filters that are placed in front of it, alternatively, such that each capture would need two shots and the mechanical switch between filters from one shot to the next. This would make the real-time capture unsuitable. One of our proposed systems (system number 1) is also a two-shot system, but the tuning of the sensitivities is so fast and easy that we could still use it for real-time spectral imaging if the conditions of the amount of light in the scene being imaged allowed it. Otherwise, we offer a single-shot alternative system (system number 10), which would reduce drastically the time needed for capturing and would still give good colorimetric and spectral results. This study demonstrates the potential of TFD sensors as candidates to be part of spectral imaging systems that are portable, real-time, versatile, and low cost, as soon as they can be implemented as part of a real capturing system.

This work was funded by the Spanish Ministry of Economy and Competitiveness through the research project DPI2011-23202. We thank our colleague A. L. Tate for revising our English text.

References

1. C. I. Chang, *Hyperspectral Data Exploitation: Theory and Applications* (Wiley, 2007).
2. H. Grahn and P. Geladi, *Techniques and Applications of Hyperspectral Image Analysis* (Wiley, 2007).
3. Y. Garini, I. T. Young, and G. McNamara, "Spectral imaging: principles and applications," Cytometry **69**, 735–747 (2006).
4. H. R. Morris, C. C. Hoyt, and P. J. Treado, "Imaging spectrometers for fluorescence and Raman microscopy: acousto-optic and liquid-crystal tunable filters," Appl. Spectrosc. **48**, 857–866 (1994).
5. S. E. Harris and R. W. Wallace, "Acousto-optic tunable filter," J. Opt. Soc. Am. **59**, 744–747 (1969).
6. I. Chang, "Acousto-optic tunable filters," Opt. Eng. **20**, 206824 (1981).
7. J. Y. Hardeberg, F. Schmitt, and H. Brettel, "Multispectral color image capture using a liquid-crystal tunable filter," Opt. Eng. **41**, 2532–2548 (2002).
8. A. Rodríguez, J. L. Nieves, E. Valero, E. Garrote, J. Hernández-Andrés, and J. Romero, "Modified fuzzy c-means applied to a Bragg-grating-based spectral imager for material clustering," in *IS&T/SPIE Electronic Imaging* (International Society for Optics and Photonics, 2012), p. 83000J.
9. N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," Opt. Eng. **52**, 090901 (2013).
10. L. Gao, R. T. Kester, N. Hagen, and T. S. Tkaczyk, "Snapshot image mapping spectrometer (IMS) with high sampling density for hyperspectral microscopy," Opt. Express **18**, 14330–14344 (2010).
11. J. Y. Hardeberg, *Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches* (Universal-Publishers, 2001).
12. J. Brauers, N. Schulte, and T. Aach, "Multispectral filter-wheel cameras: geometric distortion model and compensation algorithms," IEEE Trans. Image Process. **17**, 2368–2380 (2008).
13. R. Shrestha, J. Y. Hardeberg, and A. Mansouri, "One-shot multispectral color imaging with a stereo camera," Proc. SPIE **7876**, 787609 (2011).
14. M. Tsuchida, T. Takahashi, K. Ito, T. Kawanishi, J. Yamato, and T. Aoki, "A stereo one-shot multi-band camera system for accurate color reproduction," in *ACM SIGGRAPH 2010 Posters* (ACM, 2010), p. 66.
15. M. Tsuchida, T. Kawanishi, K. Kashino, and J. Yamato, "A stereo nine-band camera for accurate color and spectrum reproduction," in *ACM SIGGRAPH 2012 Posters* (ACM, 2012), p. 18.
16. B. Gunturk, J. Glotzbach, Y. Altunbasak, R. Schafer, and R. Mersereau, "Demosaicking: color filter array interpolation," IEEE Signal Process. Mag. **22**, 44–54 (2005).
17. Y. Murakami, M. Yamaguchi, and N. Ohyama, "Hybrid-resolution multispectral imaging using color filter array," Opt. Express **20**, 7173–7183 (2012).
18. Y. Murakami, A. Tanji, and M. Yamaguchi, "Development of a low-resolution spectral imager and its application to hybrid-resolution spectral imaging," in *Proceedings 12th International AIC Congress* (2013), pp. 363–366.
19. G. Langfelder, "Spectrally reconfigurable pixels for dual-color-mode imaging sensors," Appl. Opt. **51**, A91–A98 (2012).
20. A. Longoni, F. Zaraga, G. Langfelder, and L. Bombelli, "The transverse field detector (TFD): a novel color-sensitive CMOS device," IEEE Electron. Device Lett. **29**, 1306–1308 (2008).
21. G. Langfelder, A. Longoni, and F. Zaraga, "Implementation of a multi-spectral color imaging device without color filter array," Proc. SPIE **7876**, 787607 (2011).
22. M. Martínez-Domingo, E. Valero, J. Hernández-Andrés, and G. Langfelder, "Spectral reflectance estimation from transverse field detectors responses," in *Conference on Color in Graphics, Imaging, and Vision* (Society for Imaging Science and Technology), 2012, pp. 378–383.
23. G. D. Finlayson, J. Vazquez-Corral, S. Süsstrunk, and M. Vanrell, "Spectral sharpening by spherical sampling," J. Opt. Soc. Am. A **29**, 1199–1210 (2012).
24. E. R. Fossum, "CMOS image sensors: electronic camera-on-a-chip," IEEE Trans. Electron Dev. **44**, 1689–1698 (1997).
25. P. M. Hubel, "Foveon technology and the changing landscape of digital cameras," in *Color and Imaging Conference* (Society for Imaging Science and Technology, 2005), pp. 314–317.
26. M. S. Drew and G. D. Finlayson, "Spectral sharpening with positivity," J. Opt. Soc. Am. A **17**, 1361–1370 (2000).
27. A. Lin and F. Imai, "Efficient spectral imaging based on imaging systems with scene adaptation using tunable color pixels," in *Color and Imaging Conference* (Society for Imaging Science and Technology, 2011), pp. 332–338.
28. F. H. Imai, "Computational spectral imaging based on adaptive spectral imaging," in *Computational Color Imaging* (Springer, 2013), pp. 35–52.
29. K. Hirakawa and P. J. Wolfe, "Spatio-spectral color filter array design for optimal image recovery," IEEE Trans. Image Process. **17**, 1876–1890 (2008).
30. L. Miao, H. Qi, R. Ramanath, and W. Snyder, "Binary tree-based generic demosaicking algorithms for multispectral filter arrays," IEEE T. Image Process. **15**, 3550–3558 (2006).
31. J. Brauers and T. Aach, "A color filter array-based multispectral camera," in *Workshop Farbbildverarbeitung* (Ilmenau, 2006).
32. G. Langfelder, T. Malzbender, A. Longoni, and F. Zaraga, "A device and an algorithm for the separation of visible and near infrared signals in a monolithic silicon sensor," Proc. SPIE **7882**, 788207 (2011).
33. G. Langfelder, "CMOS pixels directly sensitive to both visible and near-infrared radiation," IEEE Trans. Electron Dev. **60**, 1695–1700 (2013).
34. M. Hashimoto and J. Kishimoto, "Two-shot type 6-band still image capturing system using commercial digital camera and custom color filter," in *Conference on Color in Graphics, Imaging, and Vision* (Society for Imaging Science and Technology, 2008), pp. 538–541.
35. M. Tsuchida, A. Takayanagi, Y. Sakaguchi, and R. Mukai, "Estimation of spectral reflectance from six-band images based on partial least-squares regression," in *Proceedings 12th International AIC Congress*, 2013, pp. 1785–1788.
36. I. Chatzis, V. Kappatos, and E. Dermatas, "Filter selection for multi-spectral image acquisition using the feature vector

analysis methods," in *Intelligent Production Machines and Systems: 2nd I* PROMS Virtual Conference* (Elsevier Science, 2007).
37. Y. Lu, I. Cohen, X. Zhou, and Q. Tian, "Feature selection using principal feature analysis," in *Proceedings of the 15th International Conference on Multimedia* (ACM, 2007), pp. 301–304.
38. Natural Color System (NCS). Scandinavian Colour Institute of Stockholm, Sweden, (1979).
39. A. Hård and L. Sivik, "NCS- natural color system: a Swedish standard for color notation," Color Res. Appl. **6**, 129–138 (1981).
40. G. Langfelder, F. Zaraga, and A. Longoni, "Tunable spectral responses in a color-sensitive CMOS pixel for imaging applications," IEEE Trans. Electron Dev. **56**, 2563–2569 (2009).
41. http://www.andovercorp.com/Web_store/Standard_BP/Std_BP_General.php.
42. J. A. Hartigan and M. A. Wong, "Algorithm as 136: a k-means clustering algorithm," J. Roy. Stat. Soc. C-App **28**, 100–108 (1979).
43. J. Hernández-Andrés, J. Romero, and R. L. Lee, "Colorimetric and spectroradiometric characteristics of narrow-field-of-view clear skylight in Granada, Spain," J. Opt. Soc. Am. A **18**, 412–420 (2001).
44. V. Heikkinen, A. Mirhashemi, and J. Alho, "Link functions and Matérn kernel in the estimation of reflectance spectra from RGB responses," J. Opt. Soc. Am. A **30**, 2444–2454 (2013).
45. F. H. Imai, M. R. Rosen, and R. S. Berns, "Comparative study of metrics for spectral match quality," in *Conference on Colour in Graphics, Image and Vision* (Society for Imaging Science and Technology, 2002), pp. 492–496.
46. G. Sharma, W. Wu, and E. Dalal, "The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations," Color Res. Appl. **30**, 21–30 (2005).
47. J. L. Nieves, E. M. Valero, S. M. Nascimento, J. Hernández-Andrés, and J. Romero, "Multispectral synthesis of daylight using a commercial digital CCD camera," Appl. Opt. **44**, 5696–5703 (2005).
48. V. Heikkinen, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, T. Jaaskelainen, and S. Lee, "Regularized learning framework in the estimation of reflectance spectra from camera responses," J. Opt. Soc. Am. A **24**, 2673–2683 (2007).
49. V. Heikkinen, "Kernel methods for estimation and classification of data from spectral imaging," Ph.D. dissertation (Faculty of Forestry and Natural Sciences, University of Eastern Finland, 2011).
50. M. Martínez-Domingo, E. Valero, V. Heikkinen, Jr., and G. Langfelder, "Design of a multispectral system based on transverse field detectors," in *Proceedings 12th International AIC Congress*, 2013, pp. 371–374.