One–Shot Multispectral Color Imaging with a Stereo Camera

Raju Shrestha1, Jon Yngve Hardeberg1 and Alamin Mansouri2
1The Norwegian Color Research Laboratory, Gjøvik University College, Gjøvik, Norway
2The University of Burgundy, Auxerre, France

ABSTRACT
Multispectral color imaging is a promising technology, which can solve many of the problems of traditional RGB color imaging. However, it still lacks widespread and general use because of its limitations. State of the art multispectral imaging systems need multiple shots making it not only slower but also incapable of capturing scenes in motion. Moreover, the systems are mostly costly and complex to operate. The purpose of the work described in this paper is to propose a one–shot six-channel multispectral color image acquisition system using a stereo camera or a pair of cameras in a stereoscopic configuration, and a pair of optical filters. The best pair of filters is selected from among readily available filters such that they modify the sensitivities of the two cameras in such a way that they get spread reasonably well throughout the visible spectrum and gives optimal reconstruction of spectral reflectance and/or color. As the cameras are in a stereoscopic configuration, the system is capable of acquiring 3D images as well, and stereo matching algorithms provide a solution to the image alignment problem. Thus the system can be used as a “two-in-one” multispectral-stereo system. However, this paper mainly focuses on the multispectral part. Both simulations and experiments have shown that the proposed system performs well spectrally and colorimetrically.

Keywords: Multispectral, stereo, 3D, imaging, optical filter

1. INTRODUCTION
With the development and advancement of digital cameras, the acquisition and use of digital images have increased tremendously these days. Conventional image acquisition systems which capture images into three color channels, usually Red, Green and Blue, popularly known as RGB, are by far the most commonly used imaging systems. However, these suffer from several limitations like: they provide only color information, suffer from metamerism, limited to visual range and captured images are environment dependent. Multispectral imaging is an answer to these problems. Multispectral imaging systems capture image data at specific wavelengths across the electromagnetic spectrum. From a spectroscopic viewpoint, in the multispectral image, we obtain a spectrum in each pixel. Multispectral imaging provides more than color information while avoiding the metamerism and unlike conventional digital cameras, it is not limited to the visual range, rather it can also be used in near infrared, infrared and ultraviolet spectrum as well. It can significantly improve the color accuracy and makes color reproduction under different illumination environments possible with reasonably good accuracy.

Despite all these benefits of multispectral imaging technology, their use is still mostly limited to high-end airborne satellite imaging and research only. This is because of the limitations of the current state-of-the-art multispectral imaging systems. There are different types of multispectral imaging, among them most of them are filter-based which are also the one of our interest in this paper. In a typical filter-based imaging systems, either a set of traditional optical filters in a filter wheel, or a tunable filter capable of many different configurations are employed. It acquires a multispectral image by taking multiple shots, one at a time with one filter. A sensor used in a multispectral system may be a linear array as in CRISATEL where the images are acquired by scanning line-by-line. With a matrix sensor (CCD or CMOS) like in a monochrome camera, a whole image scene can be captured at once without need of scanning but it still needs multiple shots, one channel at a time. With a high quality tri-chromatic digital camera in conjunction with a set of appropriate optical filters allows to acquire unique spectral information. This method enables three channels of data to be captured per exposure as opposed to one. With a total of \( n \) colored filters, there are \( 3n + 3 \) camera responses for each pixel (including responses with no colored filters), correspondingly giving rise to a \( 3n + 3 \) channel multispectral image. This greatly increases the speed of capture and allows the use of technology that is readily and cheaply available and does not need to be specialized. However, these too need multiple shots to acquire a multispectral color image. Thus, most of these current state of the art multispectral imaging systems need several shots to capture multispectral images, with each chosen filter. Moreover, the construction and operation of such systems are complex and also costly. These severely limit the use of the multispectral systems. Several systems have been proposed aiming to circumvent these limitations.
Hashimoto proposed a two-shot 6-band still image capturing system using a commercial digital camera and a custom color filter. It captures a multispectral image in two shots, one with and one without the filter, thus resulting in a 6-channel output. The filter is custom designed in such a way that it cuts off the left side (short wavelength domain) of the peak of original spectral sensitivity of blue and red, and also cuts off the right side (long wavelength domain) of the green. It shows the 6-channel system results more accurate color and wider color range. The problem with this is that it still needs two shots and is therefore incapable of capturing objects in motion.

Ohsawa et al. proposed a one-shot 6-band HDTV camera system. In their system, the light is divided into two optical paths by a half mirror, and incident on two conventional three-charge coupled device (CCD) cameras after transmission through the specially designed interference filters inserted in each optical path. The two HDTV cameras capture three–band images in sync to compose each frame of the six band image. The total spectral sensitivities of the six band camera are the convoluted spectral characteristics of the optical components: the objective lens, the half mirror, the IR cutoff filter, the interference filters, the CCD sensors, etc. It allows one-shot image capture. However, it needs custom designed filters and complex optics making it still far from being practical.

This paper proposes a fast and practical solution to multispectral imaging with the use of two modern digital cameras and a pair of readily available optical filters.

2. PROPOSED ONE–SHOT ACQUISITION WITH A STEREO CAMERA

The proposed multispectral system uses a stereo camera or two modern digital (RGB) cameras in a stereoscopic configuration, and a pair of appropriate optical filters in front of each camera. The idea is to select one or two appropriate optical filters from among a set of readily available filters, so that they will modify the sensitivities of one or two cameras to give six, if not quite, but reasonably well spaced channels in the visible spectrum so as to give optimal reconstruction of the scene spectral reflectances and/or the color. Selection of the filters can be done using a filter selection method presented in Section 2.1. Depending upon the sensitivities of the two cameras, it is also possible to select a single filter for one camera while the other camera left without any filter. The subsequent combination of the images from the two cameras provides a multispectral image of the acquired scene. The multispectral system thus constructed will be of six channel, 3 each contributed from the two cameras. It is faster, as the image can be acquired in one-shot, thus making it capable of taking multispectral images of objects in motion as well.

For natural and man-made surfaces whose reflectance are more or less smooth, it is recommended to use as few channels as possible and we have found that six channels can be good enough in such cases. Moreover, by joining the two cameras in a stereoscopic configuration, it allows us to capture 3D stereo images also. This makes the system capable of acquiring both the multispectral and stereo 3D images simultaneously. Stereo imaging needs extra procedure and precision issues to be taken into account and this paper assumes that these issues are taken care of. The stereoscopic configuration allows the alignment of the images captured by two cameras through the use of one or more stereo matching algorithms. Figure 1 shows a multispectral-stereo system constructed from a modern stereo camera - Fujiﬁlm FinePix REAL 3D W1 (Fujifilm3D) and two optical filters in front of the two lenses. One–shot acquisition can be made possible by using two cameras with a sync controller available in the market. The system, thus, acts as a two-in-one multispectral-stereo system. The two cameras need not be of same type, instead, any two cameras can be used, provided the two are operated in the same resolution. Thus, the proposed multispectral system is a faster, cheaper and practical solution, as it is the one–shot acquisition which can be constructed from even commercial digital cameras and readily available filters; and additionally, capable of capturing 3D images at the same time.

The proposed multispectral system has been investigated with both simulation and experimental approaches. Section presents the proposed multispectral system model. Before this, we present the algorithm used for the selection of filters.
2.1 Selection of Filters

Several methods have been proposed for the selection of filters, particularly for multi-shot based multispectral color imaging\cite{15,30,31}. In our study, as we have to choose just two filters from the set of filters, the exhaustive search method is feasible and a logical choice because of its guaranteed optimal results. For selecting $k$ (here $k=2$) filters from given set of $n$ filters, it requires $P(n,k) = \frac{n!}{(n-k)!}$ permutations. When two same type of cameras are used, the problem reduces to combinations instead of permutations, i.e. $C(n,k) = \frac{n!}{k!(n-k)!}$ combinations. In order to reduce the computational complexity, infeasible filter pairs are excluded based on a secondary criterion: “Filter pairs that result in a maximum transmission factor of less than forty percent, and less than ten percent of the maximum transmission factor in one or more channels are excluded.”

2.2 Multispectral System Model

Let $s_i$ be the spectral sensitivity of the $i$th channel, $t$ 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 $n$ denotes the acquisition noise. The camera response corresponding to the $i$th channel $C_i$ is then, given by the multispectral camera model as

$$C_i = S_i^T \text{Diag}(L)R + n_i; \quad i = 1, 2, ..., K,$$

(1)

where $S_i = \text{Diag}(t)s_i$, $n_i$ is the channel acquisition noise, and $K$ is the number of channels, which is 6 here in our system.

Spectral reflectances of the image scene can be reconstructed from the six-channel multispectral camera response $C$ as discussed in Section 3.

3. SPECTRAL RECONSTRUCTION AND EVALUATION

The reconstructed/estimated reflectance ($\tilde{R}$) is obtained for the corresponding original reflectance ($R$) from the camera responses for the training and test targets $C$ and $C_{\text{train}}$ respectively, using different estimation methods. In the simulation approach, results with three popular methods, Maloney and Wandell (MW)\cite{32} Imai and Berns (IB)\cite{17} and Constrained Least Square-Wiener (Wiener)\cite{33} are investigated. The estimated reflectances with these three methods are given by the following equations:

$$\tilde{R}_{\text{MW}} = B(E^TB)^+C$$

(2)

$$\tilde{R}_{\text{IB}} = BB^T R_{\text{train}} C_{\text{train}}$$

(3)

$$\tilde{R}_{\text{Wiener}} = R_{\text{train}} R_{\text{train}}^T E (E^T R_{\text{train}} R_{\text{train}}^T E)^{-1} C$$

(4)

Here $E = \text{Diag}(L)S$, and $B$ contains the basis vectors obtained by Singular Value Decomposition (SVD) of the training reflectances $R_{\text{train}}$. The number of basis vectors to be used are determined by optimization of the estimation errors. $X^+$ denotes the pseudo-inverse of $X$. To handle ill-posed and over fitting problem in the pseudo-inverse computation, Tikhonov regularization\cite{34} has been implemented.

In the experimental approach, the polynomial (PN)\cite{35,36} and neural network (NN)\cite{37} methods of reconstructions are used as they do not require the spectral sensitivities of the camera. The estimated reflectance with these methods are given by the following equations:

$$\tilde{R}_{\text{PN}} = MC_p$$

(5)

$$\tilde{R}_{\text{NN}} = wC$$

(6)

where $M = R_{\text{train}} C_{p,\text{train}}^+$, $C_p$ is the polynomials of $n$ degree formed from the channel responses in the camera response matrix $C$, and $w$ is obtained from the training reflectance $R_{\text{train}}$ and corresponding camera response $C_{\text{train}}$ from the neural network Delta rule.
The reconstructed reflectance are evaluated using spectral as well as colorimetric metrics. Most commonly used RMS (Root Mean Square) error has been used as the spectral metric, and $\Delta E_{ab}^*$ (CIELAB Color Difference) as the colorimetric metric. These metrics are given by the equations:

$$RMS = \sqrt{\frac{1}{n} \sum_{j=1}^{n} [\tilde{R}(\lambda_j) - R(\lambda_j)]^2}$$

(7)

$$\Delta E_{ab}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}$$

(8)

CIE D50 illuminant and CIE XYZ 1964 are used for computation of $\Delta E_{ab}^*$.

4. SIMULATION AND RESULTS

Simulation has been carried out with different stereo cameras (or camera pairs) whose spectral sensitivities are known or measured. It takes a pair of filters at a time, computes the camera responses, obtains the reconstructed spectral reflectances and evaluates the estimation errors (spectral and colorimetric) as discussed in Section 3. As there is always acquisition noise introduced into the camera outputs, to make the simulation more realistic, simulated random shot noise and quantization noise are introduced. Recent measurements of noise levels in a trichromatic camera suggest that realistic levels of shot noise are between 1% and 2%. So 2% normally distributed Gaussian noise is introduced as a random shot noise in the simulation. And 12-bit quantization noise is incorporated by directly quantizing the simulated responses after the application of shot noise.

The simulation study has been conducted with a pair of Nikon D70 cameras, Nikon D70 and Canon 20D pair, and Fujifilm3D stereo camera. Previously measured spectral sensitivities of the Nikon D70 and Canon 20D cameras are used, and those of the Fujifilm3D camera are measured using Bentham TMc300 monochromator. Figure 2 shows these spectral sensitivities. 265 optical filters of three different types: exciter, dichroic, and emitter from Omega are used in the simulation. Rather than mixing filters from different vendors, one vendor has been chosen as a one point solution for the filter, and Omega has been chosen as it has a large selection of filters and data is available online. The Gretag Macbeth Color Checker DC has been used as the training target and the classic Macbeth Color Checker as the test target in the simulation. The outer surrounding achromatic patches and the glossy patches in the S-column of the DC chart have been omitted from the training dataset.

The estimated reflectances are evaluated using the two evaluation metrics: spectral RMS and $\Delta E_{ab}^*$. CIE D50 illuminant and CIE XYZ 1964 color matching functions are used for color computation. The best pair of filters is exhaustively searched as discussed in Section 2.1 according to each of the evaluation metrics, from among all available filters with which the multispectral system can optimally reconstruct the reflectances of the twenty four classic Macbeth Color Checker patches. The results corresponding to minimum mean of the evaluation metrics are obtained. To speed up the process, the filter combinations not fulfilling the constraints described in Section 2.1 are skipped. The 265 filters leads to more than 70,000 possible permutations. The constraints introduced reduce the processing down to less than 20,000 permutations.

Figure 2. Normalized spectral sensitivities of the cameras
Table 1 shows statistics of estimation errors (RMS and $\Delta E^*_{\text{ab}}$) and selected filter pairs in all cases. RMS % in the table means the RMS value multiplied by 100. For a given stereo camera or a pair of cameras and for a given evaluation metric, all three estimation methods pick the same pair of filters with the exception of NikonD70-Canon20D pair for $\Delta E^*_{\text{ab}}$ metric, still the filter types are quite similar. The simulation picks different filter pairs for different stereo cameras, and the different evaluation metrics.

### Table 1. Statistics of estimation errors

<table>
<thead>
<tr>
<th>Camera System</th>
<th>Method</th>
<th>3-Channel System</th>
<th>6-Channel System</th>
<th>Filter</th>
<th>3-Channel System</th>
<th>6-Channel System</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>NikonD70</td>
<td>MW</td>
<td>10.60</td>
<td>6.82</td>
<td>2.04</td>
<td>2.51</td>
<td>1.02</td>
<td>1.04</td>
</tr>
<tr>
<td>NikonD70</td>
<td>IB</td>
<td>11.37</td>
<td>10.23</td>
<td>2.75</td>
<td>2.54</td>
<td>1.02</td>
<td>1.04</td>
</tr>
<tr>
<td>NikonD70</td>
<td>Wiener</td>
<td>11.83</td>
<td>6.26</td>
<td>1.98</td>
<td>2.55</td>
<td>1.02</td>
<td>1.05</td>
</tr>
<tr>
<td>NikonD70-Canon20D</td>
<td>MW</td>
<td>10.60</td>
<td>6.82</td>
<td>2.04</td>
<td>2.50</td>
<td>1.00</td>
<td>1.58</td>
</tr>
<tr>
<td>NikonD70-Canon20D</td>
<td>IB</td>
<td>11.37</td>
<td>10.23</td>
<td>2.75</td>
<td>2.54</td>
<td>1.00</td>
<td>1.62</td>
</tr>
<tr>
<td>NikonD70-Canon20D</td>
<td>Wiener</td>
<td>11.83</td>
<td>6.26</td>
<td>1.98</td>
<td>2.57</td>
<td>1.01</td>
<td>1.57</td>
</tr>
<tr>
<td>Fujifilm 3D</td>
<td>MW</td>
<td>8.73</td>
<td>19.61</td>
<td>4.59</td>
<td>2.55</td>
<td>1.03</td>
<td>2.29</td>
</tr>
<tr>
<td>Fujifilm 3D</td>
<td>IB</td>
<td>9.11</td>
<td>19.26</td>
<td>4.65</td>
<td>2.58</td>
<td>1.04</td>
<td>2.30</td>
</tr>
<tr>
<td>Fujifilm 3D</td>
<td>Wiener</td>
<td>9.56</td>
<td>20.70</td>
<td>4.62</td>
<td>2.68</td>
<td>1.04</td>
<td>2.12</td>
</tr>
</tbody>
</table>

Figures 3, 5 and 7 show the transmittances of the filters selected for the optimal results of the two metrics used in the multispectral-stereo systems constructed with NikonD70-NikonD70, NikonD70-Canon20D and Fujifilm3D cameras respectively. And the corresponding normalized effective channel sensitivities of the multispectral systems are shown in Figures 4, 6 and 8.
Figure 4. Multispectral 6-channel normalized sensitivities (NikonD70–NikonD70)

Figure 5. Transmittance of the filter pairs selected for minimum \( \Delta E_{ab}^* \) (NikonD70–Canon20D)

Figure 6. Multispectral 6-channel normalized sensitivities (NikonD70–Canon20D)
Figure 7. Transmittances of the filters selected (Fujifilm3D)

Figure 8. Multispectral 6-channel normalized sensitivities (Fujifilm3D)

As an illustration, Figure 9 shows the measured original, and the estimated spectral reflectances of the 16 test patches of the classic Macbeth Color Checker captured with the six-channel multispectral system constructed from Fujifilm3D, and using Wiener estimation method for minimum mean RMS error. The simulation results show that the proposed six-channel multispectral system outperforms classical 3-channel camera systems, both spectrally and colorimetrically. The improvement is significant with the decrease in mean RMS error from about 3% to 1%, and by 3 $\Delta E_{ab}^*$ units on the average. And, the reconstructed spectral reflectances, as can be seen in Figure 9, is very close to the original ones. The simulation results, thus, show promising results in favor of our proposed approach of multispectral imaging with two RGB cameras or a stereo camera. We can see that sensitivities of the cameras are modified in a sensible way and that the dominant wavelengths are quite nicely spread out in the spectrum. This clearly shows that the use of two RGB cameras or a stereo camera with appropriate use of filters can function well as a multispectral system.
5. EXPERIMENT AND RESULTS

Promising results from the simulation encouraged to extend the study of the system with real experiments. Experiments were carried out with the Fujifilm3D stereo camera and a Schott BG-36 multiband filter (see Figure 10). Like in the simulation, Macbeth Color Checker DC has been used as the training target and classic Macbeth Color Checker as the test target. The multispectral camera system has been constructed by placing the Schott BG-36 filter in front of one lens of the stereo camera, and the other lens left open. Even though the filter used is not optimal, if gives promising results, we can expect that the system performs much better with an optimal filter selection. During the experiments, the camera has been set to a fixed configuration (mode: manual, flash: off, ISO: 100, exposure time: 1/50s, aperture: F3.7, white balance: fine, 3D file format: MPO, image size: 3648 × 2736).

The spectral power distribution of the light source (Daylight D50 simulator, GretagMacbeth SpectraLight III) under which the experiments have been carried out has been measured with Minolta CS-1000 spectroradiometer. Spectral reflectances of the color charts have been measured with X-Rite Eye One Pro spectrophotometer. Spectral characterization of the Fujifilm3D camera has been done with Bentham TMc300 monochromator. Both the left and right cameras have been corrected for linearity, DC noise and non-uniformity. The system then acquires the images of the color charts, and the estimated spectral reflectances of the test patches have been obtained using the polynomial and neural network approaches of spectral reconstruction discussed in Section 3. The degree of polynomial and the neural network parameters have been optimized and the the polynomial of degree 2 has been found to give the best result. Statistics of estimation errors for the two methods are given in Table 2.
Table 2. Statistics of estimation errors

<table>
<thead>
<tr>
<th>Method</th>
<th>3-Channel System</th>
<th>6-Channel System</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMS %</td>
<td>$\Delta E^{*}_{ab}$</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>Mean</td>
</tr>
<tr>
<td>Polynomial (degree 2)</td>
<td>9.36</td>
<td>4.38</td>
</tr>
<tr>
<td>Neural Network</td>
<td>15.50</td>
<td>5.55</td>
</tr>
</tbody>
</table>

As an illustration, the measured original and the estimated spectral reflectances of the 16 test patches of the classic Macbeth Color Checker with the neural network approach have been given in Figure 11.

![Figure 11. Reflectances with 6-channel system (Neural Network method)](image)

Statistics of estimation errors from the results shows that the proposed six-channel multispectral system performs better than the 3-channel system both spectrally and colorimetrically in terms of mean metric values. Both polynomial and neural network methods result in favor of proposed six-channel multispectral-stereo system with improvement in error estimations from 4.9% down to 3.7% mean RMS, and from 7.1 down to 6.4 mean $\Delta E^{*}_{ab}$ on the average.

6. CONCLUSION

In this paper, we have proposed a one–shot multispectral image acquisition with a stereo camera. The proposed system is simple to construct out of commercial digital cameras, and a pair of filters selected from readily available filters in the market. It, therefore, could be a fast, practical, and cheaper solution to multispectral imaging, useful in a variety of applications. Both the simulation and experimental results show that the six-channel multispectral system perform better than the traditional 3-channel cameras both spectrally and colorimetrically. Moreover, stereo configuration allows to acquire stereo 3D images simultaneously along with the multispectral image. We assume that the well known problem of occlusion in 3D imaging would be handled to the possible extent with the use of best known stereo matching algorithm, and this could itself be an extension work to this study in the future.
REFERENCES


