

Simultaneous Multispectral Imaging and Illuminant Estimation Using a Stereo Camera

Raju Shrestha and Jon Yngve Hardeberg

The Norwegian Color Research Laboratory, Gjøvik University College,
Gjøvik, Norway

raju.shrestha@hig.no
<http://www.colorlab.no>

Abstract. We propose here a novel approach to acquire a multispectral image and at the same time estimate the illuminant with the use of a stereo camera. Two images of a scene: one normal RGB and one filtered image with an appropriate optical filter selected from among readily available filters placed in front of a lens of the stereo camera are acquired. The spectral reflectance and/or color at each pixel on the scene are estimated from the corresponding outputs in the two images. In the mean time, the illuminant used during the image capture is estimated using chromagenic illuminant estimation method. Experiments with the simulated data show that this is a promising technique for simultaneous multispectral imaging and the illuminant estimation. Today's increasing commercial availability of digital stereo cameras makes the proposed solution a viable one for many applications.

1 Introduction

Multispectral imaging provides a solution to the limitations of conventional three channel (usually RGB) color imaging like metamerism and environment dependency. There are different types of multispectral imaging systems, most of them are filter-based which use additional filters to expand the number of color channels. The state-of-the art filter-based multispectral imaging systems [6,8,19] acquire images in multiple shots. Their use is, therefore, limited to static scenes only, thus making them less useful and less practical. Shrestha et al. [16,17] has made a comprehensive study and proposed a practical and feasible 6-channel multispectral imaging system with the use of a stereo camera. Depending upon the sensitivities of the two cameras in the stereo system, one or two appropriate optical filters are selected from among a set of readily available filters and placed in front of one or both lenses of the stereo camera, so that they will modify the sensitivities of one or two cameras to produce six channels (three each contributed from the two cameras) in the visible spectrum so as to give optimal estimation of the scene spectral reflectance and/or the color. This produces a 6-channel multispectral system.

Color constancy is yet another important issue in color imaging. It is the ability to account for the color of the light source which allows seeing the color of an object more or less the same under different lighting conditions. Human vision is said to be color constant as it has a natural tendency to correct for the effects of the color of the light source. Computational color constancy tries to emulate the color constancy in color imaging, and this is one of the fundamental requirements in many color imaging and computer vision applications. Computational color constancy models, in general, comprise of two steps: illuminant estimation and color correction. The illuminant estimation is the primary task in a computational color constancy algorithm. Knowing the estimated illuminant, the effects of the color of the illuminant are corrected to obtain the desired color constant image. Many methods have been proposed for the illuminant estimation, and these methods are typically based on the assumption of spatially uniform color of the light source across the scene. Some example methods are gray-world, max-RGB, a gamut based algorithm, neural networks, color-by-correlation, a Bayesian method. Yet another color constancy algorithm, known as the chromagenic color constancy, has been proposed by Finlayson et al. [4] that uses a special color filter which they named as chromagenic. This algorithm estimates the illuminant from two images: a normal RGB and a filtered RGB, of a scene. The algorithm has registration problems and also requires two shot images.

In this paper, we extend the multispectral imaging proposed by Shrestha et al. [16, 17] by making the system capable of acquiring multispectral image and at the same time estimating the illuminant under which the image has been acquired. For this, instead of two, only one filter is used in front of one of the lenses of the stereo camera. It thus acquires two images: one normal RGB image and one filtered image, in a single shot. The 6-channel multispectral image is estimated from these two images, and at the same time the illuminant is estimated using the same two images. The proposed system is thus capable not only of acquiring the multispectral image but also acquiring the normal RGB image, and at the same time capable of estimating the illuminant under which the images are taken. This gives users not only the flexibility to choose among a normal RGB and a multispectral image, but also to obtain a color constant image with the use of the estimated illuminant. The proposed system would therefore be useful in many applications where multispectral images and color constancy are applicable. We have performed experiments with the simulated data and they produce promising results.

After this introduction, we present the proposed system along with the method of multispectral imaging and the illuminant estimation in Section 2. We then present the experiments and results in Section 3. The results will be discussed next in Section 4, and finally we conclude the paper in Section 5.

2 Proposed System

The proposed system is constructed from a stereo camera and an appropriate optical filter in front of one of the lenses of the stereo camera (Fig. 1). Either a commercial stereo camera or two modern digital (RGB) cameras joined in a stereoscopic configuration can be used. An optimal filter is selected from the given set of filters through an exhaustive search. Since only one filter needs to be selected, the computational complexity is just $\mathcal{O}(n)$. We select a filter that produces minimum estimation errors (spectral or color) with regards to both the multispectral output and the color constancy output. We use here in this paper the minimum color estimation error as the criteria for the multispectral output for more accurate color reproduction, and use the most widely used median error [1, 7] for more accurate illuminant estimation. However, depending on the application, spectral estimation error criteria could also be used. Section 3.2 describes the filter selection in the experiment in details.

The selected filter modifies the sensitivities of the filtered side of the camera producing six channels (three each contributed from the two cameras) in the visible spectrum. The system allows capturing two images, one normal RGB and one filtered RGB images of a scene. Furthermore, knowing the geometry of the stereo camera, not only the two images can be registered rather more precisely but also 3D information of the scene can be obtained. However 3D acquisition is beyond the scope of this paper. Among many registration techniques proposed in the literature, the phase-only correlation method (POC) [18] could be the one for precision registration. From the two registered images of a scene, the multispectral image is obtained and at the same the illuminant is estimated. The subsections below discuss the multispectral system model and the illuminant estimation with the proposed system.



Fig. 1. Illustration of a propose multispectral camera

2.1 Multispectral System Model

In order to model the proposed multispectral system, let $S = [s_R, s_G, s_B]$ denotes the matrix of spectral sensitivities of the three channels of the normal RGB camera of the stereo camera, and similarly $S^F = [s_R^F, s_G^F, s_B^F]$ is the matrix of the spectral sensitivities of the three channels of the filtered side of the stereo camera. Let t is the spectral transmittance of the selected filter, L is the spectral power distribution of the light source, and R is the spectral reflectance of the surface captured by the camera. As there is always acquisition noise introduced into the camera outputs, let n and n^F denote the noise vectors corresponding to the acquisition noise in the three channels of the normal and the filtered side of the stereo camera respectively. The camera responses of the normal (C) and the filtered (C^F) sides of the stereo camera are respectively given by:

$$C = S' \text{Diag}(L)R + n \quad (1)$$

$$\text{and } C^F = (S^F)' \text{Diag}(L)R + n^F, \quad (2)$$

where X' denotes the transpose of the matrix X . The combined response $\begin{bmatrix} C \\ C^F \end{bmatrix}$ of the two cameras gives six responses. The estimated reflectance (\tilde{R}) is obtained for the corresponding original reflectance (R) from these camera responses for the training and the test targets C_{train} and C respectively, using different spectral estimation methods. Training targets are the database of surface reflectance functions from which basis functions are generated and test targets are used to validate the performance of the device. Among many estimation algorithms proposed in the literature, we have investigated the performance of the proposed system with four different estimation methods: Imai and Berns (IB) [9], Linear Regression, Polynomial (PN) [3] and Neural Network (NN) [12]. These methods are described in details by Shrestha et al. [17].

The estimated reflectances are evaluated using spectral as well as colorimetric metrics. Two different metrics: GFC (Goodness of Fit Coefficient) and RMS (Root Mean Square) error have been used as spectral metrics, and ΔE_{ab}^* (CIELAB Color Difference) as the colorimetric metric. For the details on these metrics also, we refer to Shrestha et al. [17].

2.2 Illuminant Estimation

The two images of a scene allow estimating the illuminant as well. We use the chromagenic illuminant estimation method proposed by Finlayson et al. [4]. The chromagenic algorithm is based on the first approximation that the image formed by placing a colored filter in front of the camera is the same as changing the illumination impinging on the scene. In other words, the responses of the camera with and without the filter can be considered as the responses of a single surface under two different illuminants. When the same surfaces are viewed under two light sources, the corresponding camera responses, to a good approximation, can be related by a linear transform [11]. Therefore, if C and C^F denote the unfiltered and the filtered camera responses respectively, then these responses can be related by the following equation:

$$C^F = MC, \quad (3)$$

where M is a 3×3 linear transformation matrix. M depends on both the illuminant and the filter used, and it can be computed as:

$$M = C^F C^+. \quad (4)$$

The transformation matrix M can be described as the transform that maps, in a least square sense, unfiltered to filtered responses of the camera under a given illuminant. A linear model of illuminant change is not perfect and may result in large estimation errors. More accurate mapping with a reduced estimation error can be obtained with a convex relational model by expressing an RGB image as a convex combination of the training RGBs for each training light in turn. The chromagenic illuminant estimation method is, therefore, based on the assumption that we know all possible illuminants a priori. The transforms M_i

are different for different illuminants l_i ; the matrix M_i is determined for each of these illuminants. This property of chromagenic camera responses is used to identify the illuminant in a scene, i.e., to solve the color constancy problem. Let $l_i(\lambda)$, $i = 1, \dots, m$ are the spectral power distributions of the possible known illuminants, and $r_j(\lambda)$, $i = 1, \dots, n$ is the reflectances of the n representative real world surfaces. For each illuminant i , we determine the camera responses without and with the chromagenic filter: C_i , and C_i^F respectively, which are $3 \times n$ matrices whose j^{th} column contains the camera responses of the j^{th} surface under the i^{th} illuminant. The transformation matrix M_i for the i^{th} illuminant is obtained using Eq. 4.

For a given test illuminant, we select an illuminant $l_{\text{est}}(\lambda)$ from all plausible illuminants l_i as the estimated illuminant, which gives the minimum error:

$$\text{est} = \underset{i}{\text{argmin}}(e_i), \quad i = 1, \dots, m \quad (5)$$

where e_i is the fitting error that can be calculated as:

$$e_i = \|M_i C - C^F\|, \quad i = 1, \dots, m. \quad (6)$$

The illuminant estimation algorithms are evaluated using the same framework as proposed by Hordley and Finlayson [7]. They recommended using the median angular error over the mean root mean square (RMS) error. Angular error is intensity independent and it has been widely used in the literature [1, 7]. Let $C_{l_{\text{est}}}$ and $C_{l_{\text{act}}}$ be the camera responses of a white reflectance under the estimated and the actual illuminant respectively, then the angular error e_{ang} is calculated as:

$$e_{\text{ang}} = \text{acos} \left(\frac{C_{l_{\text{act}}}^T C_{l_{\text{est}}}}{\|C_{l_{\text{act}}}\| \|C_{l_{\text{est}}}\|} \right) \quad (7)$$

3 Experiments

Here we first discuss the experimental setup and then present three different experiments and the results they produced: first the filter selection (Section 3.2), then the multispectral imaging (Section 3.3) and finally the illuminant estimation (Section 3.4).

3.1 Experimental Setup

The experimental setup comprises of a camera, filters, illuminants, reflectance data and test targets. A modern digital stereo camera from Fujifilm: the *Fujifilm FinePix REAL 3D W1* (in short, Fujifilm 3D) has been used. Fig. 1 illustrates a multispectral camera system constructed from this camera with a filter in front of one of its lenses. The sensitivities of this camera as measured and used by Shrestha et al. [16] has been used. The sensitivities of its left and right cameras are shown in Fig. 2.

To make the simulated multispectral system more realistic, as much as 4% normally distributed Gaussian noise is introduced as a random shot noise and 12-bit quantization noise is incorporated by directly quantizing the simulated camera responses after the application of the shot noise. Simulated D50 illuminant, and the CIE 1964 10° color matching functions are used for color computation as it is the logical choice for each color checker patches subtends more than 2° from the lens position. In order to evaluate the system, 63 patches of the Gretag Macbeth Color Checker DC has been used as the training target; and 122 patches remained after omitting the outer surrounding achromatic patches, multiple white patches at the center, and the glossy patches in the S-column of the DC chart have been used as the test target. The training patches have been selected using linear distance minimization method (LDMM) proposed by Pellegrini et al. [15]. A color whose associated system output vector has maximum norm among all the target colors is selected first. The method then chooses the colors of the training set iteratively based on their distance from those already chosen; the maximum absolute difference is used as the distance metric.

For the illuminant estimation, the 1995 Munsell surface reflectances (denoted as R) and the illuminants: the 87 measured training illuminants (L_{87}) and the 287 test illuminants (L_{287}), all the data from Barnard et al. [1] have been used. 265 optical filters of three different types: exciter, dichroic, and emitter from Omega are used. Transmittances of the filters available in the company website [14] have been used. One supplier has been chosen as a one point solution for the filters, and the Omega Optical Inc. has been chosen as they have a large selection of filters and data is available online.

3.2 Experiment I: Filter Selection

As discussed in Section 2, an optimal filter that produces the minimum estimation errors is selected. We have used the color estimation error (ΔE_{ab}^*) as the criteria, and the filter that produces the minimum ΔE_{ab}^* by the multispectral system and the minimum illuminant estimation (e_{ang}) error that lies within a acceptable error threshold values is selected from a given set of filters through exhaustive search.

For the illuminant estimation, the transformation matrices $M_i, i = 1 \dots 87$ are computed by imaging the whole surface reflectances, R under the training illuminants L_{87} , and they will be used to estimate the test illuminants. The test illuminants are estimated using the chromagenic algorithm.

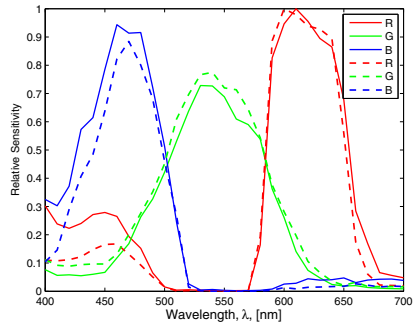


Fig. 2. Spectral sensitivities of the Fujifilm 3D camera (Left - solid, Right - dotted)

Depending on the application, an appropriate threshold values can be set for the color and the illuminant estimation errors. Here, as an illustration, we have chosen the $\Delta E_{ab}^* < 1.25$ and $e_{ang} < 2^\circ$ as the threshold values. Fig. 3 shows the XY plot of estimation errors with the 265 Omega filters, with the color estimation error along the X-axis and the angular error along the Y-axis. The filter selection algorithm chooses the XF1078 filter as shown in the figure. This filter has been selected with all the four spectral estimation methods discussed above. Fig. 4 shows the transmittance of the filter. This filter is then used in the next two experiments for the multispectral imaging and the illuminant estimation.

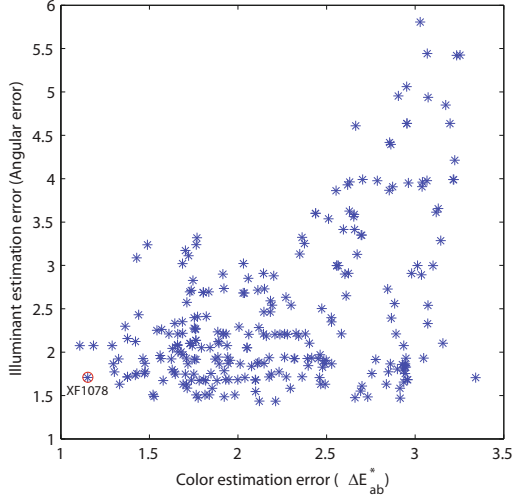


Fig. 3. Selection of a filter (red circled) from a set of 265 Omega filters

3.3 Experiment II: Multispectral Imaging

This experiment evaluates the proposed multispectral system constructed from the Fujifilm 3D and the selected filter, the Omega XF1078. Fig. 5 shows the normalized spectral sensitivities of the multispectral imaging system. The performance of the system has been investigated with all four spectral estimation methods (IM, LR, PN and NN) discussed in Section 2.1 and the results from three evaluation metrics (GFC, RMS and ΔE_{ab}^*) are reported. Table 1 shows the statistics (maximum/minimum, mean and standard deviation) of estimation errors side-by-side for the 3-channel and the proposed 6-channel systems.

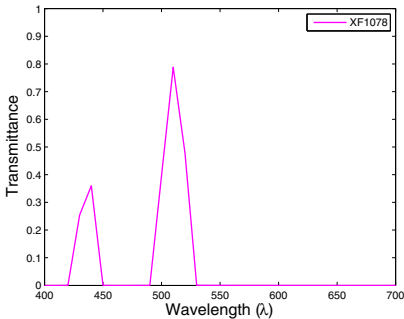


Fig. 4. Transmittance of the Omega XF1078 filter

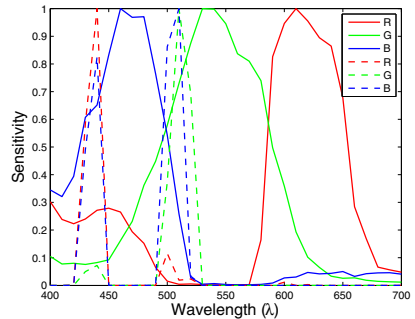


Fig. 5. Normalized spectral sensitivities of the 6-channel multispectral system

Table 1. Statistics of the estimation errors produced by the 6-channel system

Metric		3-Channel System				6-Channel System			
		IB	LR	PN	NN	IB	LR	PN	NN
GFC	Min	0.890	0.889	0.898	0.889	0.910	0.900	0.902	0.880
	Mean	0.990	0.990	0.990	0.990	0.996	0.996	0.996	0.996
	STD	0.017	0.017	0.015	0.017	0.009	0.010	0.010	0.012
RMS	Max	0.151	0.152	0.148	0.151	0.126	0.132	0.132	0.136
	Mean	0.031	0.031	0.029	0.031	0.021	0.020	0.020	0.021
	STD	0.023	0.023	0.021	0.023	0.018	0.019	0.018	0.020
ΔE_{ab}^*	Max	16.052	16.174	13.631	16.035	4.926	4.592	4.561	5.666
	Mean	3.486	3.542	3.552	3.484	1.153	1.130	1.196	1.050
	STD	3.320	3.359	2.766	3.316	0.958	0.816	0.825	0.872

The results show that the proposed 6-channel multispectral system outperforms the classic 3-channel system in terms of all the three metrics. We can also see that the performance of the four spectral estimation methods are comparable.

3.4 Experiment III: Illuminant Estimation

In this experiment, we use the real images generated from hyperspectral images of the eight natural scenes from Nascimento et al. [13]. The RGB images generated from the hyperspectral images using the Fujifilm3D camera and one of the 87 possible illuminant L_{87} are shown in Fig. 6. These hyperspectral images are available online in $820 \times 820 \times 33$ over 400-700nm bands in 10nm steps. However, the real image contents are less than 820×820 , but padded with zeros. Those padded empty data are removed and only real image contents are used. From these hyperspectral images, we obtain the unfiltered and the filtered versions of each image for every test illuminant L_{287} . The test illuminant is estimated in each case with the chromagenic algorithms using the transformation matrices $M_i, i = 1 \dots 87$ computed as discussed in Section 3.2. The median angular errors produced by the chromagenic algorithms along with the gray world [2] and the max-RGB [10] methods are given in Table 2. The results show significantly better estimation of the illuminant with the proposed system compared to the grayworld and max-RGB methods.

4 Discussion

The optimal filter used to construct the proposed system has been selected by setting acceptable error threshold values for the color and the illuminant estimation errors. As a further work, this selection could be made automatic with a single combined metric. Our experiments here are based on simulated images, and we assume that the two images are perfectly registered and there is no occlusion. It would be interesting to work further on experimental validation with real experiments taking into account the geometry of the stereo camera and two images from slightly different perspectives.



Fig. 6. The RGB images obtained from hyperspectral images of the 8 natural scenes from Nascimento et al. [13]

Table 2. Median angular errors for the 8 images generated from hyperspectral data of the scenes

Scene #	Gray World	Max RGB	Chromag.
1	5.50	4.75	3.52
2	9.86	21.85	6.19
3	9.45	3.20	4.62
4	5.50	4.75	3.52
5	7.32	11.04	2.05
6	2.83	6.94	2.21
7	0.99	2.12	1.64
8	2.87	3.10	3.49
Average	5.54	7.22	3.40

The experimental results show that the 6-channel multispectral system outperforms the 3-channel system significantly with all the four spectral estimation methods. As an example, the 6-channel system produces mean ΔE_{ab}^* of 1.05 with the NN method, while the 3-channel system produces 3.484. Moreover, the illuminant estimation with the chromagenic algorithm produces better results than the gray world and the max RGB methods. Finlayson et al. [4] have shown that the chromagenic based algorithms outperforms other color constancy algorithms like neural network, LP gamut mapping, Bayesian method, and color by correlation. The experimental results infer that simply selecting an optimal filter from a set of available filters produces promising results not only in the spectral and color reproduction from the multispectral imaging but also in the illuminant estimation. The performance could be improved further by using a set of a large number of filters, possibly from different manufacturers. The performance could also be improved significantly by using a custom designed filter [5].

The proposed system could be useful in digital color imaging where more accurate color imaging is required, and in color vision and robotics where color constant imaging is essential. Moreover, as the system is capable of acquiring multispectral and 3D images, it could also be used in multispectral imaging, for example, of culture and heritage.

5 Conclusion

This paper proposes a system capable of acquiring both a normal RGB image as well as a 6-channel multispectral image in a single shot, and at the same time capable of estimating the illuminant under which the image has been acquired. The system can be constructed from a off-the-shelf commercial stereo camera and

a filter. An optimal filter could either be selected from a set of available filters or custom designed. This allows a user flexibility to capture a color constant RGB image or a multispectral image or both. The system could be useful in many color imaging applications and computer vision.

References

1. Barnard, K., Cardei, V.C., Funt, B.: A comparison of computational color constancy algorithms. i: Methodology and experiments with synthesized data. *IEEE Transactions on Image Processing* 11(9), 972–984 (2002)
2. Buchsbaum, G.: A spatial processor model for object colour perception. *Journal of the Franklin Institute* 310(1), 1–26 (1980)
3. Connah, D.R., Hardeberg, J.Y.: Spectral recovery using polynomial models. In: *Color Imaging X: Processing, Hardcopy, and Applications*. SPIE Proceedings, vol. 5667, pp. 65–75 (2005)
4. Finlayson, G.D., Hordley, S.D., Morovic, P.: Chromagenic colour constancy. In: 10th Congress of the International Colour Association (AIC), Granada, Spain, pp. 8–13 (May 2005)
5. Finlayson, G.D., Hordley, S.D., Morovic, P.: Chromagenic filter design. In: 10th Congress of the International Colour Association (AIC), Granada, Spain, pp. 1079–1083 (May 2005)
6. Hardeberg, J.Y., Schmitt, F., Brettel, H.: Multispectral color image capture using a liquid crystal tunable filter. *Optical Engineering* 41(10), 2532–2548 (2002)
7. Hordley, S.D., Finlayson, G.D.: Reevaluation of color constancy algorithm performance. *J. Opt. Soc. Am. A* 23(5), 1008–1020 (2006)
8. Huang, H.H.: Acquisition of multispectral images using digital cameras. In: *Asian Association on Remote Sensing, ACRS* (2004)
9. Imai, F.H., Berns, R.S.: Spectral estimation using trichromatic digital cameras. In: *International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, pp. 42–49 (1999)
10. Land, E.H.: The retinex theory of color vision. *Scientific American* 237(6), 108–128 (1977)
11. Maloney, L.T., Wandell, B.A.: Color constancy: a method for recovering surface spectral reflectance. *J. Opt. Soc. Am. A* 3(1), 29–33 (1986)
12. Mansouri, A., Marzani, F.S., Gouton, P.: Neural networks in two cascade algorithms for spectral reflectance reconstruction. In: *IEEE International Conference on Image Processing*, pp. 2053–2056 (2005)
13. Nascimento, S.M.C., Ferreira, F.P., Foster, D.H.: Statistics of spatial cone-excitation ratios in natural scenes. *J. Opt. Soc. Am. A* 19(8), 1484–1490 (2002)
14. Omega: Omega filters. Omega Optical, Inc., <https://www.omegafilters.com/Products/Curvomatic> (last visited: February 2012)
15. Pellegrini, P., Novati, G., Schettini, R.: Selection of training sets for the characterisation of multispectral imaging systems. In: *PICS*, pp. 461–466 (2003)
16. Shrestha, R., Hardeberg, J.Y., Mansouri, A.: One-shot multispectral color imaging with a stereo camera. In: *Digital Photography VII, Electronic Imaging, Proceedings of SPIE/IS&T Electronic Imaging*, vol. 7876, p. 787609. SPIE, San Francisco (2011)

17. Shrestha, R., Mansouri, A., Hardeberg, J.Y.: Multispectral imaging using a stereo camera: Concept, design and assessment. *EURASIP Journal on Advances in Signal Processing* 2011(1) (September 2011)
18. Takita, K., Aoki, T., Sasaki, Y., Higuchi, T., Kobayashi, K.: High-accuracy subpixel image registration based on phase-only correlation. *IEICE Trans. Fundamentals* E86-A(8), 1925–1934 (2003)
19. Yamaguchi, M., Haneishi, H., Ohyama, N.: Beyond Red–Green–Blue (RGB): Spectrum-based color imaging technology. *Journal of Imaging Science and Technology* 52(1), 10201 (2008)