Hybrid-resolution multispectral imaging based on color filter array: basic principles and computer simulation

Yuri Murakami,[*] Masahiro Yamaguchi,[*] and Nagaaki Ohyama[**]; [*]Global Scientific Information and Computing Center, Tokyo Institute of Technology; Tokyo, Japan, [**]Imaging Science and Engineering Laboratory, Tokyo Institute of Technology; Yokohama, Japan

Abstract

Hybrid-resolution multispectral imaging has been proposed for reconstructing multispectral images from two types of data with different spatial and spectral resolutions. This paper proposes a new scheme for hybrid-resolution multispectral imaging based on image sensors with a multispectral filter array (MSFA). The proposed method reconstructs multispectral images from a high-resolution RGB image and image data from an MSFA with narrow-band color filters. The MSFA samples each band of a multispectral image sparsely in space, while RGB images are obtained at high spatial resolution; i.e., this dataset is regarded as a hybrid-resolution measurement. In the multispectral image reconstruction, each band of a multispectral image is reconstructed as a linear combination of high-resolution RGB images. The weighting coefficients of the linear combination are determined on the basis of MSFA data. By determining the weighting coefficients for each appropriate small area, overall multispectral images can be reconstructed with high accuracy. In this paper, typical configurations of image sensors for the hybrid-resolution multispectral imaging are also proposed. The effectiveness of the proposed method was verified for 16-band images through computer simulations in the case of a single-panel image sensor with an MSFA. As a result, it was confirmed that the proposed method reconstructed 16-band images at higher spatial resolution and with lower normalized root mean squared error than did a simple linear interpolation of the data from an equidensity-sampling MSFA.

Introduction

Multispectral imaging techniques are utilized as an effective tool to measure and analyze targets noninvasively. However, multispectral image acquisition requires bulky imaging systems or time-consuming processes that include data scanning along either the wavelength or spatial axis. As a result, the applications are limited to several typical fields, such as remote sensing and microscopic imaging. To utilize multispectral imaging techniques in wider applications, compact and one-shot spectral imaging instruments are strongly anticipated.

Image sensors based on color filter arrays (CFAs) are widely used to capture color images, and the interpolation techniques for each type of CFA are vigorously being investigated. Recently, multispectral image acquisition techniques using a multispectral filter array (MSFA) were proposed as an extension of CFA-based methods [1-2]. However, such an approach is inappropriate when the number of spectral bands is large (e.g., more than ten bands), because the spatial sampling becomes too sparse to reconstruct spectral images by interpolation-based methods.

A hybrid-resolution spectral imaging scheme has been proposed as a new type of spectral imaging technique [3-8]. In this scheme, spectral images are estimated from two types of data with different spatial and spectral resolutions: for instance, high-resolution RGB images and corresponding low-resolution spectral data. Various estimation methods have been proposed, mainly for the abovementioned combination of data [4-6,8]. Although the data obtained by an MSFA-type image sensor can be regarded as a kind of low-resolution spectral data, a hybrid-resolution approach has never been applied.

This paper proposes a new scheme for hybrid-resolution multispectral imaging based on an MSFA. The proposed method reconstructs multispectral images from a high-resolution RGB image and the data from an MSFA with narrow-band color filters. Because the previously proposed methods presuppose low-resolution spectral data measured at the same position for all bands, this paper presents a new reconstruction method that can be applied to MSFA data. By using an MSFA-type image sensor, compact configurations of multispectral cameras can be realized.

Below, a model of multispectral images is first introduced, and the proposed observation and reconstruction methods for multispectral images are derived on the basis of that model. Three typical configurations of image sensors for the proposed method are also presented. After that, the effectiveness of the proposed MSFA-based multispectral imaging is shown by computer simulations.

Hybrid-Resolution Multispectral Imaging based on Color Filter Array

Multispectral image modeling

A widely used approach is the representation of spectral reflectance functions by a linear combination of a relatively small number of basis functions. The required number of basis functions depends on the target data and the desired accuracy of the representation. If we consider the spectral reflectance functions involved by a relatively small area of a spectral image, the variation in those spectral reflectance functions is limited. As a result, it is expected that only a small number of basis functions should be needed to represent the spectral reflectance functions in that area.

Based on this assumption, we derive a model for multispectral images in relation to RGB images. In the model, we focus on each appropriately selected small area, which below is simply called a small area. The description below is related to a small area of a multispectral image, and N is the number of pixels in this area. In addition, the original spectral reflectance image is represented in a discrete form by sampling in both the spatial and spectral domains.

A spectral reflectance function $\mathbf{r}_i$ involved by a small area can be approximately represented by a linear combination of three spectral basis vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ as

$$\mathbf{r}_i \cong \sum_{i=1}^{3} b_{ij}\mathbf{u}_j \ , \tag{1}$$

where b_{ij} is a weighting coefficient, $\mathbf{r}_i$ and $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are L-dimensional column vectors, and L is the number of spectral samples; e.g., L = 81 in the case of sampling with a 5-nm interval for 380–780 nm. The generalization of Eq. (1) for the whole area can be written as

$$\mathbf{R} \cong \mathbf{BU} \ , \tag{2}$$

where

$$\mathbf{U} = \begin{pmatrix} \mathbf{u}_1{}^T \\ \mathbf{u}_2{}^T \\ \mathbf{u}_3{}^T \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} \mathbf{r}_1{}^T \\ \vdots \\ \mathbf{r}_N{}^T \end{pmatrix}, \tag{3}$$

and $\mathbf{B}$ is the $N \times 3$ matrix that consists of b_{ij}.

Let us consider K-band multispectral data

$$\mathbf{f}_i = \mathbf{H}_{MS}{}^T \mathbf{r}_i \ , \tag{4}$$

where $\mathbf{H}_{MS}$ is an $L \times K$ matrix that represents the spectral sensitivity of the multispectral imaging system, and the spatial resolution of the multispectral image is assumed to be the same as that of the original spectral reflectance image. A multispectral image corresponding to the same small area is represented by

$$\mathbf{F} = \mathbf{RH}_{MS} \ , \tag{5}$$

where

$$\mathbf{F} = \begin{pmatrix} \mathbf{f}_1{}^T \\ \vdots \\ \mathbf{f}_N{}^T \end{pmatrix}. \tag{6}$$

Substituting Eq. (2) into Eq. (5), we have

$$\mathbf{F} \cong \mathbf{BUH}_{MS} \ . \tag{7}$$

If an RGB camera system can be modeled by a linear system, then $\mathbf{G} = \begin{pmatrix} \mathbf{g}_R & \mathbf{g}_G & \mathbf{g}_B \end{pmatrix}$, the RGB image of $\mathbf{R}$, is represented by

$$\mathbf{G} = \mathbf{RH}_{RGB} \ , \tag{8}$$

where $\mathbf{H}_{RGB}$ is an $L \times 3$ matrix representing the spectral characteristics of the camera system. Substituting Eq. (2) into Eq. (8),

$$\mathbf{G} \cong \mathbf{BUH}_{RGB} \ . \tag{9}$$

Except for cases in which the 3×3 matrix $\mathbf{UH}_{RGB}$ is singular,

$$\mathbf{B} \cong \mathbf{G}\left(\mathbf{UH}_{RGB}\right)^{-1} . \tag{10}$$

Substituting Eq. (10) into Eq. (7), we have

$$\begin{aligned} \mathbf{F} &\cong \mathbf{G}\left(\mathbf{UH}_{RGB}\right)^{-1}\mathbf{UH}_{MS} \ , \\ &= \mathbf{GA} \end{aligned} \tag{11}$$

where $\mathbf{A} = \left(\mathbf{UH}_{RGB}\right)^{-1}\mathbf{UH}_{MS}$. Equation (11) can be rewritten as

$$\widetilde{\mathbf{f}}_k \cong \mathbf{Ga}_k \ , \tag{12}$$

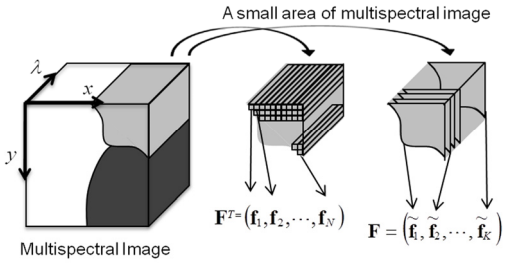

Figure 1. *Two types of notation for a small area of a multispectral image: collection of spectral data (left) and collection of image bands (right).*

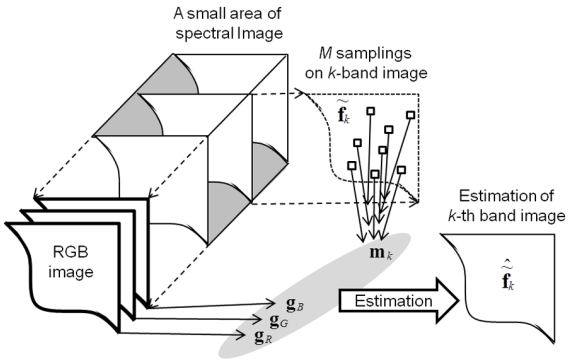

Figure 2.*Two types of observational data required to reconstruct a small area of a multispectral image: RGB image for corresponding area and a small number of samples from each image band.*

where $\widetilde{\mathbf{f}}_k$ and $\mathbf{a}_k$ are the kth column vectors of matrices $\mathbf{F}$ and **A,** respectively. The N-dimensional column vector $\widetilde{\mathbf{f}}_k$ is the kth-band image of the small area. The relation between $\mathbf{f}_k$ and $\widetilde{\mathbf{f}}_k$ is depicted in Fig. 1. Equation (12) indicates that any image band of a multispectral image is approximately represented by a linear combination of the corresponding areas of the RGB image bands, $\mathbf{g}_R, \mathbf{g}_G, \mathbf{g}_B$.

Considering that $\mathbf{g}_R, \mathbf{g}_G, \mathbf{g}_B$ serve as a spatial basis for each one-band image of this area, we can extend the set of spatial basis vectors by adding some versatile spatial basis vector to $\mathbf{G}$. The simplest extension is $\mathbf{G} \to \begin{bmatrix} \mathbf{G} & \mathbf{1}_N \end{bmatrix}$, where $\mathbf{1}_N$ is the N-dimensional column vector whose elements all are unity. With this extension, $\mathbf{a}_k$ becomes a 4-dimensional column vector.

Multispectral image observation and reconstruction methods

The model of Eq. (12) indicates that any image band of a multispectral image is represented by the corresponding RGB image $\mathbf{G}$ and three weighting coefficients $\mathbf{a}_k$. Therefore, if we have $\mathbf{G}$, all we need is the information required for estimating $\mathbf{a}_k$. Let us consider that we obtain M independent linear observations of $\widetilde{\mathbf{f}}_k$,

$$\begin{aligned} \mathbf{m}_k &= \mathbf{S}_k\widetilde{\mathbf{f}}_k \\ &\cong \left(\mathbf{S}_k\mathbf{G}\right)\mathbf{a}_k \end{aligned} \ , \tag{13}$$

where $\mathbf{S}_k$ is the $M \times N$ system matrix of the observations. If $M \geq 3$, the three unknowns of $\mathbf{a}_k$ can be estimated by any regression-based method; e.g., by simple multiple linear regression analysis we have an estimation of $\mathbf{a}_k$ as

$$\hat{\mathbf{a}}_k = \left[(\mathbf{S}_k\mathbf{G})^T(\mathbf{S}_k\mathbf{G})\right]^{-1}(\mathbf{S}_k\mathbf{G})^T\mathbf{m}_k . \quad (14)$$

Then, an estimation of $\tilde{\mathbf{f}}_k$ is derived as

$$\hat{\tilde{\mathbf{f}}}_k = \mathbf{G}\hat{\mathbf{a}}_k . \quad (15)$$

Note that calculation of Eqs. (14) and (15) is done for each small area.

A simple implementation of M independent observations is M-pixel sampling of $\tilde{\mathbf{f}}_k$. Below, only this implementation is focused upon. Figure 2 is a conceptual diagram of the proposed observation and reconstruction of multispectral images. As shown in Fig. 2, each image band can be reconstructed independently, which means that the sampling position for $\mathbf{m}_k$ can be different for each band. Therefore, an MSFA-type sensor is utilized to obtain $\mathbf{m}_k$. In the next subsection, typical configurations of image sensors are proposed to obtain both high-resolution RGB images and MSFA data.

To end this subsection, let us discuss how to determine the small area. The smaller a region, the higher the probability that it involves only spectra with similar features. As a result, the condition in Eq. (1) is easily satisfied. However, as the region becomes smaller (i.e., as M becomes smaller), the regression to derive Eq. (14) becomes unstable. Therefore, these opposite influences should be considered in deciding the size of the regions. Regardless of the size, there are several ways to divide images into regions. A simple method is that an image be divided horizontally and vertically into small blocks. Because neighboring pixels in multispectral images are more likely to have similar spectral characteristics, simple block division can be effective. Another method is that pixels with similar RGB values be collected into a single region.

Configurations of image sensors with multispectral filter array

There are several implementation methods to realize the observations discussed in the previous section. Here, three typical configurations of image sensors are proposed, taking the case of $K = 16$ as an example (Fig. 3).

The first configuration consists of four image sensors (top of Fig. 3). Three of these are used to capture an RGB image, while the remaining one is an MSFA-type image sensor. The spectral sensitivity of the R, G, and B channels can be implemented by absorption color filters or a dichroic prism. The MSFA consists of 16 narrow-band color filters with different peak wavelengths, which are placed with equal density and in some arbitrary arrangement.

The second configuration consists of two image sensors (middle of Fig. 3). One of these is an RGB CFA-type image sensor, and the other is an MSFA-type image sensor. By combining the R, G, and B image sensors into one, the number of image sensors is reduced, which realizes a simple optical design and a compact imaging system.

The third configuration is a single-panel image sensor with an MSFA (bottom of Fig.3). The MSFA consists of 16 narrow-band color filters with different peak wavelengths as well as R, G, and B color filters. An element pattern of the MSFA consists of $4K = 64$ pixels, which is equivalent to an 8×8 pixel area, for instance. In the element pattern, each of the RGB color filters is allocated to every other pixel along the horizontal and vertical directions. As a result, $3K = 48$ pixels are filled by the RGB color filters. The 16 narrow-band color filters are arbitrarily arranged in the remaining $K = 16$ pixels.

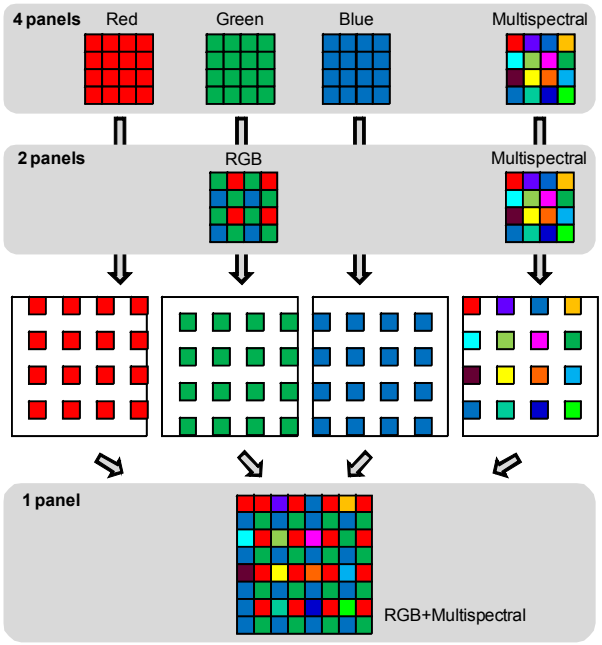

Figure 3. *Three typical configurations of image sensors with a multispectral filter array for the proposed hybrid-resolution multispectral imaging in the case of K=16.*

In the case of four image sensors, the reconstruction method of Eqs. (14) and (15) can be applied directly to the measurement data. On the other hand, in the case of two image sensors or one, every RGB image is first interpolated to have the same resolution as the image sensor. Then, multispectral images are reconstructed on the basis of the interpolated RGB images by means of Eqs. (14) and (15). For the interpolation of the RGB images, any method can be used. However, the interpolation error directly affects the image quality of the reconstructed multispectral images. Therefore, this process should be executed carefully. A typical error is that caused by aliasing distortion. One realistic and effective approach to this problem is inserting some optical elements in front of the image sensor to cut the high-frequency components of the image so that the R, G, or B sampling data can be made to satisfy the sampling theorem for the low-passed version of the image. As a result, the low-passed image can be reconstructed accurately from the sampled data of any of the R, G, and B channels. Note that when using this approach, the resolution of the multispectral image obtained is approximately half that of the image sensor.

Simulation Results

Three spectral reflectance images were prepared as original scenes, which are shown in color in Fig. 4. Each image consists of 512×512 pixels and 81 wavelength samples in the range of 380 to 780 nm with an interval of 5 nm. These spectral reflectance images were obtained through Wiener estimation from a multispectral image captured by a 16-band camera [9].

An imaging system of single-panel type is assumed; i.e., the MSFA consists of 16 narrow-band color filters as well as R, G, and B color filters. The spectral sensitivity functions of these color filters are shown in Fig. 5. An element pattern of the MSFA is shown in the left-hand side of Fig. 6. The total pixel number of the image sensor with the MSFA is 512×512,

Figure 4.*Three original spectral reflectance images presented in color: toys (top-left), scarves (top-right), and flowers (bottom).*

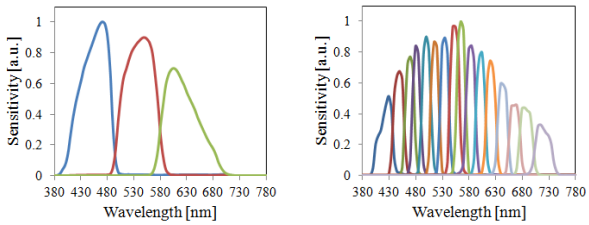

Figure 5. Spectral sensitivities of R, G, and B color filters (left) and 16 narrow-band color filters (right) used in simulation.

9	R	15	R	5	R	3	R
B	G	B	G	B	G	B	G
8	R	2	R	10	R	13	R
B	G	B	G	B	G	B	G
11	R	14	R	7	R	1	R
B	G	B	G	B	G	B	G
6	R	4	R	12	R	16	R
B	G	B	G	B	G	B	G

9	15	5	3	9	15	5	3
8	2	10	13	8	2	10	13
11	14	7	1	11	14	7	1
6	4	12	16	6	4	12	16
9	15	5	3	9	15	5	3
8	2	10	13	8	2	10	13
11	14	7	1	11	14	7	1
6	4	12	16	6	4	12	16

Figure 6. *Color filter arrangements (element patterns) used in simulation for proposed MSFA with hybrid-resolution sampling (left) and equidensity-sampling MSFA (right). The 16 narrow-band color filters are numbered by peak wavelength from the shortest to the longest.*

and the element pattern of color filters is repeated both horizontally and vertically over the whole image sensor.

The simulation flow is shown in Fig. 7. The simulation was started from a 512 × 512 spectral reflectance image. It was assumed that an optical low-pass filter was inserted in front of the image sensor to cut approximately the upper half of the spatial frequency of the original image. This avoids aliasing distortion of R, G, and B images when the 512 × 512 image is reconstructed from the sampled data. The measurement of the 512 × 512 data by the MSFA was calculated from the low-passed version of the spectral image. We simulated both cases with and without noise, where the noise was assumed to be additive white Gaussian noise at a peak signal-to-noise ratio (PSNR) of 50 dB or 40 dB. Note that the PSNRs of the 16 narrow bands were approximately 10dB less than the setting.

From the measurements by the MSFA, 512 × 512 16-band images were reconstructed. Before the image reconstruction, the RGB images were interpolated by zero padding in the frequency domain. Then, each band of the 512 × 512 image was divided into 16 × 16 pixel blocks, which corresponded to small areas, and the image data for each block were reconstructed by means of Eqs. (14) and (15).

For comparison, an MSFA with equidensity sampling was also tested. The element pattern of the MSFA is shown in the right-hand side of Fig. 6. By setting 16 narrow-band color filters of equal density, each color filter was placed in every 4 × 4 pixel region. Note that in this case the R, G, and B color filters were not used. For these data, linear interpolation was applied to each image band independently to reconstruct the full-resolution images.

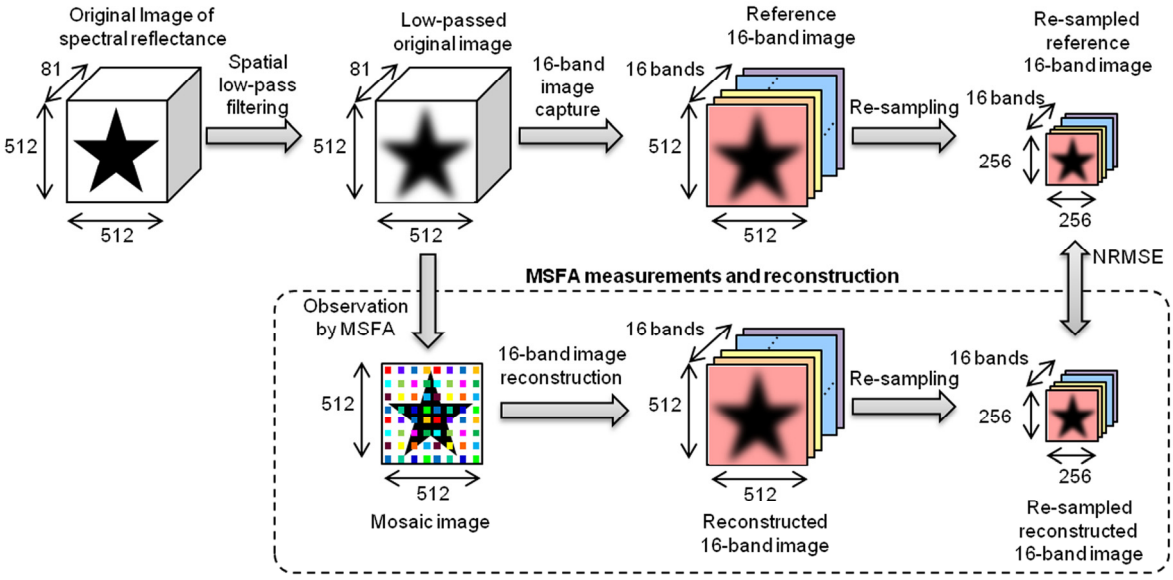

Figure 7. *Schematic diagram of simulation flow.*

For evaluation purposes, 16-band images were calculated from the low-passed original images, which were used for the reference data. Considering that the valid resolution is half the original, the reference and reconstructed 16-band images were resized to 256×256 pixel images, and the normalized root mean squared error (NRMSE) between these was calculated for evaluation. In the resizing, the images were sampled at the positions of the multispectral-filter pixels.

Figure 8 shows the NRMSEs of the proposed method (RGB+MS) and the linear interpolation of equidensity-sampling data (16MS) in the noise-free case. For all three images, the proposed method generally realizes lower NRMSE than 16MS. Only for the 16th image band is the error of the proposed method higher. When the multispectral image was reconstructed from the extended version of the spatial basis, $[\mathbf{G} \quad \mathbf{1}_N]$, the NRMSEs of the 16th image bands were reduced while the NRMSEs of other image bands were not changed significantly. The NRMSEs of the 16th image bands for the toys, scarves, and flowers became 6.8%, 7.3%, and 6.4% respectively. This result indicates that the extension of the spatial basis can reduce the error. Nonetheless, the error of the 16th image band is higher than others. One reason is presumably that the spectral transmittance of the 16th image band is not overlapped with any of the RGB color filters.

Figure 9 shows the NRMSEs for the toys in the cases with an without noise. As the noise level is increased, the error increases, especially for the proposed method. However, we can confirm that the proposed method, RGB+MS, realizes lower error than 16MS even for noise cases. The results for the images of the scarves and flowers exhibit similar behavior.

Figure 10 shows a 64×64 pixel region of a 256×256 16-band image of the scarves in the noise-free case. The top

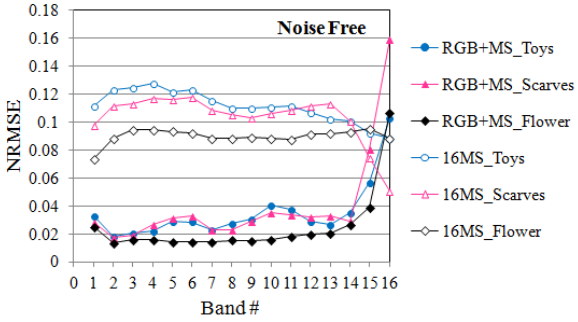

Figure 8. *Normalized root mean squared errors of 16-band images reconstructed by proposed method (RGB+MS) and linear interpolation of equally sampled data (16MS) in the noise-free case.*

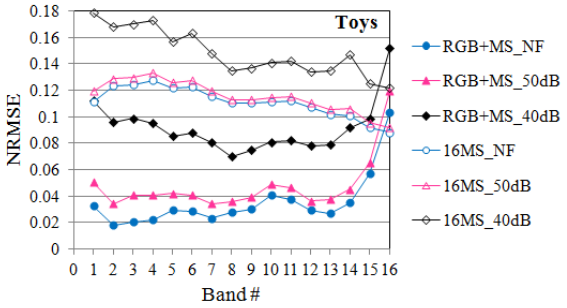

Figure 9. *Normalized root mean squared errors of 16-band images reconstructed by proposed method (RGB+MS) and linear interpolation of equally sampled data (16MS) for the image of toys in the noise-free case (NF) and with-noise cases (40 dB and 50 dB).*

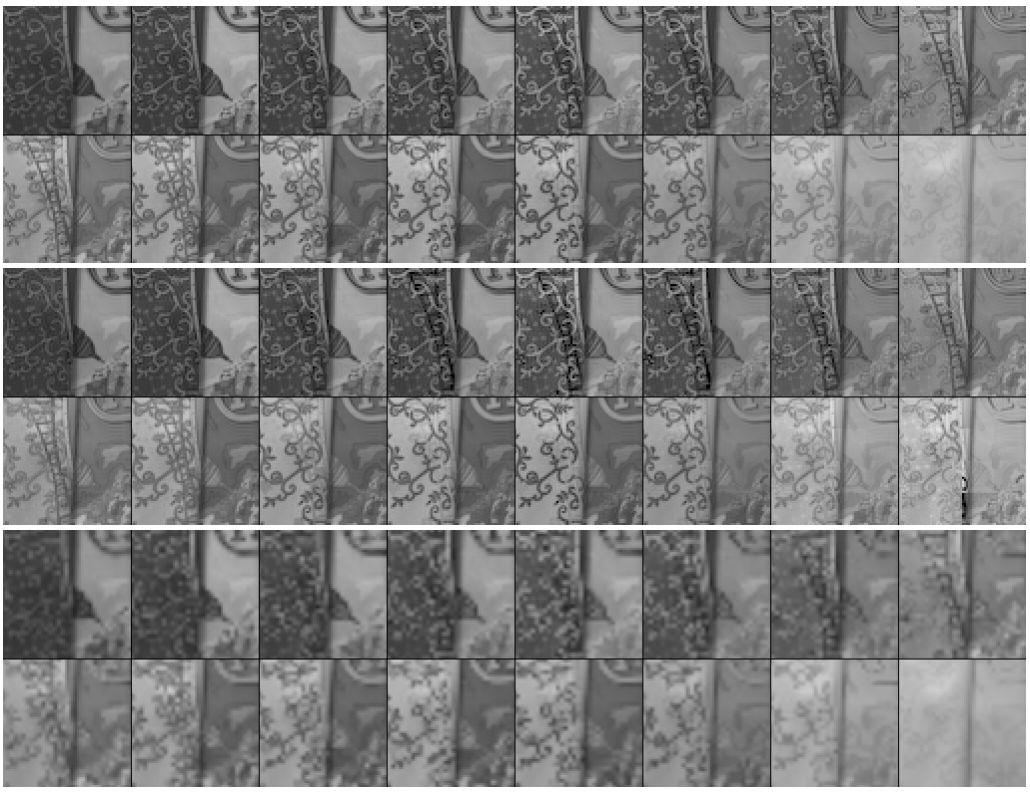

Figure 10. *Reference (top), reconstruction by proposed method (middle), and reconstruction by linear interpolation of equidensity samples (bottom) for 64 × 64 pixel region of 16-band image of scarves in noise-free case. In each panel, the upper row corresponds to image bands 1–8 and the lower row corresponds to image bands 9–16.*

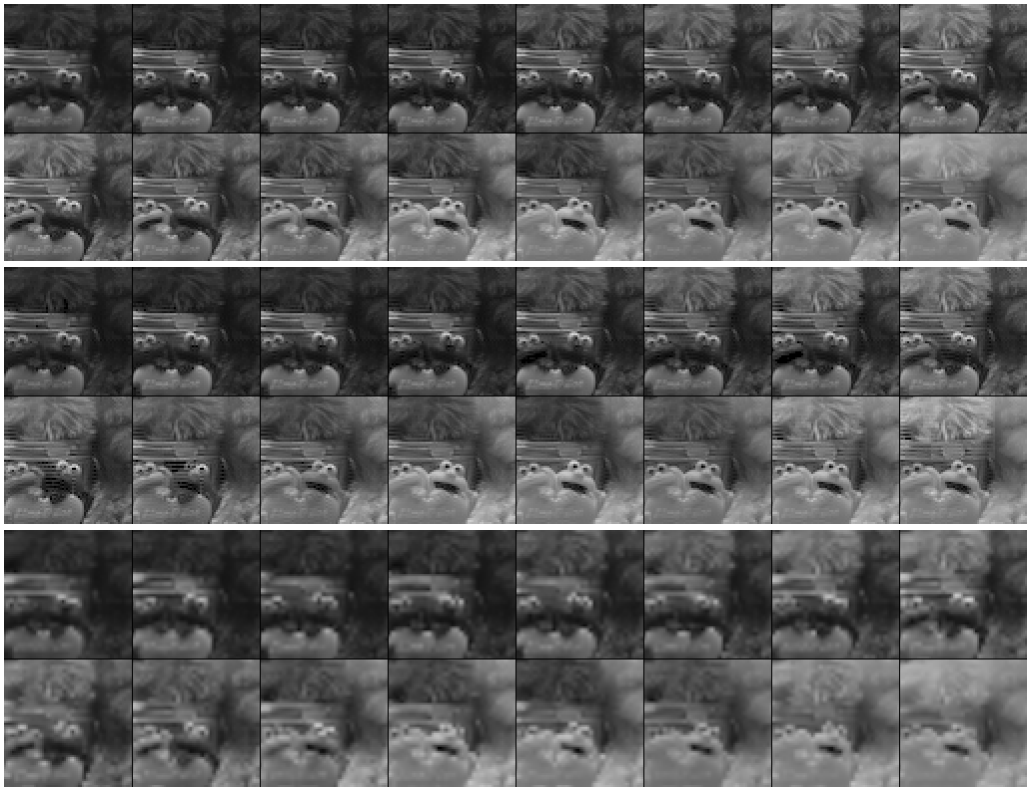

Figure 11. *Reference (top), reconstruction by proposed method (middle), and reconstruction by linear interpolation of equidensity samples (bottom) for 64 × 64 pixel region of 16-band image of toys in noise-free case. In each panel, the upper row corresponds to image bands 1–8 and the lower row corresponds to image bands 9–16.*

panel is the reference, the middle panel is the reconstruction by the proposed method, and the bottom panel is the reconstruction by linear interpolation of the equidensity samples. For each of the 16-band images, the upper row is the 1st through 8th one-band images from left to right, and the lower row is the 9th through 16th. On the whole, the spatial resolution of the image obtained is significantly higher by the proposed method than by 16MS. The 16×16 small areas, which are independently processed, are not visually recognized in most one-band images. In the 4th through 6th one-band images by the proposed method, some error is recognized (in a ladder-like structure). In addition, we can see large error in the 16th-band image, as structures that are not in the reference appear in the reconstruction, which causes the large NRMSE of the 16th image band in Fig. 8.

Figure 11 shows the results for the toys in the manner of Fig. 10. The overall tendency is similar to that in Fig. 10. However, we can see that the aliasing error caused by the RGB sampling is enlarged in the reconstruction of the 8th through 10th, 15th, and 16th one-band images. Except for the enlarged aliasing error in some areas, the overall image quality is good.

Conclusions and Future Works

This paper proposes a new scheme for hybrid-resolution multispectral imaging based on an MSFA. The observations for the proposed method can be realized by one, two, or four image sensors with an MSFA. Therefore, there is a possibility to realize a compact one-shot multispectral imaging instrument. Through computer simulations for 16-band images, it was confirmed that a multispectral image is obtained with significantly high spatial frequency, considering the low sampling ratio of each image band. The NRMSE is less than 4% for most image bands in both noise-free and 50-dB-noise cases. However, some distinctive error was observed that did not appear in the image derived by simple methods based on linear interpolation. These errors can be reduced by optimizing or improving each of the following steps of the proposed method:

- interpolation of the RGB image,
- extension of the spatial basis,
- selection of small areas,
- regression to derive the weighting coefficients, and
- selection of the spectral sensitivity of the RGB color filters.

These improvements will be made in future work.

This research is supported by KAKENHI (20-40108) and KAKENHI (23135509).

References

[1] R. Shrestha, J. Y. Hardeberg, and R. Khan, "Spatial arrangement of color filter array for multispectral image acquisition," Proc. SPIE 7875, 787502 (2011).

[2] J. Brauers and T. Aach, "A color filter array based multispectral camera," the 12. Workshop Farbbildverarbeitung (2006).

[3] F. H. Imai and R. S. Berns, "High-resolution multi-spectral image archives: A hybrid approach, Proc. CIC, 224 (1998).

[4] Y. Murakami, K. Ietomi, M. Yamaguchi, and N. Ohyama, "MAP estimation of spectral reflectance from color image and multipoint spectral measurements," Appl. Opt., 46, 7068 (2007).

[5] Y. Murakami, K. Ietomi, A. Tadano, M. Yamaguchi, and N. Ohyama, "Comparison of spectral image reconstruction methods using multipoint spectral measurements," Proc. CGIV, 591 (2008).

[6] Y. Murakami, M. Yamaguchi, and N. Ohyama, "Piecewise Wiener estimation for reconstruction of spectral reflectance image by multipoint spectral measurements," Appl. Opt., 48, 2188 (2009).
[7] O. Kohonen, "Multiresolution-based pansharpening in spectral color images," Proc. CGIV, 535 (2010).
[8] Y. Murakami, M. Yamaguchi, and N. Ohyama, "Class-based spectral reconstruction based on unmixing of low-resolution spectral information," JOSA A, 28, 1470 (2011).
[9] H. Fukuda, T. Uchiyama, H. Haneishi, M. Yamaguchi, and N. Ohyama, "Development of 16-band multispectral image archiving system," Proc. SPIE 5667, 136 (2002).

Author Biography

Yuri Murakami is a researcher at the Global Scientific Information and Computing Center at the Tokyo Institute of Technology (Tokyo Tech) in Japan. She received her M.S. and Ph.D. degrees from Tokyo Tech in 1998 and 2005. From 2000 to 2005 she was an assistant professor at Tokyo Tech, and from 2008 to 2011 she was a research fellow of the Japan Society for the Promotion of Science. Her research interests include multispectral imaging and color image reproduction.