

Evaluating Color Reproduction Accuracy of Stereo One-shot Six-band Camera System

M. Tsuchida¹, S. Sakai², K. Ito², T. Kawanishi¹, K. Kashino¹, J. Yamato¹, and T. Aoki²

1. NTT Communication science laboratories, NTT Corporation / 3-1 Morinosato-Wakamiya, Atsugi-shi, Kanagawa, Japan

2. Graduate School of Information Sciences, Tohoku University / 6-6-05, Aramaki Aza Aoba, Aoba-ku, Sendai-shi, Japan

Abstract

A stereo one-shot six-band image-capturing system that combines multiband and stereo imaging techniques has been developed. This system can acquire both spectral color information and depth information at the same time. It worked well for two-dimensional objects that have a wavy structure like a tapestry. In this paper, we discuss the accuracy of color reproduced by using this system. In the experiments, several stereo-pair images were captured while the distance between the two cameras was changed. The estimated color and spectral reflectance of a color chart are compared with the measurement results obtained with a spectrometer and estimation results obtained with the two-shot six-band camera system. It is confirmed the average color difference for 24 color patches of a color chart between measurement data and estimation result is $dE_{ab}^* = 1.21$ at maximum. Finally, demonstrations of image-capturing of three-dimensional object, image transformation, and color reproduction are presented.

Introduction

The final goal of our research project is to use imaging technology to achieve high-fidelity color, gloss, and texture reproduction and three-dimensional (3D) shape, micro-structure, and movement reproduction of objects for archiving (e.g. cultural heritage and medical applications). Multi-band imaging technology is a solution for accurate color reproduction. Although several types of multiband camera system have been developed [1-4], all of them are multi-shot type systems and they cannot take images of moving objects. Ohsawa et al. have developed a six-band HDTV camera system [5]. However, the system requires very expensive customized equipment. In order to make multiband technology pervasive, equipment costs must be reduced and the systems have to be able to take images of moving objects. To meet these requirements, a stereo one-shot six-band image capturing system that combines multiband and stereo imaging techniques has been developed [6]. This system can acquire both spectral color information and depth information at the same time. It worked well for 2D objects that have a wavy structure like a tapestry at the time. In this paper, we discuss the accuracy of color reproduced by using this system. In the experiments, several stereo-pair images were captured while the distance between two cameras was changed. The estimated color and spectral reflectance of a color chart are compared with the measurement results obtained with a spectrometer and estimation results obtained with the two-shot six-band camera systems [4].

Stereo six-band camera system

Principle of color estimation using multi-band data

As Fig. 1 shows, an object surface reflects light from illumination. Let the illumination spectrum and spectral reflectance be $W(\lambda)$ and $f(\lambda)$, respectively. The observed spectrum $I(\lambda)$ can then be represented as

$$I(\lambda) = W(\lambda)f(\lambda), \quad (1)$$

where λ is wavelength. Let us consider a situation where the reflected light is captured by an N -band sensor. Let the spectral sensitivity of the sensor be $\mathbf{S} = [S_1(\lambda), S_2(\lambda), \dots, S_N(\lambda)]^T$. Let the matrix whose diagonal elements represent the spectral power distribution of illumination be \mathbf{W} . Equation (1) can then be rewritten into a vector representation as

$$\mathbf{I} = \mathbf{Wf}. \quad (2)$$

By using the Wiener estimation method [Pratt 1976], spectral reflectance is estimated from the camera signal, $\mathbf{c} = \mathbf{SWf} = \mathbf{Hf}$, as

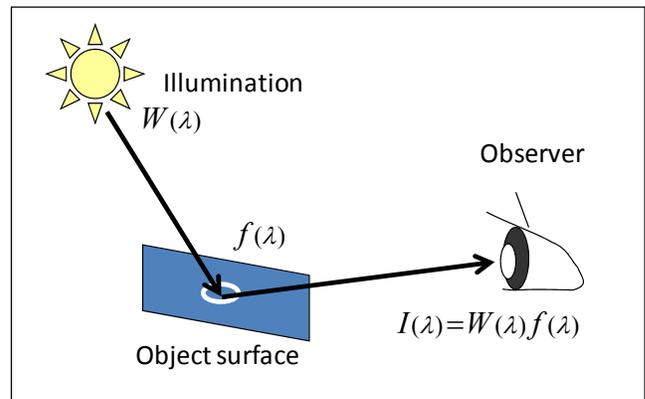


Fig.1: Geometrical setup

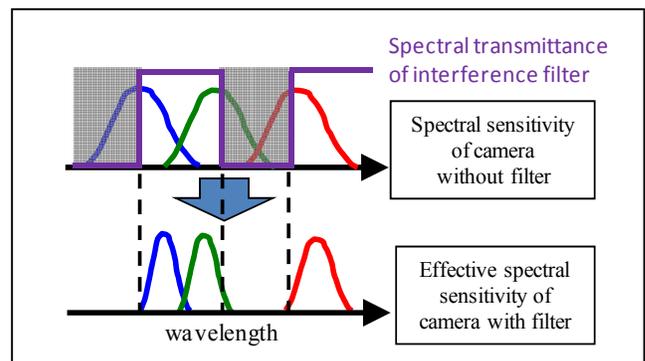


Fig.2: Principle of six-band image capturing

$$\hat{\mathbf{f}} = \mathbf{M}\mathbf{c}, \quad \mathbf{M} = \mathbf{R}\mathbf{H}^t \{\mathbf{H}\mathbf{R}\mathbf{H}^t\}^{-1} \quad (3)$$

where \mathbf{M} is the Wiener estimation matrix obtained from \mathbf{H} ,



Fig.3: Stereo six-band camera system

and \mathbf{R} is a priori knowledge about the spectral reflectance of objects.

In the Wiener estimation method, we used the correlation matrix \mathbf{R} modeled on a first-order Markov process covariance matrix in the form

$$\mathbf{R} = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ \rho & 1 & \rho & \dots & \rho^{N-2} \\ \rho^2 & \rho & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{N-1} & \rho^{N-2} & \dots & \dots & 1 \end{pmatrix}, \quad (4)$$

where $0 \leq \rho \leq 1$ is the adjacent element correlation factor and we selected $\rho = 0.999$ for our experiments.

Six-band image capturing using an interference filter whose spectral transmittance is comb-shaped

Figure 2 shows the principle of six-band image capturing using an interference filter whose spectral transmittance is comb-shaped. Our proposed camera system uses a pair of digital cameras. The filter is mounted in front of the lens of the camera and used to capture a specialized RGB image. The other camera is used to capture a normal RGB image. The filter mounted in front of the camera lens cuts off the left sides, i.e., the short-wavelength domain, of the peaks of both the blue and red in the original spectral sensitivity of the camera. It also cuts-off the green's right side, i.e., the long-wavelength domain.

Stereo one-shot six-band camera system

Our system consists of two commercial digital cameras, and a custom interference filter. The filter is mounted in front of the lens of one camera (see Fig.3). The captured two images have parallax. Therefore, to generate a six-band image from the pair of images, one image should be transformed to adjust it to the other image.

As a first step to do that, corresponding points between two images are detected. Although the two cameras take images of the



(a) Image captured without filter (b) Image captured with filter

Fig.4: Captured images.

same target object, the color balance between the two images is quite different because of the interference filter mounted in front of lens of the one camera (Fig.4). General detection methods cannot work well in such a case. Then, we use the phase-only correlation method (POC) to detect the corresponding detection [7]. POC is a scale- and rotation-invariant pattern detection method that uses phase information. POC also has robustness against color.

Next, the shape of the image captured with the interference filter is adjusted to that of the other image using the detected corresponding points. Projective transformation is a simple method and works well for two-dimensional (2D) objects. When the target object has 3D shape, nonlinear transformation is better. The thin-plate spline (TPS) model [8] was used for image transformation in this work. The resultant two three-band images are combined into a six-band image.

Phase-Only Correlation (POC) function

Consider two $N_1 \times N_2$ images, $f(n_1, n_2)$ and $g(n_1, n_2)$, where we assume that the index ranges are $n_1 = -M_1, \dots, M_1$ and $n_2 = -M_2, \dots, M_2$ for mathematical simplicity, and hence $N_1 = 2M_1 + 1$ and $N_2 = 2M_2 + 1$. Let $F(k_1, k_2)$ and $G(k_1, k_2)$ denote the 2-D discrete Fourier transforms (DFTs) of the two images. $F(k_1, k_2)$ and $G(k_1, k_2)$ are given by

$$F(k_1, k_2) = \sum_{n_1, n_2} f(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} = A_F(k_1, k_2) e^{j\theta_F(k_1, k_2)}, \quad (5)$$

$$G(k_1, k_2) = \sum_{n_1, n_2} g(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} = A_G(k_1, k_2) e^{j\theta_G(k_1, k_2)}, \quad (6)$$

where $k_1 = -M_1, \dots, M_1$, $k_2 = -M_2, \dots, M_2$, $W_{N_1} = e^{-j\frac{2\pi}{N_1}}$,

$W_{N_2} = e^{-j\frac{2\pi}{N_2}}$, and the operator \sum_{n_1, n_2} denotes

$\sum_{n1=-M1}^{M1} \sum_{n2=-M2}^{M2}$. $A_F(k1, k2)$ and $A_G(k1, k2)$ are amplitude components, and $e^{j\theta_F(k1, k2)}$ and $e^{j\theta_G(k1, k2)}$ are phase components.

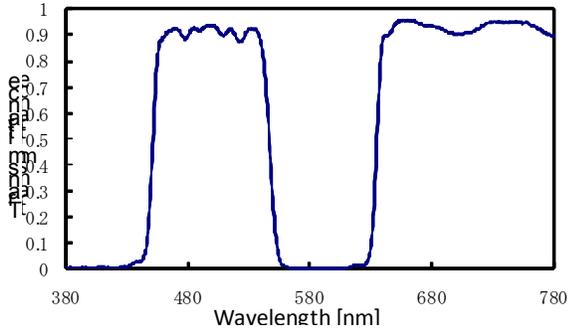


Fig.5: Spectral transmittance of interference filter.

The cross spectrum $R(k1, k2)$ between $F(k1, k2)$ and $G(k1, k2)$ is given by

$$\begin{aligned} R(k1, k2) &= \overline{F(k1, k2)G(k1, k2)}, \\ &= \overline{A_F(k1, k2)A_G(k1, k2)e^{j\theta(k1, k2)}}, \end{aligned} \quad (7)$$

where $G(k1, k2)$ denotes the complex conjugate of $F(k1, k2)$ and $\theta(k1, k2) = \theta_F(k1, k2) - \theta_G(k1, k2)$. On the other hand, the cross-phase spectrum (or normalized cross spectrum) $\hat{R}(k1, k2)$ is defined as

$$\begin{aligned} \hat{R}(k1, k2) &= \frac{\overline{F(k1, k2)G(k1, k2)}}{\overline{F(k1, k2)G(k1, k2)}} \\ &= e^{j\theta(k1, k2)}, \end{aligned} \quad (8)$$

The POC function $\hat{r}(n1, n2)$ is the 2D inverse DFD of $\hat{R}(k1, k2)$ and is given by

$$\hat{r}(n1, n2) = \frac{1}{N1N2} \sum_{k1k2} \hat{R}(k1, k2) W_{N1}^{-k1n1} W_{N2}^{-k2n2}, \quad (9)$$

where $\sum_{n1, n2}$ denotes $\sum_{n1=-M1}^{M1} \sum_{n2=-M2}^{M2}$.

Sub-pixel Image Registration

Consider $f_c(x1, x2)$ as a 2-D image defined in continuous space with real-number index $x1$ and $x2$. Lets $\delta1$ and $\delta2$ represent sub-pixel displacement of $f_c(x1, x2)$ in $x1$ and $x2$ directions, respectively. So, the displaced image can be represented as $f_c(x1 - \delta1, x2 - \delta2)$. Assume that $f(n1, n2)$ and

$g(n1, n2)$ are spatially sampled images of $f_c(x1, x2)$ and $f_c(x1 - \delta1, x2 - \delta2)$, defined as

$$f(n1, n2) = f_c(x1, x2)|_{x1=n1T1, x2=n2T2}, \quad (10)$$

$$g(n1, n2) = g_c(x1 - \delta1, x2 - \delta2)|_{x1=n1T1, x2=n2T2}, \quad (11)$$

where $T1$ and $T2$ are the spatial sampling intervals, and index ranges are given by $n1 = -M1, \dots, M1$ and $n2 = -M2, \dots, M2$. Let $F(k1, k2)$ and $G(k1, k2)$ be the 2-D DFTs of $f(n1, n2)$ and $g(n1, n2)$, respectively. Considering the difference of properties between the Fourier transform defined in continuous space and that defined in discrete space carefully, we can now say that

$$G(k1, k2) \cong F(k1, k2) \cdot e^{-j\frac{2\pi}{N1}k1\delta1} e^{-j\frac{2\pi}{N2}k2\delta2}. \quad (12)$$

Thus, $\hat{R}(k1, k2)$ is given by

$$\hat{R}(k1, k2) \cong e^{-j\frac{2\pi}{N1}k1\delta1} e^{-j\frac{2\pi}{N2}k2\delta2}. \quad (13)$$

The POC function $\hat{r}(n1, n2)$ will be the 2D inverse DFT of $\hat{R}(k1, k2)$, and is given by

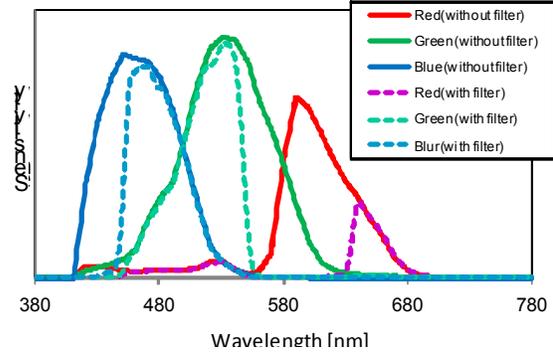


Fig.6: Spectral sensitivity of the camera system.

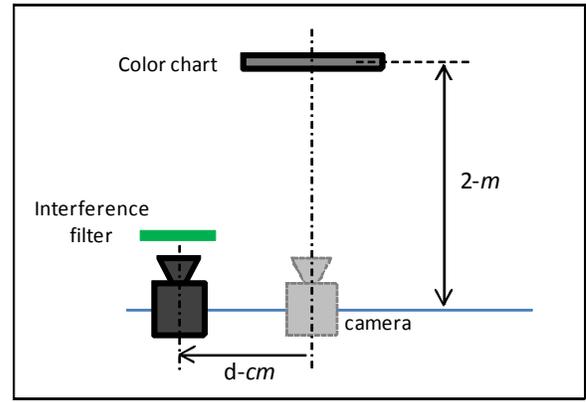


Fig.7: Experimental setup.

$$\begin{aligned} \hat{r}(n1, n2) &= \frac{1}{N1N2} \sum_{k1k2} \hat{R}(k1, k2) W_{N1}^{-k1n1} W_{N2}^{-k2n2} \\ &\cong \frac{\alpha}{N1N2} \frac{\sin\{\pi(n1 + \delta1)\}}{\sin\{\frac{\pi}{N1}(n1 + \delta1)\}} \frac{\sin\{\pi(n2 + \delta2)\}}{\sin\{\frac{\pi}{N2}(n2 + \delta2)\}}, \end{aligned} \quad (14)$$

where $\alpha = 1$. The above equation represents the shape of the peak for the POC function for common images that are minutely displaced from each other. The peak position of the POC function corresponds to the displacement between the two images. We can prove that the peak value α decreases (without changing the function shape itself), when small noise components are added to the original images. Hence, we assume $\alpha \leq 1$ in practice.

Experiments

Characteristics of the experimental equipment

We used a consumer-model digital camera (D700, NIKON) which can write out raw image data without any color correction as NEF file format. This camera can take 12-MPixel image, and its bit-depth is 14 bits. We analyzed the NEF file format and converted NEF file into general raw file format. Figure 5 and 6

show spectral transmittance of the interference filter and spectral sensitivity of the camera used in this experiment. Note that the camera does not have sensitivity lower than 400 nm and higher than 700 nm because UV- and IR-cut filters are attached to the image sensor.

Color reproduction of the color chart

We evaluate the accuracy of reproduced color and spectral reflectance when the distance between the two cameras is changed. Macbeth ColorChecker™ was used as a target object and the digital camera (Nikon D700) was used. Focus length of the lens was 105 mm. The distance between camera and color chart was 2 m. As a first step, the first image was captured without the interference filter. Then the interference filter was attached in front of the camera lens and the second image was captured. Next, the camera with the filter was moved 15-cm horizontally in 1-cm intervals (see Fig. 7) and filtered images were captured at each position. The exposure setting (shutter speed, iris, etc.) of the camera were fixed. Projective transformation was used to generate a six-band image of color chart.

Figure 8 shows the estimated spectral reflectance of the 24 color patches of Macbeth ColorChecker™. The estimation results when the camera's moving distance $d=0, 5, 10,$ and 15 cm are plotted. To evaluate the estimation results, we also measured the spectral reflectance using a spectrometer, and the measurement results are plotted on the same graphs. We can see that distance the camera is moved does not affect the estimation of spectral reflectance under this experimental geometry. Good estimation results were obtained between 400 and 700 nm wavelength. There are some errors in near UV- and near IR-wavelength domain caused by the UV- and IR-cut filter on the image sensor.

Next, color difference dE_{ab} between measured color and estimated color was calculated. Averaged color differences of 24 color patches $dE_{ab} = 0.97, 1.15$ and 1.21 when $d = 0$ cm, 5 cm, 10 cm and 15 cm.

Image capturing of 3D object

A stereo image of the Japanese doll shown in Fig. 4 was captured with the stereo six-band camera system shown in Fig. 3. The distance between the camera system and object was 2 m, the same as in the experiment described above section, and the distance between the optical axes of the two cameras was 13 cm. Since the object has a 3D shape, a nonlinear method based on TPS was used for image transformation. Two color reproduction results before and after image transformation are shown in Fig. 9. We can see some artifacts of green, which were caused by parallax between the two captured images. On the other hand, few artifacts remain on the image after image transformation.

Summary

We evaluated color reproduction results for a six-band image captured with a stereo one-shot six-band camera system while the distance between the two cameras was changes. Estimated spectral reflectance and CIE $L^*a^*b^*$ value of the color chart were compared with the measurement data obtained with a spectrometer, and it was confirmed that average color difference of the 24 color patches of color chart between the measurement data and estimation achieved was 1.21 at maximum.

References

- [1] M. Yamaguchi et al., "Natural Vision: Visual Telecommunication based on Multispectral Technology.", Proc. International Display Workshop 2000, 1115-1118. (2000).
- [2] S. Tominaga et al., "Object Recognition by Multi-Spectral Imaging with a Liquid Crystal Filter.", Proc. Conference on Pattern Recognition, vol.1, 708-711. (2000)
- [3] S. Helling et al., "Algorithms for spectral color stimulus reconstruction with a seven-channel multispectral camera." Proc. Second European Conference on Colour in Graphics, Imaging, and Vision (CGIV2004), 254-258. (2004)
- [4] M. Hashimoto, "Two-Shot type 6-band still image capturing system using Commercial Digital Camera and Custom Color Filter.", Proc. Fourth European Conference on Colour in Graphics, Imaging, and Vision (CGIV2008). (2008).
- [5] K. Ohsawa et al., "Six-band HDTV camera system for spectrum-based color reproduction.", J. of Imaging Science and Technology, 48, 2, pp.85-92. (2004)
- [6] M. Tsuchida et al., "A stereo one-shot multi-band camera system for accurate color reproduction.", Proc. ACM Siggraph, Poster. (2010)
- [7] H. Takita et al., "High-accuracy image registration based on phase-only correlation.", IEICE Trans. of Fundamentals, Vol. E86-A, no.8, pp.1925-1934. (2003).
- [8] F. L. Bookstein, "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 11, No. 16, pp.567-585. (1989).

Author Biography

Masaru Tsuchida received the B.E., M.E and Ph.D. degrees from the Tokyo Institute of Technology, Tokyo, in 1999, 2002, respectively. In 2002, he joined NTT Laboratories, where his research areas included color science, three-dimensional image processing, and computer vision. His specialty is color measurement and multiband image processing. From 2003 to 2006, he worked at the National Institute of Information and Communication Technology (NICT) as a researcher for the "Natural Vision" project. Since 2011, he has been a guest professor at Ritsumeikan University, Kyoto.

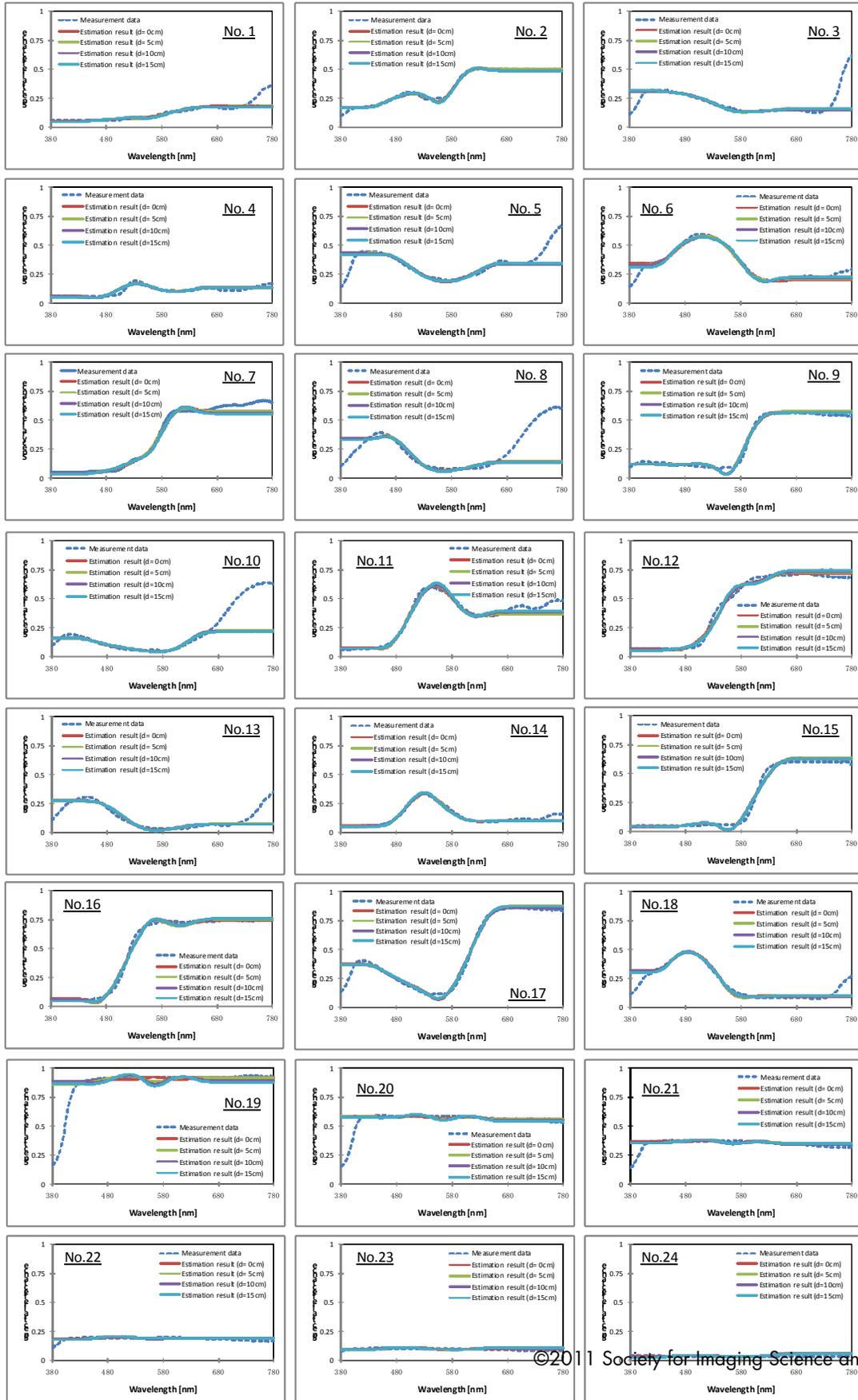
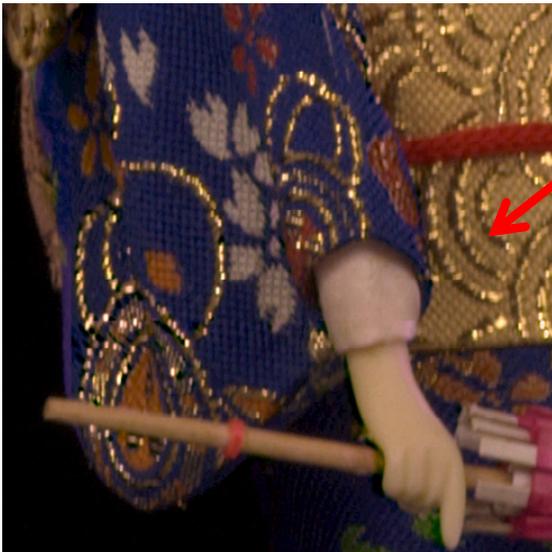
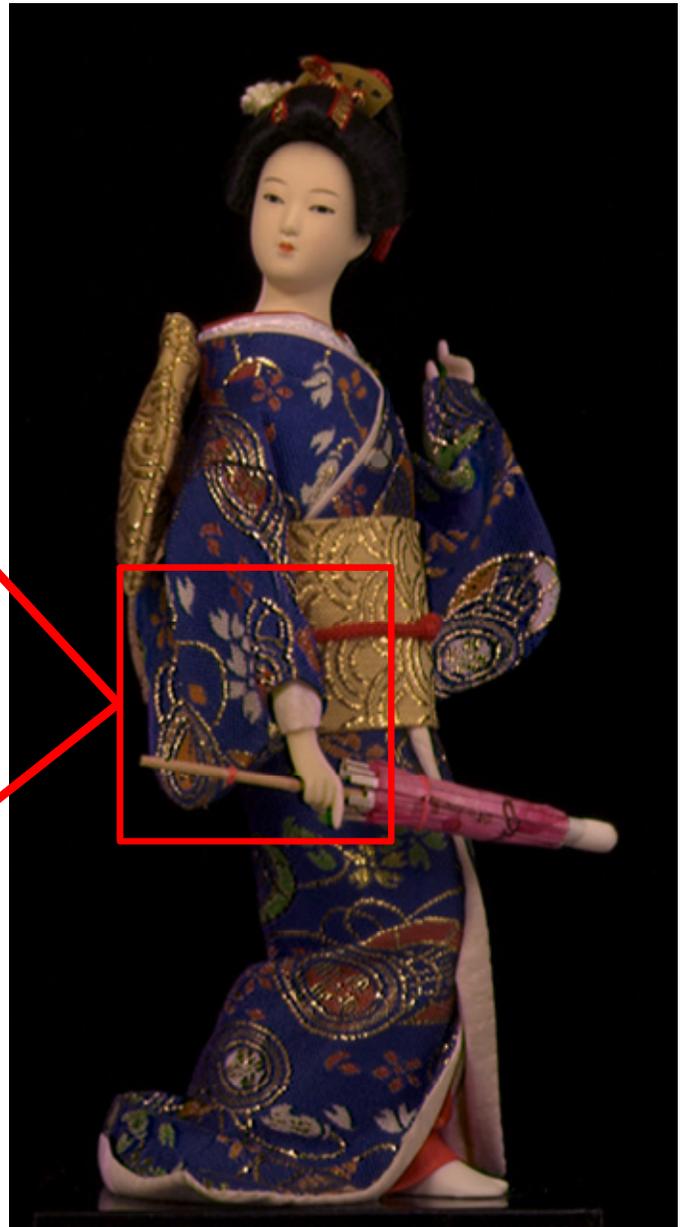


Fig.8: Estimated spectral reflectance.

Before image transformation



After image transformation



Color reproduction result

Fig.9: Color reproduction results of 3D object.