Nonlinear estimation of spectral reflectance based on Gaussian mixture distribution for color image reproduction

Yuri Murakami, Takashi Obi, Masahiro Yamaguchi, and Nagaaki Ohyama

Nonlinear estimation method of spectral reflectance from camera responses is proposed. The proposed method minimizes the mean square error of spectral reflectance when the reflectance can be regarded as a random sequence of Gaussian mixture distribution. In computer simulations, 168 samples of spectral reflectance from a color chart are estimated from their image signals obtained by three- and six-band cameras. It is confirmed that the proposed method improves the accuracy in comparison with the conventional Wiener estimation method. © 2002 Optical Society of America

OCIS codes: 100.2000, 100.3190, 300.6550, 330.1690.

1. Introduction

High fidelity color image reproduction through digital imaging devices has become important recently, especially in applications, such as telemedicine, digital art, on-line shopping, etc. The color of an object depends on the viewing illumination, and in order to reproduce the color of the object under various illuminations, information about the spectral reflectance, which can be obtained by multispectral imaging,[1–3] is required. It is reported that a large number of color bands, approximately more than fourteen,[3] is required for accurate reproduction of various objects. However, improvement in the accuracy with cameras that operate in fewer bands or ordinary trichromatic camera is also an important issue[4–6] in some practical applications. This paper presents a nonlinear estimation method of spectral reflectance, which improves the estimation accuracy especially when the number of the color bands is insufficient.

Y. Murakami and N. Ohyama are with the Tokyo Institute of Technology, Frontier Collaborative Research Center and the Telecommunications Advancement Organization of Japan, Akasaka Natural Vision Research Center. T. Obi and M. Yamaguchi are with the Tokyo Institute of Technology, Imaging Science & Engineering Laboratory and the Telecommunications Advancement Organization of Japan, Akasaka Natural Vision Research Center. Y. Murakami's e-mail address is yuri@isl.titech.ac.jp.

Received 15 January 2002; revised manuscript received 15 April 2002.

0003-6935/02/234840-08$15.00/0

The formation process of a pixel of N-band images ($N = 3$ in case of trichromatic) can often be modeled by a linear system, where the input is a spectral reflectance and the output is a N-dimensional image signal. Then, the estimation of a spectral reflectance from an image signal can be regarded as the inverse problem of a linear system. To improve estimation accuracy, the probability density of input data is often used as *a priori* information. One such method is the Wiener estimation that minimizes the mean square error of the estimate when the input is a Gaussian sequence. Even if it is not a Gaussian sequence, the Wiener estimation gives the best estimate among the linear estimates, therefore it has been widely used. However, because there are various colors and various subjects that exist in an image, it is difficult to model the probability density of the input by a Gaussian distribution. Therefore, if the mean square error based on the actual probability density can be minimized, the estimation accuracy should be improved. For this purpose, a nonlinear estimation technique is needed.

In this paper, we propose an estimation method called the Gaussian mixture distribution based (GMD-based) estimation, which minimizes the mean square error of estimates when the input is a random sequence of GMD. The proposed method is applied to the estimation problem of the spectral reflectance of various colors in the color chart from the camera responses. The results show the improvement of the estimation accuracy.

2. Theory

A. Gaussian Mixture Distribution

If the probability density distribution of a stochastic sequence $\mathbf{f}$ is a mixture distribution, it can be written as

$$P^M(\mathbf{f}) = \sum_{k=1}^{K} w_k p_k(\mathbf{f}), \tag{1}$$

where the superscript M on $P^M(\mathbf{f})$ indicates a mixture distribution, K is the number of the components, $p_k(\mathbf{f})$ is kth component density, and w_k is kth weight coefficient. The functions, $p_k(\mathbf{f})$ and w_k, satisfy

$$\int p_k(\mathbf{f})\mathrm{d}\mathbf{f} = 1, \tag{2}$$

$$\sum_{k=1}^{K} w_k = 1, \qquad 0 \le w_k \le 1, \tag{3}$$

respectively, which yield the integration of $P^M(\mathbf{f})$ becomes 1. When $p_k(\mathbf{f})$ is a Gaussian function, $P^M(\mathbf{f})$ is called a GMD.

B. Linear Inverse Problem from a Stochastic View

Let $\mathbf{f}$ and $\mathbf{g}$ be a zero mean stochastic sequence and its linear observation, where their relationship is represented with matrix $\mathbf{H}$ by

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \tag{4}$$

where $\mathbf{n}$ is a vector of an additive noise. In the case when an input sequence is not zero mean, the subtraction of the mean from the original should be applied in the same way. Then, an estimate of $\mathbf{f}$, $\hat{\mathbf{f}}$, from $\mathbf{g}$ is derived under the condition that the mean square error

$$\varepsilon = \langle \|\mathbf{f} - \hat{\mathbf{f}}\|^2 \rangle \tag{5}$$

is minimized, where $\langle\ \rangle$ is an averaging operator. It is known that $\hat{\mathbf{f}}$ can be calculated as the conditional mean of $\mathbf{f}$ given $\mathbf{g}$, that is,

$$\begin{aligned}\hat{\mathbf{f}} &= \langle \mathbf{f}|\mathbf{g} \rangle \\ &= \int P(\mathbf{f}|\mathbf{g})\mathbf{f}\mathrm{d}\mathbf{f},\end{aligned} \tag{6}$$

where $P(\mathbf{f}|\mathbf{g})$ is the conditional probability density of $\mathbf{f}$ given $\mathbf{g}$.

When $P(\mathbf{f}|\mathbf{g})$ is Gaussian, Eq. (6) can be solved linearly. One of such cases is that the probability densities of $\mathbf{f}$ and $\mathbf{n}$ are Gaussian, and they are independent. In this case $P(\mathbf{f}|\mathbf{g})$ can be written as

$$P(\mathbf{f}|\mathbf{g}) \propto P^G(\mathbf{f})\,P^G(\mathbf{n}), \tag{7}$$

where superscript G indicates a Gaussian distribution,

$$P^G(\mathbf{f}) = C_f \exp\left(-\frac{1}{2}\,\mathbf{f}^T\Sigma_f^{-1}\mathbf{f} \right), \tag{8}$$

$$P^G(\mathbf{n}) = C_n \exp\left(-\frac{1}{2}\,\mathbf{n}^T\Sigma_n^{-1}\mathbf{n} \right), \tag{9}$$

where C_f and C_n are normalization constants. From Eqs. (7)–(9), we have

$$P(\mathbf{f}|\mathbf{g}) = P^G(\mathbf{f}|\mathbf{g}) \propto \exp\left[-\frac{1}{2}\,(\mathbf{f} - \mathbf{f}^*)^T\Sigma^{*-1}(\mathbf{f} - \mathbf{f}^*) \right], \tag{10}$$

where the mean and the covariance are

$$\mathbf{f}^* = \Sigma_f\mathbf{H}^T(\mathbf{H}\Sigma_f\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{g}, \tag{11}$$

$$\Sigma^* = [\mathbf{I} - \Sigma_f\mathbf{H}^T(\mathbf{H}\Sigma_f\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{H}]\Sigma_f. \tag{12}$$

The detail derivation is demonstrated in Ref. 7. Substituting Eq. (10) into Eq. (6) gives

$$\hat{\mathbf{f}} = \mathbf{f}^*. \tag{13}$$

The vector, $\mathbf{f}^*$, is the best estimate minimizing the mean square error in this Gaussian case, and is identical to the Wiener estimate.

C. Mixture Distribution Case

Assuming that the probability density of $\mathbf{f}$ is the mixture distribution $P^M(\mathbf{f})$, the conditional probability density of $\mathbf{f}$ given $\mathbf{g}$ becomes

$$P(\mathbf{f}|\mathbf{g}) = AP^M(\mathbf{f})\,P^G(\mathbf{n}) \tag{14}$$

as Eq. (9), where A is a normalization constant such that

$$\int P(\mathbf{f}|\mathbf{g})\mathrm{d}\mathbf{f} = 1. \tag{15}$$

Substituting Eq. (1) into Eq. (14) gives

$$P(\mathbf{f}|\mathbf{g}) = A\sum_{k=1}^{K} \bar{p}_k(\mathbf{f}|\mathbf{g}), \tag{16}$$

where we use the notation $\bar{p}_k(\mathbf{f}|\mathbf{g})$ defined by

$$\bar{p}_k(\mathbf{f}|\mathbf{g}) \equiv w_k p_k(\mathbf{f})\,P(\mathbf{n}). \tag{17}$$

The function, $\bar{p}_k(\mathbf{f}|\mathbf{g})$, is not identical, but in proportion to the probability density of $\mathbf{f}$ given $\mathbf{g}$ when $\mathbf{f}$ is a random sequence of $p_k(\mathbf{f})$, $p_k(\mathbf{f}|\mathbf{g})$;

$$\bar{p}_k(\mathbf{f}|\mathbf{g}) = B_k p_k(\mathbf{f}|\mathbf{g}), \tag{18}$$

where

$$B_k \equiv \int \bar{p}_k(\mathbf{f}|\mathbf{g})\mathrm{d}\mathbf{f}. \tag{19}$$

Substituting Eqs. (16) and (18) into Eq. (6), the best estimate becomes

$$\hat{\mathbf{f}} = \int P(\mathbf{f}|\mathbf{g})\mathbf{f}d\mathbf{f} = A\sum_{k=1}^{K}\left[B_k\int p_k(\mathbf{f}|\mathbf{g})\mathbf{f}d\mathbf{f}\right]. \quad (20)$$

The vector, $\int p_k(\mathbf{f}|\mathbf{g})\mathbf{f}d\mathbf{f}$, is the best estimate of $\mathbf{f}$ if $\mathbf{f}$ is a random sequence of $p_k(\mathbf{f})$, which is denoted by $\hat{\mathbf{f}}_k$. As a result, we have

$$\hat{\mathbf{f}} = A\sum_{k=1}^{K} B_k\hat{\mathbf{f}}_k. \quad (21)$$

This equation indicates that $\hat{\mathbf{f}}$ is the sum of $\hat{\mathbf{f}}_k$ with the weight B_k.

D. Gaussian Mixture Distribution-Based Estimation

Next, let us think about the case where the probability density of $\mathbf{f}$ is a GMD, every $p_k(\mathbf{f})$ is Gaussian;

$$p_k(\mathbf{f}) = {p_k}^G(\mathbf{f}) = C_k\exp\left[-\frac{1}{2}(\mathbf{f}-\langle\mathbf{f}_k\rangle)^T{\Sigma_k}^{-1}(\mathbf{f}-\langle\mathbf{f}_k\rangle)\right], \quad (22)$$

where C_k is the normalization constant as A is in Eq. (16), and $\langle\mathbf{f}_k\rangle$ and Σ_k are mean and covariance of random sequence $\mathbf{f}$ of ${p_k}^G(\mathbf{f})$. In addition, assuming the Gaussian noise, $p_k(\mathbf{f}|\mathbf{g})$ also becomes Gaussian, the same as Eqs. (10)–(12);

$$p_k(\mathbf{f}|\mathbf{g}) = {p_k}^G(\mathbf{f}|\mathbf{g}) \propto \exp\left[-\frac{1}{2}(\mathbf{f}-\mathbf{f}_k^*)^T{\Sigma_k^*}^{-1}(\mathbf{f}-\mathbf{f}_k^*)\right], \quad (23)$$

where

$$\mathbf{f}_k^* = \langle\mathbf{f}_k\rangle + \Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T+\Sigma_n)^{-1}(\mathbf{g}-\mathbf{H}\langle\mathbf{f}_k\rangle) \quad (24)$$

and

$$\Sigma_k^* = [\mathbf{I}-\Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T+\Sigma_n)^{-1}\mathbf{H}]\Sigma_k, \quad (25)$$

where we assume that the mean vector $\langle\mathbf{f}_k\rangle$ is not zero. As Eq. (13), $\hat{\mathbf{f}}_k$ can be obtained by $\mathbf{f}_k^*$, which is identical to the Wiener estimate $\hat{\mathbf{f}}_k^{\text{Wiener}}$ using *a priori* information of ${p_k}^G(\mathbf{f})$.

To calculate B_k from Eq. (19), we need the complete form of $\bar{p}_k(\mathbf{f}|\mathbf{g})$ including the proportional constant. In Appendix A, we show the derivation. For now, we simply state the result as

$$\bar{p}_k(\mathbf{f}|\mathbf{g}) = \bar{p}_k{}^G(\mathbf{f}|\mathbf{g}) = w_kC_kC_n\exp\left\{-\frac{1}{2}D_k\right\}\times\exp\left[-\frac{1}{2}(\mathbf{f}-\mathbf{f}_k^*)^T{\Sigma_k^*}^{-1}(\mathbf{f}-\mathbf{f}_k^*)\right], \quad (26)$$

where

$$D_k = (\mathbf{g}-\mathbf{H}\langle\mathbf{f}_k\rangle)^T(\mathbf{H}\Sigma_k\mathbf{H}^T+\Sigma_n)^{-1}(\mathbf{g}-\mathbf{H}\langle\mathbf{f}_k\rangle). \quad (27)$$

From Eq. (26), we can calculate B_k as

$$\begin{aligned}B_k &= \int\left(w_kC_kC'\exp\left\{-\frac{1}{2}D_k\right\}\times\exp\left[-\frac{1}{2}(\mathbf{f}-\mathbf{f}_k^*)^T{\Sigma_k^*}^{-1}(\mathbf{f}-\mathbf{f}_k^*)\right]\right)d\mathbf{f}\\ &= w_kC_kC'\exp\left\{-\frac{1}{2}D_k\right\}\times\int\exp\left[-\frac{1}{2}(\mathbf{f}-\mathbf{f}_k^*)^T{\Sigma_k^*}^{-1}(\mathbf{f}-\mathbf{f}_k^*)\right]d\mathbf{f}\\ &= w_kC_kC'\exp\left\{-\frac{1}{2}D_k\right\}\sqrt{(2\pi)^L|\Sigma_k^*|}. \end{aligned} \quad (28)$$

A becomes

$$A = \frac{1}{\sum_{k=1}^{K}B_k}. \quad (29)$$

Therefore, finally we have the estimate

$$\hat{\mathbf{f}} = \sum_{k=1}^{K}\frac{B_k}{\sum_{j=1}^{K}B_j}\mathbf{f}_k^* = \sum_{k=1}^{K}m_k\hat{\mathbf{f}}_k^*, \quad (30)$$

where m_k indicates the mixing proportion of each estimates $\mathbf{f}_k^*$. If m_k's are equal to 0 except at $k=k'$, the estimate becomes

$$\hat{\mathbf{f}} = \mathbf{f}_{k'}^* = \hat{\mathbf{f}}_{k'}^{\text{Wiener}}. \quad (31)$$

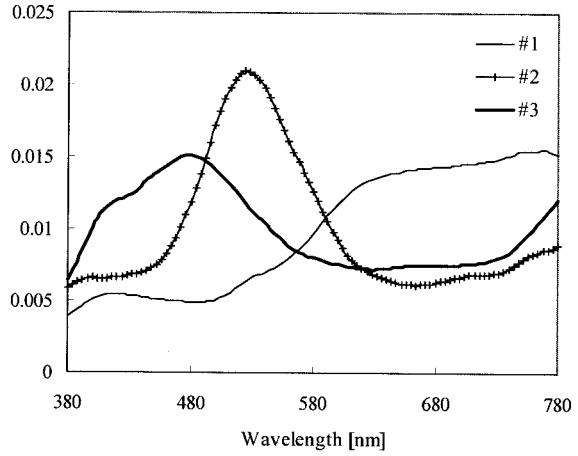
It is the same as the Wiener estimate obtained under the assumption that $\mathbf{f}$ is a random sequence of ${p_{k'}}^G(\mathbf{f})$.

3. Simulations

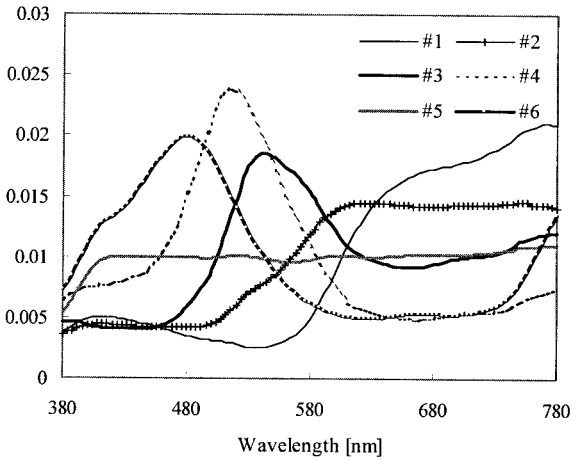
We compare the estimation accuracy of spectral reflectance from camera responses between Wiener and GMD-based estimations. In this comparison, we use samples of spectral reflectance with various colors, which were classified into six or three classes by the clustering technique.

A. Samples of Spectral Reflectance and Their Classification

We use spectral reflectances from 168 mat patches of GretagMacbeth Color Checker DC, measured by spectrophotometer (PhotoResearch PR650). The measuring range was 380–780 nm with a 4-nm wavelength interval. To classify the samples, we use K-means algorithm[8]; $K = 6$ and 3. Before clustering, all reflectances are normalized so that the integrations of them are equal to 1. The reason for this operation is that the spectral reflectances that have the same spectral shape but different power can be assumed to have the same statistical characteristics and should be placed in the same classification. Then the principal components analysis is performed to reduce the data dimension, which reduces the it-

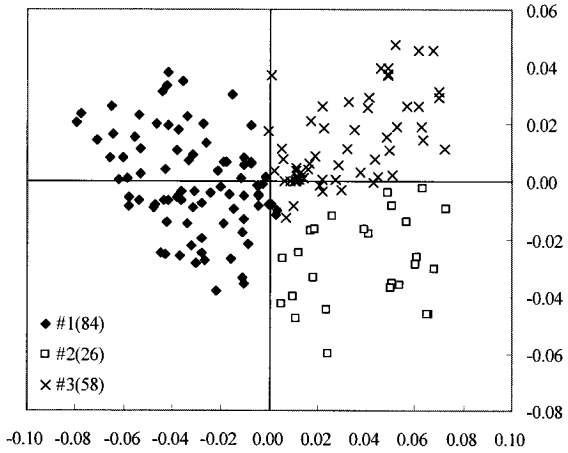


(a)

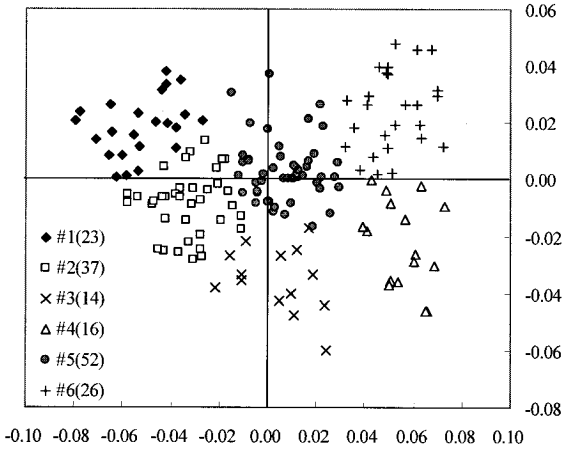


(b)

Fig. 1. Centers of the clusters in spectrum form, in the case of (a) $K = 3$ and (b) $K = 6$.



(a)



(b)

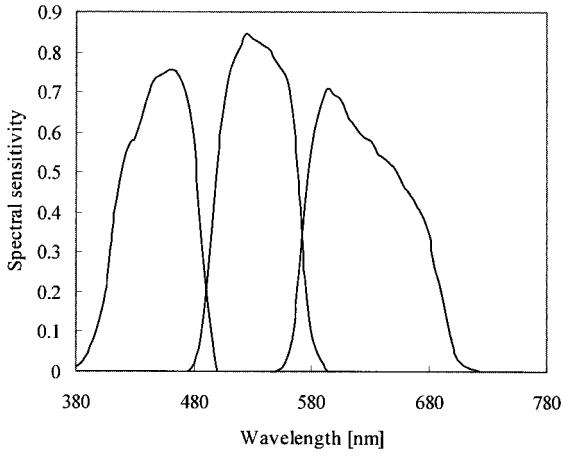
Fig. 2. Clustered samples on the plane whose horizontal and vertical axes are the first and the second principal components in the case of (a) $K = 3$ and (b) $K = 6$. The designated numbers in the parentheses are the number of the samples classified into a respective class.

eration time for clustering. Figure 1 shows the centers of the clusters that resulted from this classification. Figure 2 shows the clustered samples projected onto the plane where the horizontal and vertical axes are the first and second principal components, respectively. Though we cannot say that the GMD-based model based on either clustering result is the best to represent the data distribution, they are superior to the model by a single Gaussian from the viewpoint of the Akaike information criterion.[9]

B. Color Estimation from Camera Responses

We carried out the simulations using the spectral sensitivities of three-band and six-band cameras. As a three-band camera, the spectral sensitivity of FD420M (Flovel), shown in Fig. 3(a), is used. The sensitivities of six-band camera are shown in Fig. 3(b), which is the sensitivities of the six-band camera assembled in our laboratory from the two FD420Ms and two kinds of filters. The illumination light is assumed to be a D65 illuminant. Three- or six-dimensional image signals are calculated from the 168 spectral reflectances, and each spectral reflectance is estimated with Wiener- and GMD-based estimation methods.

The Wiener estimation is carried out with Eq. (24) instead of Eq. (11) because the average reflectance does not become a zero vector, where the covariance and average are computed from the following two-type ensembles; all 168 samples (indicated by Wiener in the results shown below) and the samples belonging to kth class [Wiener(#k)]. GMD-based estimation is carried out by use of Eqs. (24)–(30) (GMD). The kth-class covariance and average are computed from the spectral reflectance of the samples belonging to the kth class, where each class is regarded as a component of the GMD. The estimation error is measured in the spectral space by normalized root mean square error (NRMSE) and in CIE 1976 L*a*b* color space, E_{ab}*, where the color matching function



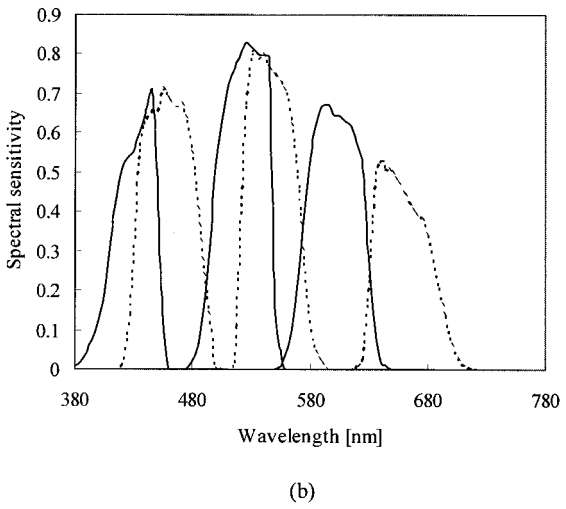


Fig. 3. Spectral sensitivities of the cameras used in the simulations: (a) three bands and (b) six bands.

is CIE 1931 and the viewing illuminants are A, D65, and F2.

Figure 4 shows the NRMSE of spectral reflectance functions for every class of samples in the case of $K = 3$ and three-band camera. In this graph, #k means the kth class samples and the label All means all 168 samples. We can see that the minimum NRMSE for kth class samples is obtained by the Wiener(#k) method. However, the errors of the samples of non-kth classes are considerably large. In contrast, GMD realizes almost the same accuracy to the corresponding Wiener(#k) in every class. As a result, for All, the error of the GMD becomes minimum.

Figure 5 shows the results of the same simulation as Fig. 4, in which the error is measured by the average E_{ab}* under viewing illuminant A. This graph shows the same tendencies as Fig. 4.

Figure 6 shows the average and the maximum E_{ab}* of 168 samples when the viewing illuminants are A, D65, and F2, the cameras are three and six bands, and the estimation methods are Wiener and GMD

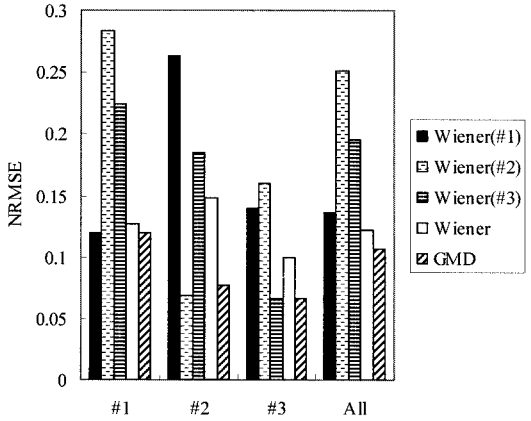


Fig. 4. NRMSE of every class of samples for the case of $K = 3$ and the three-band camera.

($K = 3$ and $K = 6$). For every illuminant, it can be confirmed that the GMD reduces both the average and the maximum error in comparison with the Wiener estimation. Especially for $K = 6$, the errors of the GMD are about half of those of the Wiener estimation for both three- and six-band cases. The reason why $K = 6$ cases are superior to $K = 3$ cases can be thought of as follows. As shown later, GMD-based estimates become similar to Wiener estimates by use of corresponding class statistics in both the cases of $K = 6$ and $K = 3$. In the Wiener estimate, the smaller the deviation of the data in a class, the higher the accuracy of the estimates become. In accordance with the fact that the deviation of a class is smaller in $K = 6$ than in $K = 3$, the accuracy of $K = 6$ can be thought of as being higher than that of $K = 3$.

Finally, let us consider the mixing parameter m_k of each sample. Figure 7 shows the proportion of m_k of each of the samples obtained through the three-band camera simulation. Samples are listed in the order of classes along the horizontal axis. The vertical axis indicates the proportion of m_k's. From the graphs, we can see that the m_k's for the samples of the kth class are dominant to other mixing parameters. This ten-

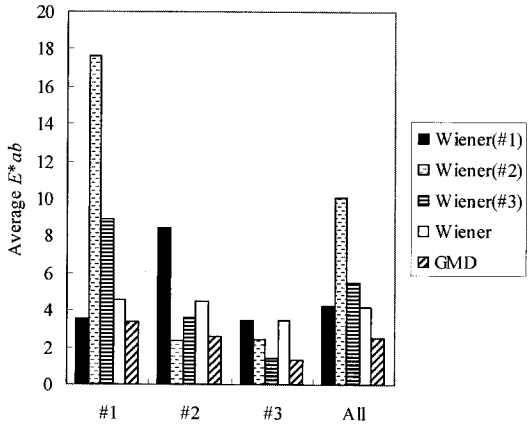


Fig. 5. Average E_{ab}* of every class of samples in the case of $K = 3$, three-band camera, and viewing illuminant A.

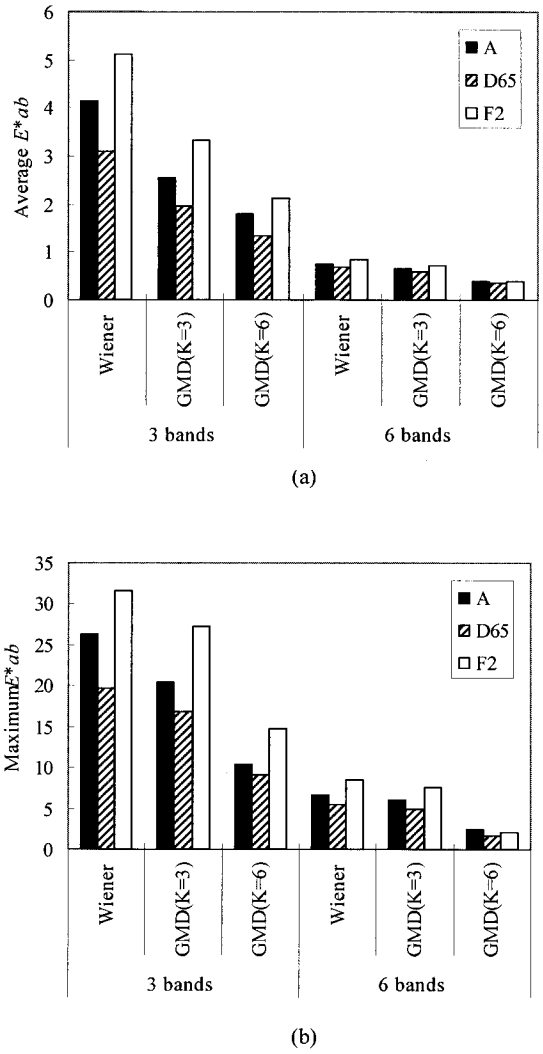


Fig. 6. Average and maximum E_{ab}^* of all samples by Wiener and GMD-based estimations.

dency can be seen in both cases of $K = 3$ and $K = 6$. These results indicate that an estimate of the samples in the kth class obtained by the GMD-based estimation becomes similar to the estimate from the Wiener estimation selectively by use of its own statistics.

4. Conclusions

We proposed a nonlinear-estimation method for an inverse problem, a GMD-based estimation, assuming that the input is a random sequence of Gaussian mixture distribution. The significant property of the proposed method is to be derived through analytical procedures in spite of the nonlinear estimation. That is, the GMD estimate is a linear combination of the Wiener estimates, which are derived by use of the probability distribution of one of the components as *a priori* information.

The proposed method is applied to the problem of a spectral-reflectance estimation from camera responses. As a result, the GMD-based estimation reduces the estimation error by up to half in comparison with the conventional method, the Wiener estimation. The improvement in accuracy depends on the accuracy of the model of the probability density of spectral reflectance in the proposed method. Then, it becomes an important issue that how to make an appropriate model for target ensemble based on GMD.

Appendix A: Derivation of the Formulas

In this appendix, we show that the complete form of the probability density $\bar{p}_k{}^G(\mathbf{f}|\mathbf{g})$ is given by Eqs. (26) and (27). Ref. 7 derives the relative form of $\bar{p}_k(\mathbf{f}|\mathbf{g})$, which is the same as $p_k(\mathbf{f}|\mathbf{g})$ in Eq. (23), then it is enough to derive only the constant term.

Substitute Eqs. (4), (9), and (22) into Eq. (17), we have

$$\begin{aligned}\bar{p}_k{}^G(\mathbf{f}|\mathbf{g}) &= w_k C_k \exp\left[-\frac{1}{2}(\mathbf{f}-\langle\mathbf{f}_k\rangle)^T\Sigma_k{}^{-1}(\mathbf{f}-\langle\mathbf{f}_k\rangle)\right] \\ &\quad\times C_n \exp\left[-\frac{1}{2}(\mathbf{g}-\mathbf{H}\mathbf{f})^T\Sigma_n{}^{-1}(\mathbf{g}-\mathbf{H}\mathbf{f})\right] \\ &= w_k C_k C_n \exp\left[-\frac{1}{2}[(\mathbf{f}-\langle\mathbf{f}_k\rangle)^T\Sigma_k{}^{-1}(\mathbf{f}-\langle\mathbf{f}_k\rangle)\right. \\ &\quad\left.+(\mathbf{g}-\mathbf{H}\mathbf{f})^T\Sigma_n{}^{-1}(\mathbf{g}-\mathbf{H}\mathbf{f})]\right]. \qquad \text{(A1)}\end{aligned}$$

We now consider the argument of the exponential not including variable $\mathbf{f}$:

$$\langle\mathbf{f}_k\rangle^T\Sigma_k{}^{-1}\langle\mathbf{f}_k\rangle + \mathbf{g}^T\Sigma_n{}^{-1}\mathbf{g}, \qquad \text{(A2)}$$

where the factor of $-1/2$ is omitted. We want to show that this expression equals to

$$\mathbf{f}_k{}^{*T}\Sigma_k{}^{*-1}\mathbf{f}_k{}^* + D_k, \qquad \text{(A3)}$$

the argument of the exponential of Eq. (26) not including $\mathbf{f}$.

Before entering the derivation, we make some preparations. First, we shall show two useful general matrix identities[7]:

$$\begin{aligned}&\mathbf{R}_2 - \mathbf{R}_2\mathbf{M}^T(\mathbf{R}_1 + \mathbf{M}\mathbf{R}_2\mathbf{M}^T)^{-1}\mathbf{M}\mathbf{R}_2 \\ &\qquad = (\mathbf{R}_2{}^{-1} + \mathbf{M}^T\mathbf{R}_1{}^{-1}\mathbf{M})^{-1}, \qquad \text{(A4)}\end{aligned}$$

$$\begin{aligned}&\mathbf{R}_2\mathbf{M}^T(\mathbf{R}_1 + \mathbf{M}\mathbf{R}_2\mathbf{M}^T)^{-1} \\ &\qquad = (\mathbf{R}_2{}^{-1} + \mathbf{M}^T\mathbf{R}_1{}^{-1}\mathbf{M})^{-1}\mathbf{M}^T\mathbf{R}_1{}^{-1}, \qquad \text{(A5)}\end{aligned}$$

where $\mathbf{R}_1$ and $\mathbf{R}_2$ are symmetric matrices whose inverse exits, and $\mathbf{M}$ is a third matrix. Applying those to the related terms, we have three equations:

$$\begin{aligned}&\Sigma_k - \Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{H}\Sigma_k \\ &\qquad = (\Sigma_k{}^{-1} + \mathbf{H}^T\Sigma_n{}^{-1}\mathbf{H})^{-1}, \qquad \text{(A6)}\end{aligned}$$

$$\begin{aligned}&\Sigma_n{}^{-1} - \Sigma_n{}^{-1}\mathbf{H}(\Sigma_k{}^{-1} + \mathbf{H}^T\Sigma_n{}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\Sigma_n{}^{-1} \\ &\qquad = (\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}, \qquad \text{(A7)}\end{aligned}$$

$$\Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1} = (\Sigma_k{}^{-1} + \mathbf{H}^T\Sigma_n\mathbf{H})^{-1}\mathbf{H}^T\Sigma_n{}^{-1}. \qquad \text{(A8)}$$

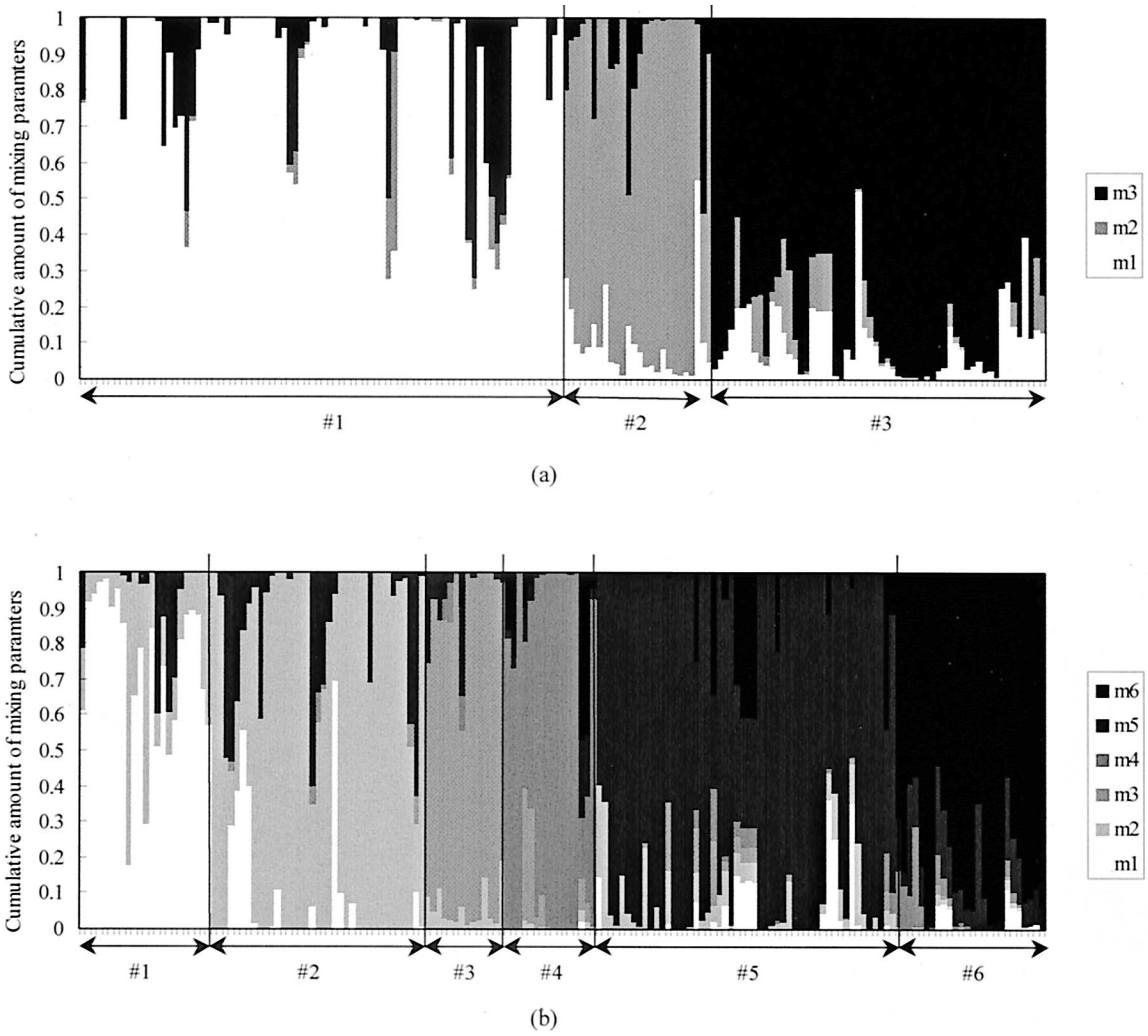


Fig. 7. Proportion of mixing parameters for all samples: (a) $K = 3$ and (b) $K = 6$. Samples are listed in the order of classes along the horizontal axis.

Because Σ_k^* is represented by Eq. (25), $\Sigma_k^* = \Sigma_k - \Sigma_k \mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{H}\Sigma_k$, Eqs. (A6)–(A8) can be rewritten using Σ_k^* as follows:

$$\Sigma_k - \Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{H}\Sigma_k = (\Sigma_k^{-1} + \mathbf{H}^T\Sigma_n^{-1}\mathbf{H})^{-1} = \Sigma_k^*, \quad \text{(A9)}$$

$$\Sigma_n^{-1} - \Sigma_n^{-1}\mathbf{H}\Sigma_k^*\mathbf{H}^T\Sigma_n^{-1} = (\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}, \quad \text{(A10)}$$

$$\Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1} = \Sigma_k^*\mathbf{H}^T\Sigma_n^{-1}. \quad \text{(A11)}$$

This is the end of the preparation, and Eqs. (A9)–(A11) are used in the following.

We now start from Eq. (A3). The vector, $\mathbf{f}_k^*$, which first appears in Eq. (24), can be rewritten as

$$\mathbf{f}_k^* = \langle\mathbf{f}_k\rangle + \Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}(\mathbf{g} - \mathbf{H}\langle\mathbf{f}_k\rangle)$$
$$= \{\mathbf{I} - \Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{H}\}\langle\mathbf{f}_k\rangle + \Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{g}. \quad \text{(A12)}$$

Using Eq. (A9) and (A11) for the first and second terms, respectively, we have

$$\mathbf{f}_k^* = \Sigma_k^*\Sigma_k^{-1}\langle\mathbf{f}_k\rangle + \Sigma_k^*\mathbf{H}^T\Sigma_n^{-1}\mathbf{g}. \quad \text{(A13)}$$

Substituting Eq. (A13) into $\mathbf{f}_k^{*T}\Sigma_k^{*-1}\mathbf{f}_k^*$ gives

$$\mathbf{f}_k^{*T}\Sigma_k^{*-1}\mathbf{f}_k^* = \langle\mathbf{f}_k\rangle^T\Sigma_k^{-1}\Sigma_k^*\Sigma_k^{-1}\langle\mathbf{f}_k\rangle + 2\langle\mathbf{f}_k\rangle^T\Sigma_k^{-1}\Sigma_k^*\mathbf{H}^T\Sigma_n^{-1}\mathbf{g} + \mathbf{g}^T\Sigma_n^{-1}\mathbf{H}\Sigma_k^*\mathbf{H}^T\Sigma_n^{-1}\mathbf{g}. \quad \text{(A14)}$$

D_k is represented by Eq. (27), which is expanded as

$$D_k = (\mathbf{g} - \mathbf{H}\langle\mathbf{f}_k\rangle)^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}(\mathbf{g} - \mathbf{H}\langle\mathbf{f}_k\rangle)$$
$$= \mathbf{g}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{g} - 2\langle\mathbf{f}_k\rangle^T\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{g} + \langle\mathbf{f}_k\rangle^T\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{H}\langle\mathbf{f}_k\rangle. \quad \text{(A15)}$$

Substituting Eq. (A10), the first term of Eq. (A15) can be rewritten as

$$\mathbf{g}^T(\mathbf{H}\Sigma_f\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{g} = \mathbf{g}^T(\Sigma_n^{-1} - \Sigma_n^{-1}\mathbf{H}\Sigma_k^*\mathbf{H}^T\Sigma_n^{-1})\mathbf{g}. \quad \text{(A16)}$$

Substituting Eq. (A11) multiplied by Σ_k^{-1} from the left-hand side into the second term of Eq. (A15) gives

$$-2\langle \mathbf{f}_k\rangle^T \mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{g} = -2\langle \mathbf{f}_k\rangle^T \Sigma_k^{-1}\Sigma_k^* \mathbf{H}^T\Sigma_n^{-1}\mathbf{g}. \quad \text{(A17)}$$

By use of Eq. (A9), the third term of Eq. (A15) becomes

$$\begin{aligned}\langle \mathbf{f}_k\rangle^T \mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\mathbf{H}\langle \mathbf{f}_k\rangle &= \langle \mathbf{f}_k\rangle^T \Sigma_k^{-1}\Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\Sigma_k\Sigma_k^{-1}\mathbf{H}\langle \mathbf{f}_k\rangle \\ &= \langle \mathbf{f}_k\rangle^T \Sigma_k^{-1}\{\Sigma_k\mathbf{H}^T(\mathbf{H}\Sigma_k\mathbf{H}^T + \Sigma_n)^{-1}\Sigma_k\}\Sigma_k^{-1}\mathbf{H}\langle \mathbf{f}_k\rangle \\ &= \langle \mathbf{f}_k\rangle^T \Sigma_k^{-1}(\Sigma_k - \Sigma_k^*)\Sigma_k^{-1}\langle \mathbf{f}_k\rangle. \quad \text{(A18)}\end{aligned}$$

Then we have

$$\begin{aligned}D_k = \mathbf{g}^T(\Sigma_n^{-1} - \Sigma_n^{-1}\mathbf{H}\Sigma_k^*\mathbf{H}^T\Sigma_n^{-1})\mathbf{g} \\ - 2\langle \mathbf{f}_k\rangle^T \Sigma_k^{-1}\Sigma_k^*\mathbf{H}^T\Sigma_n^{-1}\mathbf{g} \\ + \langle \mathbf{f}_k\rangle^T \Sigma_k^{-1}(\Sigma_k - \Sigma_k^*)\Sigma_k^{-1}\langle \mathbf{f}_k\rangle. \quad \text{(A20)}\end{aligned}$$

Adding $\mathbf{f}_k^{*T}\Sigma_k^{*-1}\mathbf{f}_k^*$ to D_k, which cancels some terms, gives

$$\mathbf{f}_k^{*T}\Sigma_k^{*-1}\mathbf{f}_k^* + D_k = \langle \mathbf{f}_k\rangle^T \Sigma_k^{-1}\langle \mathbf{f}_k\rangle + \mathbf{g}^T\Sigma_n^{-1}\mathbf{g}. \quad \text{(A21)}$$

The right-hand side of the equation is identical to the expression of Eq. (A3). We have, therefore, finished the derivation that we want to show.

References

1. Y. Ohya, T. Obi, M. Yamaguchi, N. Ohyama, and Y. Komiya, "Natural color reproduction of human skin for telemedicine," in *Medical Imaging 1998: Image Display*, Y. Kim and S. K. Mun, eds. Proc. SPIE **3335,** 263–270 (1998).
2. H. Haneishi, T. Hasegawa, A. Hosoi, Y. Yokoyama, N. Tsumura, and Y. Miyake, "System design for accurately estimating spectral reflectance of art paintings," Appl. Opt. **39,** 6621–6632 (2000).
3. Th. Keusen, "Multispectral color system with an encoding format compatible with the conventional tristimulus model," J. Imaging Sci. Technol. **40,** 510–515 (1996).
4. F. Koenig, "Reconstruction of natural spectra from color sensor using nonlinear estimation methods," in *Proceedings of IS&T's 50th Annual Conference: A Celebration of All Imaging*, (The Society for Imaging Science and Technology, Springfield, Va., 1997), pp. 454–458.
5. N. Tsumura, H. Sato, T. Hasegawa, H. Haneishi, and Y. Miyake, "Limitation of color samples for spectral estimation from sensor responses in fine art painting," Opt. Rev. **6,** 57–61 (1999).
6. F. H. Imai, R. S. Berns, and Di-Y. Tzeng, "A comparative analysis of spectral reflecntace estimated in various spaces using a trichromatic camera system," J. Imaging Sci. Technol. **44,** 280–287 (2000).
7. W. Menke, "Geophysical Data Analysis: Discrete Inverse Theory," (Academic Press, Inc., San Diego, Calif., 1989), pp. 92–99.
8. J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, (University of California Press, Berkeley, 1967) Vol. 1, pp. 281–297.
9. H. Akaike, "Factor analysis and AIC," Psychometrika, **52,** 317–332 (1987).