Novel Mixture Model for Mixed Pixel Classification of Multispectral Image Data

MITSUHIRO TOMOSADA[1] and HIROE TSUBAKI[2]

[1]Central Research Institute of Electric Power Industry, 2-11-1. Iwado Kita, Komae-shi, Tokyo 201-8511, Japan tomosada@criepi.denken.or.jp

[2]The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo, 106-8569, Japan

***Abstract*—**We propose a novel mixture model for use in the mixed pixel classification (MPC) of a multispectral image such as remotely sensed multispectral image data and Magnetic Resonance Image (MRI). Although the MPC method utilizes a generic statistical model of mixture such as a linear mixture model or a finite mixture model, the proposed mixture model of a single pixel in this paper is established on the basis of the process of mixed pixel generation in a real multispectral image. The variance-covariance structure for a pixel vector is considerably different from the variance-covariance structure derived from existing mixture models. Furthermore, we present an MPC method using the generalized method of moments (GMM), which satisfies the proposed mixture model and estimates the mixing ratio for each component in a single pixel, and the expected value and variance-covariance matrix of the pixel vector for each component.

First, the MPC method is applied to a simulated image data. Estimated parameters close to actual values are obtained, and the simulated image data is found to be in agreement with the constructed mixture model since the evaluation function is close to the actual value. The proposed mixture model and MPC method are applied to real multispectral image data acquired by Enhanced Thematic Mapper Plus (ETM^+) onboard Landsat-7satellite as one example of a multispectral image. As a result, it was found that pixels in an ETM^+ image, for which the mixing ratio of one component is high, are consistent with pixels in an image which are assumed to have the same component by visual inspection.

***Keywords*—**Generalized method of moment, Image processing, Mixed pixel classification, Mixture model, Multispectral image.

I. Introduction

We propose a new mixture model for use in the mixed pixel classification (MPC) of a remotely sensed multispectral image. A multispectral image is composed of images taken by different wavelength for same target area. In this paper, an image is indicated to a digitized image. A Multispectral image is used for many fields. For example, in a field of environmental study, a remotely sensed multispectral image acquired sensor onboard satellite is used for an analysis of environmental variation. And, in a field of medical science, a Magnetic Resonance Image (MRI) is used for recognition of disease. The smallest piece of information in a digitized multispectral image is a pixel. The values of a single pixel in a multispectral image are represented by a pixel vector. In a target area of a single pixel that is called the spatial resolution of sensor, signals derived from some components on the target area are frequently detected by a single pixel. For example, a pixel vector of visible and near infrared band image data, which are observed by the Enhanced Thematic Mapper Plus (ETM^+) sensor onboard Landsat-7, are acquired as the averaged signal of a 30 m × 30 m area on the earth's surface. In a multispectral image, a single pixel that detects signals derived from several components such as a forest, the sea, and a farm field in the target area of pixel is called a mixed pixel. And, for MRI, sensor detects signals derived from a blood vessel, a fat, bone, and so on. On the other hand, a pixel signal derived from only one component is called a pure pixel. Even with sensors of relatively high resolution, there are cases that mixed pixel is existed in a multispectral image. As one example of actual utilization, a single pixel in a multispectral image data is classified to a certain component. However, in the case that a single pixel, which is actually a mixed pixel, is assumed to a pure pixel, this single pixel can't be classified correctly. So, one expects to have to deal with mixed pixels. Therefore, it is necessary to conduct sub-pixel analysis to obtain precise information of target area in a multispectral image data.

There have been some existing studies on MPC methods. MPC is conducted as follows. First, mixture model to describe a mixed pixel is set, and then mixing ratios of components included in a target area of a single pixel are estimated. Examples of previous mixture models include a linear mixture model for Independent Component Analysis (ICA) [1] and a finite mixture probability distribution model [2]. These mixture models are general statistical mixture models. Whether these mixture models are appropriate for representing a mixed pixel of a multispectral image has not been sufficiently discussed. The mixture model proposed in this study has been established on the basis of the process of mixed pixel generation in an actual multispectral image. In this paper, first, the difference between existing mixture models used in MPC methods and our proposed mixture model is presented. Then, an MPC method

based on the proposed mixture model is presented.

In the next section, existing mixture models used in the MPC method are introduced. We describe in detail the proposed mixture model and MPC method in Section Ⅲ. In Section Ⅳ, the proposed MPC method is applied to simulated multispectral image data and to an actual Landsat-7 ETM$^+$ image as one example of a multispectral image. Finally, we present our conclusions in Section Ⅴ.

II. Existing Mixture Model

In this paper, the number of spectral bands is denoted by P, $\mathbf{y}$ is a pixel vector of a P-dimensional multispectral image, Q is the number of components within a single pixel, a_q is a mixing ratio of a component q in a single pixel, $\mathbf{y}_q$ is a vector of a pure pixel for component q, and $\boldsymbol{\mu}_q$ and $\boldsymbol{\Sigma}_q$ are a mean vector and a variance-covariance matrix of $\mathbf{y}_q$. The boldface is used for the vector and matrix. Superscript T denotes the transpose of a matrix.

In an analysis using multispectral image, linear spectral mixing is a widely used approach for determining and quantifying multiple components in an area detected by a single pixel. The linear mixture model [1][3][4][5][6][7], for example, can be expressed as

$$\mathbf{y} = \sum_{q=1}^{Q} a_q \cdot \mathbf{y}_q + \boldsymbol{\varepsilon}_y \ , \quad \boldsymbol{\varepsilon}_y \sim N(\mathbf{0}, \boldsymbol{\Omega}_y) \tag{1}$$

where $\boldsymbol{\varepsilon}_y$, whose expected vector is equal to $\mathbf{0}(=(0\ldots0)^T)$ and whose variance-covariance matrix is $\boldsymbol{\Omega}_y$, is a noise vector that can be interpreted as a measurement error. A classical approach [3] to solve mixed pixel classification is to estimate a_q in (1) by the constrained least-squares method and weighted least-squares method. Constraints are given as

$$\sum_{q=1}^{Q} a_q = 1, \tag{2}$$

and for all q

$$0 \le a_q. \tag{3}$$

Choi *et al*. [4] assumed that the conditional distribution of the observation $\mathbf{y}$ is a P-dimensional Gaussian with the mean vector

$$\boldsymbol{\mu} = \sum_{q=1}^{Q} a_q \boldsymbol{\mu}_q \ , \tag{4}$$

and with the variance-covariance matrix

$$\boldsymbol{\Omega} = \sum_{q=1}^{Q} a_q^2 \boldsymbol{\Sigma}_q + \boldsymbol{\Omega}_y \ , \tag{5}$$

where $\boldsymbol{\mu}$ be a P-dimensional mean vector of pixel, $\boldsymbol{\mu}_q$ be mean pixel vector of component q, and $\boldsymbol{\Sigma}_q$ is variance-covariance matrix of component q. H. S. Choi *et al*., who assumed that $\boldsymbol{\Sigma}_q$ is negligible, estimated a_q and $\boldsymbol{\mu}_q$ for all q by iterative calculation. P. Bosdogianni *et al*. [8], who calculated the first moment (4) and the second moment (5) of (1), estimated a_q using the Least Square Error method. Mixture model used in [8] is not included $\boldsymbol{\varepsilon}_y$ in (1). Furthermore, Chang [1] estimated a_q and $\boldsymbol{\mu}_q$ by applying Independent Component Analysis (ICA) that uses the higher order moment as the third- and fourth-order moments.

Another mixture model of a mixed pixel is the finite mixture model. The finite mixture model can be expressed as

$$f(\mathbf{y}) = \sum_{q=1}^{Q} a_q \cdot f_q(\mathbf{y}_q), \tag{6}$$

where $f(\mathbf{y})$ is the probability distribution function of a pixel vector $\mathbf{y}$, and $f_q(\mathbf{y}_q)$ is the probability distribution function of the pixel vector $\mathbf{y}_q$ for a pure pixel of component q. Santago *et al*. [2] estimated a_q from the difference between the histogram of $\mathbf{y}$ and $f(\mathbf{y})$ under the condition that (2) and (3) are satisfied. Another approach [9] using the finite mixture model to estimate mixing ratios of components is EM algorithm. The EM algorithm is applied only when the probability distribution function of each component is a normal distribution. Therefore, in the case where the EM algorithm is applied to multispectral images, there is a constraint in that the probability distribution function of pixel vector for each component is assumed to be a normal distribution.

The expected value and variance-covariance matrix for each component are usually obtained by calculating from the 'almost' pure pixel following a visual inspection. Since it is difficult to extract perfectly pure pixels, the expected value and variance-covariance matrix include some errors. Essentially, the expected value and variance-covariance matrix have to be estimated. In reference [10], the 'almost' pure pixels for a certain component q are considered complete data, namely a_q is determined to 1. Mixed pixel is regarded as incomplete data, and $\boldsymbol{\mu}_q$ for all q is estimated. It is considered that the set of complete data also includes other components to some extent, and $\boldsymbol{\Sigma}_q$ are not estimated. Therefore it is necessary for a given method to estimate $\boldsymbol{\mu}_q$ and $\boldsymbol{\Sigma}_q$ from a whole multispectral image.

III. Proposition

First, a mixture model is proposed in Section Ⅲ*A*, and then an MPC method using the mixture model proposed in Section Ⅲ*A* is presented in Section Ⅲ*B*. Finally, evaluation method of MPC is described in Ⅲ*C*.

A. Mixture model

We assume that a pixel vector of a mixed pixel in a multispectral image is obtained as the averaged signal in an area of signal sources on the target area detected by a single pixel. Signals detected by a single pixel derived from each component on the target area are assumed to be independent. A signal detected by a mixed pixel is defined as the sum of signals for elements of each component in an observation area of a single pixel. Although not actually conducted, suppose P-variate pixel vector $\mathbf{y}$ is decomposed into K same-size sub-pixels (Fig.1). A decomposed sub-pixel is called a micro-pixel in this study to distinguish it from a single pixel of a multispectral image. Each micro-pixel is assumed to be a single component. The number of micro-pixels belonging to a component q in a single pixel is denoted by K_q. The relationship between K and K_q is given by

$$\sum_{q=1}^{Q} K_q = K \ . \tag{7}$$

A P-variate vector of a micro-pixel for the k-th element of

component q in a single pixel is denoted by $\mathbf{x}_{qk}$, which is the realized value of random vector $\mathbf{X}_q$ with the P-variate probability distribution function $g_q(\mathbf{x}_q)$. A P-variate pixel vector $\mathbf{y}$ of a multispectral image is assumed to be obtained as the sum of $\mathbf{x}_{qk}$, that is, $\mathbf{y}$ is given by

$$\mathbf{y} = \sum_{q=1}^{Q} \sum_{k=1}^{Kq} \mathbf{x}_{qk} \,. \tag{8}$$

A pure-pixel vector $\mathbf{y}_q$ of component q is represented as all of the micro-pixels belonging to component q, where $\mathbf{y}_q$ is given by

$$\mathbf{y}_q = \sum_{k=1}^{K} \mathbf{x}_{qk} \,. \tag{9}$$

Let the expected value of a certain vector $\mathbf{v}$ be denoted by $E[\mathbf{v}]$, and $\mathbf{t}$ is an arbitrary real number vector. The moment generating function of $\mathbf{x}_{qk}$, is written as

$$\phi_q = E[\exp(\mathbf{t}^T \mathbf{x}_{qk})] \,. \tag{10}$$

$\mathbf{y}_q$ is the realized value of random vector $\mathbf{Y}_q$ with the P-variate probability distribution function $f_q(\mathbf{y}_q)$. The moment generating function of $\mathbf{y}_q$ derived from (9), which is written as $\varphi_q(\mathbf{t})$, is given by

$$\varphi_q(\mathbf{t}) = \phi_q(\mathbf{t})^K \,. \tag{11}$$

The moment generating function of $\mathbf{x}_{qk}$ is given by

$$\phi_q(\mathbf{t}) = \varphi_q(\mathbf{t})^{1/K} \,. \tag{12}$$

Let a_q be given by

$$a_q = \frac{K_q}{K} \,. \tag{13}$$

The moment generating function of $\sum_{k=1}^{K_q} \mathbf{x}_{qk}$ in (8) is $\varphi_q(\mathbf{t})^{K_q/K}$. The moment generating function of P-variate vector $\mathbf{y}$ is

$$\varphi_y(\mathbf{t}) = \prod_{q=1}^{Q} \varphi_q(\mathbf{t})^{a_q} \,. \tag{14}$$

As indicated above, a mixed pixel described by (8) is modeled by (14), and (14) is the mixture model proposed in this paper. The expected value $\boldsymbol{\mu}$ of $\mathbf{y}$ is derived from the first cumulant of (14). The variance-covariance matrix $\boldsymbol{\Omega}$ of $\mathbf{y}$ is derived from the second cumulant of (14). The cumulant is given by the logarithm of the moment generating function. The values of $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ of $\mathbf{y}$ are given by

$$\boldsymbol{\mu} = \sum_{q=1}^{Q} a_q \cdot \boldsymbol{\mu}_q \,, \tag{15}$$

and

$$\boldsymbol{\Omega} = \sum_{q=1}^{Q} a_q \cdot \boldsymbol{\Sigma}_q \,. \tag{16}$$

The mixture model, which is established based on the pixel of a real multispectral image, is obtained from (14). The expected value of $\mathbf{y}$ is consistent with (15), and the variance-covariance matrix of $\mathbf{y}$ is consistent with (16). For example, when $\mathbf{y}_q$ follows a P-variate normal distribution, the distribution of $\mathbf{y}$ also follows a P-variate normal distribution, whose expected value is equal to (15) and whose variance-covariance matrix is equal to (16). A pixel vector $\mathbf{y}_i$ is given by

$$\mathbf{y}_i = \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_i \,, \tag{17}$$

where $\boldsymbol{\varepsilon}_i$ denotes an error vector with the mean vector $\mathbf{0}$ and the variance-covariance matrix $\boldsymbol{\Omega}_i$ given by (16). In the existing mixture model such as finite mixture model and the linear mixture model, it is necessary to assume the probability distribution function f_q of $\mathbf{y}_q$. However, a mixed-pixel vector $\mathbf{y}$ is consistent with (14) without depending on the probability distribution of $f_q(\mathbf{y}_q)$. Therefore, it is not necessary to assume the probability distribution of $\mathbf{y}_q$.

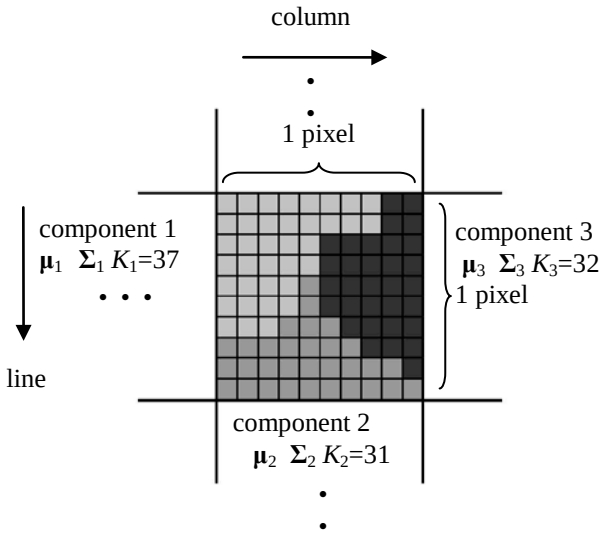


Fig.1 Overview of micro-pixel. A single pixel is depicted as an area surrounded by bold lines and micro-pixels are depicted as areas surrounded by light lines. Taking the figure above as an example, Q=3 and K=100.

The expected value of $\mathbf{y}$ derived from the mixture models expressed by (1), (6), and (14) is given by (15). However, the variance-covariance matrix of $\mathbf{y}$, which is derived from the finite mixture model (6), is given by

$$\boldsymbol{\Omega} = \sum_{q=1}^{Q} a_q \cdot \boldsymbol{\Sigma}_q + \sum_{q=1}^{Q} a_q \cdot \boldsymbol{\mu}_q \cdot \boldsymbol{\mu}_q^T - \sum_{q=1}^{Q} \sum_{r=1}^{Q} a_q \cdot a_r \cdot \boldsymbol{\mu}_q \cdot \boldsymbol{\mu}_r^T \tag{18}$$

When $f_q(\mathbf{y}_q)$ in (6) follows the P-variate normal distribution, the distribution of $\mathbf{y}$ does not follow the P-variate normal distribution. The variance-covariance matrix given by (18) is larger than that given by (16). Note that whether one matrix $\mathbf{A}$ is larger than another matrix $\mathbf{B}$ is judged as follows. A quadratic form is built from $\mathbf{A}$-$\mathbf{B}$. If quadratic form is positive definite then the matrix $\mathbf{A}$ is larger than $\mathbf{B}$. In contrast, the variance-covariance matrix (5) is also different from (16), and the variance-covariance matrix of (5) is smaller than the variance-covariance (16) derived from (14) when $\boldsymbol{\Omega}_y$ in (5) is smaller considerably than $\Sigma a_q^2 \boldsymbol{\Sigma}_q$. Therefore, the uncertainty of the variance-covariance matrix increases with the number of components in a single pixel. As for a mixed pixel in a multispectral image, the variance-covariance structure described by existing mixture models cannot satisfactorily represent the mixed pixel. The proposed mixture model (14) is

better able to express a mixed pixel than existing mixture models.

B. Mixed Pixel Classification

An MPC method satisfying the proposed mixture model of (14) is presented. Let N be the number of pixels in a multispectral image, $\mathbf{a}^*$, $\mathbf{\Sigma}^*$, $\boldsymbol{\mu}^*$, and $\mathbf{y}^*$ denote $\mathbf{a}^* = (\mathbf{a}_1 \mathbf{a}_2 \cdots \mathbf{a}_N)^T$, $\mathbf{\Sigma}^* = (\mathbf{\Sigma}_1 \mathbf{\Sigma}_2 \cdots \mathbf{\Sigma}_Q)^T$, $\boldsymbol{\mu}^* = (\boldsymbol{\mu}_1 \boldsymbol{\mu}_2 \cdots \boldsymbol{\mu}_Q)$, and $\mathbf{y}^* = (\mathbf{y}_1^T \mathbf{y}_2^T \cdot\cdot \mathbf{y}_N^T)^T$, respectively. The proposed MPC method estimates $\mathbf{a}^*$, $\mathbf{\Sigma}^*$, and $\boldsymbol{\mu}^*$ by iterative calculation using the generalized method of moments (GMM) [11][12] so as to satisfy (15) and (16) for all pixel vectors in a multispectral image. The number of iterations k is given as a superscript in parentheses for each parameter.

GMM was proposed by Hansen (1982) [11]. Let, i denote the i-th pixel, $\boldsymbol{\mu}_i$ is mean pixel vector of i-th pixel calculated by (15), $\mathbf{m}^* = (\boldsymbol{\mu}_1^T \boldsymbol{\mu}_2^T \cdots \boldsymbol{\mu}_N^T)^T$, and $\boldsymbol{\theta} = [\mathbf{a}^*, \mathbf{\Sigma}^*, \boldsymbol{\mu}^*]$. The idea behind GMM is to choose so as to make sample moment $g(\boldsymbol{\theta}|\, \mathbf{y}^*)(= \mathbf{y}^* - \mathbf{m}^*)$ as to close as possible to population moment of zero; that is, the GMM estimator $\hat{\boldsymbol{\theta}}$ is the values of $\boldsymbol{\theta}$ that minimizes the scalar

$$Q_e(\boldsymbol{\theta} \mid \mathbf{y}^*) = [\mathbf{g}(\boldsymbol{\theta} \mid \mathbf{y}^*)]^T \mathbf{W} [\mathbf{g}(\boldsymbol{\theta} \mid \mathbf{y}^*)], \tag{19}$$

where $\mathbf{W}$ is a sequence of (N×P×N×P) positive definite weighting matrices. The optimal value for the weighting matrix $\mathbf{W}$ in (19) is given by $\mathbf{S}^{-1}$, the inverse of the asymptotic variance matrix. The minimum asymptotic variance for the GMM estimator $\hat{\boldsymbol{\theta}}$ is obtained when $\boldsymbol{\theta}$ is chosen to minimize

$$Q_e(\boldsymbol{\theta} \mid \mathbf{y}^*) = [\mathbf{g}(\boldsymbol{\theta} \mid \mathbf{y}^*)]^T \mathbf{S}^{-1} [\mathbf{g}(\boldsymbol{\theta} \mid \mathbf{y}^*)]. \tag{20}$$

$\mathbf{m}^*$ is calculated using (15) and the parameter $\hat{\mathbf{a}}^*$, and then $\mathbf{S}$ in (20) is estimated using $\hat{\boldsymbol{\theta}}$ as next equation for $\hat{\boldsymbol{\theta}}$ any consistent estimate of the true value $\boldsymbol{\theta}$.

$$\hat{\mathbf{S}} = [\mathbf{g}(\hat{\boldsymbol{\theta}} \mid \mathbf{y}^*)][\mathbf{g}(\hat{\boldsymbol{\theta}} \mid \mathbf{y}^*)]^T \xrightarrow{p} \mathbf{S}, \tag{21}$$

where the sign $\xrightarrow{p}$ denotes the convergence in probability.

The best estimates of $\mathbf{m}^*$ is obtained by Generalized Least Squares (GLS). The GLS estimate of $\boldsymbol{\mu}^*$ is the value that minimizes

$$(\mathbf{y}^* - \mathbf{m}^*)^T (\mathbf{\Omega}^*)^{-1} (\mathbf{y}^* - \mathbf{m}^*), \tag{22}$$

where

$$\mathbf{\Omega}^* = E[(\mathbf{y}^* - \mathbf{m}^*)(\mathbf{y}^* - \mathbf{m}^*)^T]. \tag{23}$$

The sign $\mathbf{\Omega}^*$ denotes the variance-covariance matrix of ($\mathbf{y}^*$ - $\mathbf{m}^*$). In this study, it is assumed that pixel vectors in image are statistically independent of each other. That is, it is also assumed that vector ($\mathbf{y}_i - \boldsymbol{\mu}_i$) for i-th pixel is uncorrelated with vectors for all other pixels. Therefore, $\mathbf{\Omega}^*$ is represented as a block diagonal matrix ($\mathbf{\Omega}^*$=*diag*($\mathbf{\Omega}_i$), i=1 to N) with dimensions of (N×P×N×P). The sign $\mathbf{\Omega}_i$ denotes variance-covariance matrix of i-th pixel vector. $\mathbf{\Omega}_i$ is obtained by $(\mathbf{y}_i - \boldsymbol{\mu}_i)(\mathbf{y}_i - \boldsymbol{\mu}_i)^T$. $\mathbf{\Omega}_i$ is also obtained calculating by (16). The symbol of *diag* implies that $\mathbf{\Omega}_i$ for i-th pixel is arranged as block diagonal components. Since $\hat{\mathbf{S}}$ is also calculated by *diag*($\mathbf{\Omega}_i$), $\hat{\mathbf{S}}$ in (21) and $\mathbf{\Omega}^*$ in (23) are calculated as the same matrix.

GMM estimates $\boldsymbol{\theta}$ so as to be minimum for the expected value of the sample moment $Q_e(\boldsymbol{\theta} \mid \mathbf{y}^*)$, which is given by $E[Q_e(\boldsymbol{\theta} \mid \mathbf{y}^*)]$. GMM is used to estimate the minimum $\boldsymbol{\theta}$ by following asymptotic variance Q_e.

$$Q_e(\boldsymbol{\theta} \mid \mathbf{y}^*) = \sum_{i=1}^{N} q_i(\boldsymbol{\theta} \mid \mathbf{y}_i) \tag{24}$$

$$q_i(\boldsymbol{\theta} \mid \mathbf{y}_i) = (\mathbf{y}_i - \boldsymbol{\mu}_i)^T \mathbf{\Omega}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) \tag{25}$$

$\mathbf{a}^{*(k+1)}$, $\Sigma^{*(k+1)}$, and $\boldsymbol{\mu}^{*(k+1)}$ are estimated from $\mathbf{a}^{*(k)}$, $\mathbf{\Sigma}^{*(k)}$, and $\boldsymbol{\mu}^{*(k)}$ as follows. First, $\mathbf{a}^*$ is estimated from $\mathbf{a}^{*(k)}$, $\mathbf{\Sigma}^{*(k)}$, and $\boldsymbol{\mu}^{*(k)}$ by using the method described in Section Ⅲ*B*.1, and estimated $\mathbf{a}^*$ is given by $\mathbf{a}^{*(k+1)}$. Then, $\mathbf{\Sigma}^*$ is estimated from $\mathbf{a}^{*(k+1)}$ and $\boldsymbol{\mu}^{*(k)}$ by using the method described in Section Ⅲ*B*.2, and estimated $\mathbf{\Sigma}^*$ is given by $\mathbf{\Sigma}^{*(k+1)}$. Next, $\boldsymbol{\mu}^*$ is estimated using $\mathbf{a}^{*(k+1)}$, $\mathbf{\Sigma}^{*(k+1)}$, and $\boldsymbol{\mu}^{*(k)}$ by using the method described in Section Ⅲ*B*.3, and estimated $\boldsymbol{\mu}^*$ is given by $\boldsymbol{\mu}^{*(k+1)}$. In this way, $\mathbf{a}^*$, $\mathbf{\Sigma}^*$, and $\boldsymbol{\mu}^*$ are estimated in turn until the following equation is satisfied.

$$\left| \frac{\theta^{(k+1)} - \theta^{(k)}}{\theta^{(k+1)}} \right| < 10^{-6}, \tag{26}$$

where θ is element of $\boldsymbol{\theta}$. The estimation methods for each of $\mathbf{a}^*$, $\mathbf{\Sigma}^*$, and $\boldsymbol{\mu}^*$ are described respectively in Section Ⅲ*B*.1, Section Ⅲ*B*.2, and Section Ⅲ*B*.3.

B.1 *Method of estimating* $\mathbf{a}^*$

Let i represent the i-th pixel, $\mathbf{a}^*$ is obtained by estimating $\mathbf{a}_i$ for all i. We estimate $\mathbf{a}_i$ by the following equation (27), which is obtained by introducing the Lagrange multiplier to impose constraints (3).

$$\hat{\mathbf{a}}_i = \left(\mathbf{I}_Q - \frac{\left(\boldsymbol{\mu}^{*T} \mathbf{\Omega}_i^{-1} \boldsymbol{\mu}^*\right)^{-1} \mathbf{1}\mathbf{1}^T}{\mathbf{1}^T \left(\boldsymbol{\mu}^{*T} \mathbf{\Omega}_i^{-1} \boldsymbol{\mu}^*\right)^{-1} \mathbf{1}} \right) \left(\boldsymbol{\mu}^{*T} \mathbf{\Omega}_i^{-1} \boldsymbol{\mu}^*\right)^{-1} \boldsymbol{\mu}^{*T} \mathbf{\Omega}_i^{-1} \mathbf{y}_i + \frac{\left(\boldsymbol{\mu}^{*T} \mathbf{\Omega}_i^{-1} \boldsymbol{\mu}^*\right)^{-1} \mathbf{1}}{\mathbf{1}^T \left(\boldsymbol{\mu}^{*T} \mathbf{\Omega}_i^{-1} \boldsymbol{\mu}^*\right)^{-1} \mathbf{1}}, \tag{27}$$

where $\mathbf{1}^T$=(1 1 …1) is the Q-variate vector, and $\mathbf{I}_Q$ is represented by a Q×Q unit matrix. When $\mathbf{a}^{*(k+1)}$ is obtained by (27), $\mathbf{\Omega}_i$ in (27) is calculated by using (16), $\mathbf{a}^{*(k)}$, and $\mathbf{\Sigma}^{*(k)}$. If estimated $\mathbf{a}^*$ is different from $\mathbf{a}^{*(k)}$, then $\mathbf{a}^*$ is estimated again using $\mathbf{\Omega}_i$ calculated from estimated $\mathbf{a}^*$. When the difference in $\mathbf{a}^*$ between before estimation $\mathbf{a}_i$ and after estimation $\hat{\mathbf{a}}_i$ is almost equal to $\mathbf{0}$, estimated $\mathbf{a}^*$ is given by $\mathbf{a}^{*(k+1)}$. Actually, this process is continued until following equation is satisfied.

$$\left\| \hat{\mathbf{a}}_i - \mathbf{a}_i \right\| < 10^{-8}, \tag{28}$$

where the sign $\|\mathbf{x}\|$ denotes the infinity norm of vector $\mathbf{x}$.

Note that it is necessary to pay attention to the multicollinearity of pixel values between bands. The problem faced by applied researchers when regressors are highly, although not perfectly, related included the following symptoms [13].

1) Small changes in the data produce wide swings in the parameter estimates.

2) Coefficients may have very high standard errors and low significance levels even though they are jointly significant and

R^2 for the regression is quite high.

3) Coefficients may have the "wrong" sign or implausible magnitude.

Since nonexperimental data will never be orthogonal, to some extent multicollinearity will be present. Several strategies have been proposed for finding and coping with multicollinearity. The obvious practical remedy (and surely the most frequently used) is to drop variables suspected of causing the problem from the regression.

In this study, $\mathbf{a}_i$ is estimated by referring 1) as follows. One of the methods used to avoid the effect of multicollinearity is to reduce the variates of $\mathbf{y}$, $\mathbf{\Omega}_i$, and $\boldsymbol{\mu}^*$ for certain bands since $\mathbf{a}_i$ can be estimated in the case where $P \geq Q$. So, in this study, the number of bands to estimate $\mathbf{a}_i$ is set to four. For all combinations of four bands in six bands, when a band is excluded from four bands, I adopt $\hat{\mathbf{a}}_i$ whose variation is small. Actually, estimated $\mathbf{a}_i$ is acquired as follows. For l-th combination of four bands, estimated $\mathbf{a}_i$ using four bands is written by $\hat{\mathbf{a}}_i^l$, and $\hat{\mathbf{a}}_i$ estimated by using three bands excluding j-th band in four bands is written by $\hat{\mathbf{a}}_{i-j}^l$. A band which is excluded from four bands is changed in turn, and $\hat{\mathbf{a}}_{i-j}^l$, j=1to 4, are acquired. When the value for l-th combination of four bands

$$q_l = \sum_{j=1}^{4} \left\| \hat{\mathbf{a}}_{i-j}^l - \hat{\mathbf{a}}_i^l \right\| \tag{29}$$

is smallest in any other combination of four bands, $\hat{\mathbf{a}}_i^l$ estimated from l-th combination of four bands is adopted as the estimated value of $\mathbf{a}_i$.

*B.*2 *Method of estimating* $\mathbf{\Sigma}^*$

A consistent estimator can be constructed by using the ordinary least-squares method. $\mathbf{\Omega}_i$ is given by

$$\mathbf{\Omega}_i = (\mathbf{y}_i - \boldsymbol{\mu}_i)(\mathbf{y}_i - \boldsymbol{\mu}_i)^T , \tag{30}$$

and we let $\mathbf{\Omega}^{**} = (\mathbf{\Omega}_1 \mathbf{\Omega}_2 \cdots \mathbf{\Omega}_N)^T$. The relationship between $\mathbf{\Omega}^{**}$ and $\mathbf{\Sigma}^*$ is given by

$$\mathbf{\Omega}^{**} = (\mathbf{a}^* \otimes \mathbf{I}_P)\mathbf{\Sigma}^* , \tag{31}$$

where the $\otimes$ symbol denotes the Kronecker product, and $\mathbf{I}_P$ is a $P \times P$ dimensional unit matrix. The least-squares estimator of $\mathbf{\Sigma}^*$ is given by

$$\hat{\mathbf{\Sigma}}^* = \left\{ \left(\mathbf{a}^{*T} \mathbf{a}^* \right)^{-1} \mathbf{a}^{*T} \otimes \mathbf{I}_p \right\} \mathbf{\Omega}^{**} . \tag{32}$$

Although it is necessary that $\mathbf{\Sigma}_q$ for all $q \in \{1, \cdots, Q\}$ are positive definite, $\mathbf{\Sigma}_q$ obtained from estimated $\hat{\mathbf{\Sigma}}^*$, which is positive definite, is not guaranteed. Therefore, it is necessary to confirm whether $\mathbf{\Sigma}_q$ for all $q \in \{1, \cdots, Q\}$ are positive definite.

In this study, we estimate $\mathbf{\Sigma}_q$ by following method as one of the method to overcome this problem. It is assumed that actual $\mathbf{\Sigma}_q$ is existed to the direction from $\mathbf{\Sigma}_q^{(k)}$ to $\mathbf{\Sigma}_q^{(k+1)}$. Estimated $\mathbf{\Sigma}_q^{(k+1)}$ is given by

$$\mathbf{\Sigma}_q^{(k+1)'} = (\mathbf{\Sigma}_q^{(k+1)} - \mathbf{\Sigma}_q^{(k)}) \cdot r + \mathbf{\Sigma}_q^{(k)} , \tag{33}$$

where r in (33) is determined that $\mathbf{\Sigma}_q^{(k+1)'}$ for all q are positive definite and maximum value of r which is ranged from 0 to 1.

*B.*3 *Method of estimating* $\boldsymbol{\mu}^*$

We describe the problem with a multivariate linear model as

$$\mathbf{y}^* = \left(\mathbf{a}^* \otimes \mathbf{I}_p \right) \mathbf{m}^* + \boldsymbol{\varepsilon} , \tag{34}$$

where $\mathbf{m}^*$ is $(\boldsymbol{\mu}_1^T \boldsymbol{\mu}_2^T \cdots \boldsymbol{\mu}_Q^T)^T$, and $\boldsymbol{\varepsilon}$ is the $(N \times P) \times 1$-variate error vector corresponding to the expected value of $\mathbf{0}$ and the variance-covariance matrix of $\mathbf{\Omega}^*$. $\mathbf{\Omega}^*$ is represented as a block diagonal matrix ($\mathbf{\Omega}^* = diag(\mathbf{\Omega}_i)$, i=1 to N) with dimensions of $(N \times P) \times (N \times P)$. Using (34), the weighted least-squares estimator $\hat{\mathbf{m}}^*$ of $\mathbf{m}^*$ is given by

$$\hat{\mathbf{m}}^* = \left\{ \left(\mathbf{a}^* \otimes \mathbf{I}_p \right)^T (\mathbf{\Omega}^*)^{-1} \left(\mathbf{a}^* \otimes \mathbf{I}_p \right) \right\} \left(\mathbf{a}^* \otimes \mathbf{I}_p \right)^T (\mathbf{\Omega}^*)^{-1} \mathbf{y}^* . \tag{35}$$

Estimated $\boldsymbol{\mu}^*$, which is denoted by $\hat{\boldsymbol{\mu}}^*$, is obtained by arranging $\hat{\mathbf{m}}^*$.

C. Evaluation

Evaluation of estimation is conducted by using χ^2 statistics as follows. We assume that the empirical moments ($\mathbf{y}_i$ - $\boldsymbol{\mu}_i$) obey a central limit theorem [13]. This assumes that the moments have a finite asymptotic variance-covariance matrix $\mathbf{\Omega}_i$, so that ($\mathbf{y}_i$ - $\boldsymbol{\mu}_i$) is independent vectors distributed normally with mean $\mathbf{0}$ and variance $\mathbf{\Omega}_i$. That is,

$$g(\mathbf{a}_i; \mathbf{y}_i) = (\mathbf{y}_i - \boldsymbol{\mu}_i) \xrightarrow{d} N(\mathbf{0}, \mathbf{\Omega}_i) . \tag{36}$$

Equation (36) is rewritten by

$$\mathbf{z}_i = (\mathbf{y}_i - \boldsymbol{\mu}_i)\mathbf{\Omega}_i^{-1/2} \xrightarrow{d} N(\mathbf{0}, \mathbf{1}) \tag{37}$$

Furthermore, $\mathbf{z}^T\mathbf{z}$ is given by

$$\mathbf{z}_i^T \mathbf{z}_i = (\mathbf{y}_i - \boldsymbol{\mu}_i)\mathbf{\Omega}^{-1}(\mathbf{y}_i - \boldsymbol{\mu}_i) . \tag{38}$$

Since $\mathbf{z}_i \xrightarrow{d} N(\mathbf{0}, \mathbf{1})$,

$$\mathbf{z}_i^T \mathbf{z}_i \sim \chi^2(P) . \tag{39}$$

From equation (24),

$$Q_e = \sum_{i=1}^{N} \mathbf{z}_i^T \mathbf{z}_i \sim \chi^2(\sum_{i=1}^{N} P) = \chi^2(N \times P) \tag{40}$$

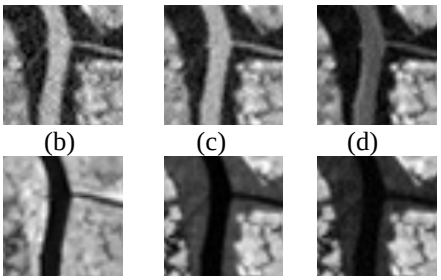
So, the estimation result is evaluated by χ^2 value. Q_e might be expected to follow the χ^2 distribution with $N \times P$ degrees of freedom. Whether or not estimators are optimal is judged on the basis of Q_e.When $N \times P$ is enough large such as more than 50, χ^2 value that the probability of observing is 50% is nearly equal to $N \times P$.

IV. Application and Result

The proposed mixture model and MPC method are applied to a multispectral image data. In this study, a multispectral image taken by the Enhanced Thematic Mapper Plus (ETM^+) onboard Landsat-7 is analyzed. The ETM^+ sensor can detect electromagnetic radiation in eight bands. In this study, six bands (P=6) are used (bands 6 and 8 are excluded because their spatial resolution is different from the other bands). An ETM^+ image included in a target area and images of the target area for the 6 bands used here are shown in Fig.2. These images were taken in Burma (Myanmar) on 21 January, 2003. The target area of the

application is the area enclosed by a white line in Fig. 2(a). This area is assumed to be covered by three components (Q=3) derived from visual inspection. The number of pixels N in the target area is 1600 pixels (=40×40 pixels). The values of the band spectra for each pixel in the image are represented by digital numbers between 0 and 255.

(a) Band 1 image included in target area



(b) (c) (d)

(e) (f) (g)

Fig.2 Landsat-7 ETM+ image taken in Burma (Myanmar) on 21 January, 2003.

Table 1. Calculated initial values of $\boldsymbol{\mu}_q$ and $\boldsymbol{\Sigma}_q$ for all $q \in \{1, 2, 3\}$.

	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
$\boldsymbol{\mu}_1$	38.44	29.88	22.08	110.96	42.92	20.24
$\boldsymbol{\Sigma}_1$	0.73	-0.07	-0.36	0.34	0.20	-0.03
	-0.07	1.31	0.37	2.12	0.15	-0.25
	-0.36	0.37	1.03	-0.32	-0.27	0.06
	0.34	2.12	-0.32	16.20	4.00	-0.19
	0.20	0.15	-0.27	4.00	2.55	0.58
	-0.03	-0.25	0.06	-0.19	0.58	0.90
$\boldsymbol{\mu}_2$	50.60	44.44	53.68	92.68	122.04	77.44
$\boldsymbol{\Sigma}_2$	2.16	1.38	2.47	2.43	5.46	3.22
	1.38	2.09	2.82	2.50	5.38	3.69
	2.47	2.82	5.58	3.94	8.93	6.66
	2.43	2.50	3.94	8.86	6.61	4.66
	5.46	5.38	8.93	6.61	25.00	16.70
	3.22	3.69	6.66	4.66	16.70	13.69
$\boldsymbol{\mu}_3$	48.88	39.96	32.60	21.76	11.84	10.28
$\boldsymbol{\Sigma}_3$	0.75	0.04	-0.05	0.21	-0.10	-0.29
	0.04	0.60	-0.14	0.39	0.43	0.25
	-0.05	-0.14	1.04	0.46	0.34	-0.01
	0.21	0.39	0.46	2.82	0.64	-0.29
	-0.10	0.43	0.34	0.64	1.73	0.28
	-0.29	0.25	-0.01	-0.29	0.28	1.72

The initial values of mean vector $\boldsymbol{\mu}_q$ and variance-covariance $\boldsymbol{\Sigma}_q$ for all $q \in \{1, 2, 3\}$ are calculated from pixel vectors in Site 1, Site 2, and Site 3 in Fig. 2 (a). The number of pixels for each site is 25. The calculated initial values of $\boldsymbol{\mu}_q$ and $\boldsymbol{\Sigma}_q$ for all $q \in \{1, 2, 3\}$ are shown in Table 1.

It is difficult to evaluate the estimated $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$, since actual values of $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$ are unknown. Therefore, first, the proposed MPC method in this study is applied to simulated image data (Section IV*A*). Then, the proposed MPC method is applied to actual Landsat-7 ETM$^+$ image data (Section IV*B*).

A. Application to simulated image data

First, a method of generating simulated image data is presented in Section IV*A*.1, and then application results are shown in Section IV*A*.2.

A.1 *Generation of simulated image data*

The number of pixels N is set to 1600. The number of components in a simulated image is three (Q=3); thus, the number of components for each single pixel is less than three. Actual $\boldsymbol{\mu}_q$ and $\boldsymbol{\Sigma}_q$ for all $q \in \{1, 2, 3\}$ are set to the values given in Table 1. $\mathbf{y}^*$ is set by generating N of $\mathbf{y}$, and the pixel vector of the i-th pixel $\mathbf{y}_i$ is generated from $\boldsymbol{\Sigma}_q$ and $\boldsymbol{\mu}_q$ given in Table 1 by using (8) in the following way. The number of micro-pixels K in a single pixel is set to 10,000. The mixing ratio $\mathbf{a}_i$ for the i-th pixel is generated by a Dirichlet random number generator [14]; $\mathbf{a}^*$ with $N \times Q$ dimensional matrix is set from the generated N of a. The histograms of the generated mixing ratios for each component are shown in Fig. 3. K_q in (8) is obtained from $a_q \times K$. In this study, it is assumed that $\mathbf{x}_{qk}$ in (8) follows the P-dimensional normal distribution. The component $\mathbf{x}_{kq}$ in (8) is generated by a P-dimensional normal random number generator with the expected value

$$E\left[\mathbf{x}_{kq}\right] = \frac{\boldsymbol{\mu}_q^*}{K}, \tag{41}$$

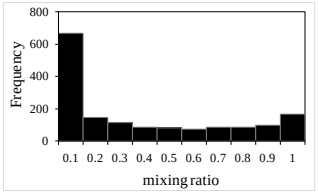
and the variance-covariance matrix

$$Var\left[\mathbf{x}_{kq}\right] = \frac{\boldsymbol{\Sigma}_q^*}{K}. \tag{42}$$

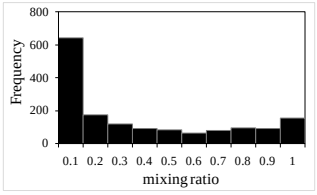
The component $\mathbf{x}_q$ in (8) is given by

$$\mathbf{x}_q = \frac{1}{K}\left(\boldsymbol{\mu}_q + \boldsymbol{\Sigma}_q^{1/2} \cdot N(\mathbf{0}, \mathbf{I}_p)\right), \tag{43}$$

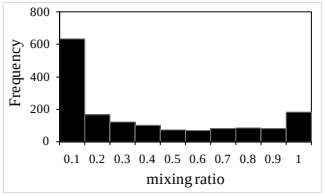
where $N(\mathbf{0}, \mathbf{I}_p)$ is a P-variate normal random number generator. P-variate simulated image data $\mathbf{y}^*$ is generated from N of $\mathbf{y}$ using different values generated from $N(\mathbf{0}, \mathbf{I}_p)$. Q_e calculated from (24) using $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$ is 9714, which is about 29% of values drawn from the normal distribution away from $N \times P$ (=9,600). And, when mixture model is used linear mixture model and finite mixture model, Q_e calculated using $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$ are 16927.77 and 7041.42, respectively.



(a) Component 1



(b) Component 2



(c) Component 3

Fig.3 Frequencies of mixing ratios.

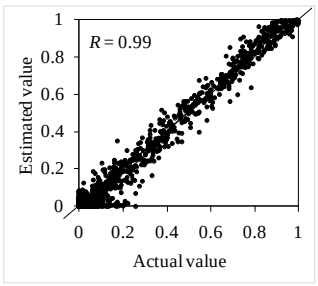
A.2 Application results

Application results are detailed. $\boldsymbol{\mu}^{*(0)}$ and $\boldsymbol{\Sigma}^{*(0)}$ are obtained as follows. We assume a single pixel to be a pure pixel of component q whose aq is larger than 0.8. From these pixels, $\boldsymbol{\mu}^{*(0)}$ and $\boldsymbol{\Sigma}^{*(0)}$ are calculated, and are shown in Table 2.

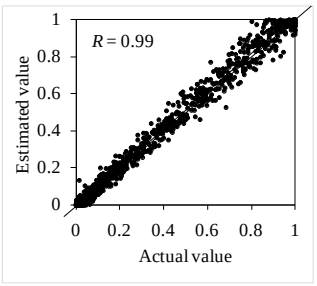
The proposed MPC method described in Section Ⅲ*B* is applied to simulated image data. $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$ are estimated, and simulated image data is modeled. The algorithm converges after 114 iterations. The scatter plots between actual mixing ratios and estimated mixing ratios are shown in Fig. 4. Plots are scattered along the straight line with zero intercept and unit slope. The correlation coefficients *R* for all components are almost 1. Estimated expected values and variance-covariance for all components are shown in Table 3. Almost all of the components of estimated $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$ are closer to the true values given in Table 1 than the initial values $\boldsymbol{\mu}^{*(0)}$ and $\boldsymbol{\Sigma}^{*(0)}$ given in Table 2. Q_e calculated from estimated $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$ is 9878.45, which is close to the actual value of Q_e. Therefore, simulated image data is in agreement with the model constructed by $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$.

Table 2 Initial values of mean vector $\boldsymbol{\mu}_q$ and variance-covariance $\boldsymbol{\Sigma}_q$ for all $q \in \{1, 2, 3\}$.

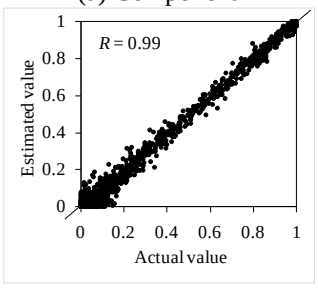
	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
$\boldsymbol{\mu}_1$	39.36	30.44	23.63	105.93	44.73	22.03
$\boldsymbol{\Sigma}_1$	1.18	0.41	0.73	-1.59	1.86	1.39
	0.41	1.79	1.66	-0.37	2.39	1.77
	0.73	1.66	4.08	-2.79	6.36	5.13
	-1.59	-0.37	-2.79	35.00	11.72	2.68
	1.86	2.39	6.36	11.72	26.00	16.04
	1.39	1.77	5.13	2.68	16.04	11.56
$\boldsymbol{\mu}_2$	49.98	43.34	51.26	89.67	114.04	71.93
$\boldsymbol{\Sigma}_2$	2.71	1.88	3.45	2.10	7.51	4.53
	1.88	2.55	3.81	2.53	7.54	4.94
	3.45	3.81	8.25	6.46	16.98	11.67
	2.10	2.53	6.46	24.97	21.78	12.68
	7.51	7.54	16.98	21.78	57.06	35.93
	4.53	4.94	11.67	12.68	35.93	25.46
$\boldsymbol{\mu}_3$	48.52	39.55	33.02	26.95	17.03	12.96
$\boldsymbol{\Sigma}_3$	1.12	0.36	0.26	-1.28	-0.12	-0.12
	0.36	0.84	0.35	-0.71	0.91	0.74
	0.26	0.35	2.34	1.92	5.32	3.31
	-1.28	-0.71	1.92	25.97	20.20	10.11
	-0.12	0.91	5.32	20.20	30.12	17.31
	-0.12	0.74	3.31	10.11	17.31	12.22



(a) Component 1



(b) Component 2



(c) Component 3

Fig.4 Scatter plots of mixing ratios for three components showing the relationship between actual values and estimated values.

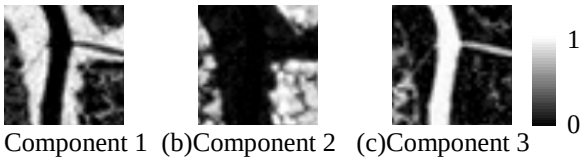
Table 3 Estimated mean vector $\boldsymbol{\mu}_q$ and variance-covariance $\boldsymbol{\Sigma}_q$ for all $q \in \{1, 2, 3\}$.

	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
$\boldsymbol{\mu}_1$	38.14	29.21	21.52	112.21	43.18	20.04
$\boldsymbol{\Sigma}_1$	0.87	0.00	-0.36	-0.37	0.16	-0.09
	0.00	1.34	0.28	1.45	0.18	-0.25
	-0.36	0.28	0.76	-0.01	-0.07	-0.02
	-0.37	1.45	-0.01	11.49	0.93	-0.72
	0.16	0.18	-0.07	0.93	0.63	-0.02
	-0.09	-0.25	-0.02	-0.72	-0.02	0.40
$\boldsymbol{\mu}_2$	51.03	44.68	54.58	90.79	124.16	78.79
$\boldsymbol{\Sigma}_2$	1.11	0.29	0.32	1.64	0.55	-0.37
	0.29	0.82	0.45	1.24	-0.02	-0.35
	0.32	0.45	1.30	1.51	-0.49	-0.42
	1.64	1.24	1.51	24.48	0.48	-2.04
	0.55	-0.02	-0.49	0.48	5.02	1.25
	-0.37	-0.35	-0.42	-2.04	1.25	2.52
$\boldsymbol{\mu}_3$	48.79	39.63	32.44	20.93	10.97	9.52
$\boldsymbol{\Sigma}_3$	0.78	0.05	-0.03	0.21	-0.04	-0.27
	0.05	0.56	-0.20	-0.21	-0.10	-0.07
	-0.03	-0.20	1.02	-0.23	-0.01	-0.23
	0.21	-0.21	-0.23	1.23	-0.23	-0.29
	-0.04	-0.10	-0.01	-0.23	1.71	0.31
	-0.27	-0.07	-0.23	-0.29	0.31	1.69

B. Application to Landsat ETM$^+$ image data

To estimate $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$, iterative calculation is performed. The values shown in Table 1 are used as the initial values of $\boldsymbol{\Sigma}^*$ and $\boldsymbol{\mu}^*$. The initial value of $\mathbf{a}_i$, which is $\mathbf{a}_i^{(0)}$, is set to $1/Q$ for all elements of $\mathbf{a}_i$. As a result, the number of iterations is 145; estimated $\boldsymbol{\Sigma}^*$ and $\boldsymbol{\mu}^*$ for all $q \in \{1, 2, 3\}$ are shown in Table 4. Estimated mixing ratios for three components are shown in Fig. 5. In Fig. 5, the larger the mixing ratio, the lighter the pixel looks.

Q_e calculated from (24) using estimated $\mathbf{a}^*$, $\boldsymbol{\Sigma}^*$, and $\boldsymbol{\mu}^*$ is 10500.40 which is slightly larger than actual value. Though Q is assumed three, it is considered that other components except for three components are included in some pixels. Pixels in Fig. 5 whose mixing ratio of one component is high are consistent with pixels in Fig. 2 which are assumed to have the same component by visual inspection



(a) Component 1 (b)Component 2 (c)Component 3

Fig.5 Estimated mixing ratios for three components.

Table 4 Estimated results of mean vector $\boldsymbol{\mu}_q$ and variance-covariance $\boldsymbol{\Sigma}_q$ for $q \in \{1, 2, 3\}$.

	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
$\boldsymbol{\mu}_1$	37.59	26.64	19.80	89.17	42.02	20.52
$\boldsymbol{\Sigma}_1$	0.63	-0.20	-0.16	0.01	0.27	-0.13
	-0.20	0.42	-0.15	1.00	0.09	-0.08
	-0.16	-0.15	0.51	-0.61	-0.03	-0.02
	0.01	1.00	-0.61	72.59	-1.48	-7.13
	0.27	0.09	-0.03	-1.48	4.50	1.40
	-0.13	-0.08	-0.02	-7.13	1.40	2.82
$\boldsymbol{\mu}_2$	48.68	41.69	52.32	81.50	124.86	83.05
$\boldsymbol{\Sigma}_2$	1.40	1.00	1.33	4.26	-0.67	-0.55
	1.00	2.56	2.94	6.92	-1.31	-2.23
	1.33	2.94	6.18	10.21	-3.11	-4.04
	4.26	6.92	10.21	54.89	-2.40	-13.15
	-0.67	-1.31	-3.11	-2.40	22.13	14.43
	-0.55	-2.23	-4.04	-13.15	14.43	19.59
$\boldsymbol{\mu}_3$	47.30	39.51	34.54	20.83	9.45	9.66
$\boldsymbol{\Sigma}_3$	1.43	0.34	-0.82	-1.00	0.88	-0.03
	0.34	0.80	-0.40	-1.09	0.68	0.23
	-0.82	-0.40	3.61	-1.15	-0.01	0.47
	-1.00	-1.09	-1.15	5.94	-0.01	0.19
	0.88	0.68	-0.01	-0.01	2.91	0.48
	-0.03	0.23	0.47	0.19	0.48	1.93

V. Conclusion

In this paper, a new mixture model for describing a mixture state in a single pixel of a multispectral image data has been proposed. The proposed mixture model is established on the basis of the process of mixed pixel generation in an actual multispectral image. The variance-covariance structure of a pixel vector which is derived from the proposed mixture model is different from the variance-covariance derived from the finite mixture model and the linear mixture model. The proposed mixture model does not specify the probability distribution of single-component data.

The proposed mixture model and MPC method were applied to simulated image data, first. Estimated parameters close to actual values were obtained. Next, we applied the proposed mixture model and MPC method to real Landsat-7 ETM^+ image data. As a result, it was found that pixels in the ETM^+ image whose mixing ratio of one component was high were consistent with pixels in an image which were assumed to have the same component.

The proposed technique is applied under the assumption that the number of categories within a single pixel is known a priori. It is therefore necessary to develop a method to determine the number of categories. A method to achieve this, which is based on the mixture model proposed in this paper, will be proposed at a later date.

References

[1] Chang C., Chiang S., Smith J. A., and Ginsberg I. W., "Linear spectral random mixture analysis for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens*, vol. 40, no.2, 2002, pp. 375-392.

[2] Santago P. and Gage H. D., "Quantification of MR brain images by mixture density and partial volume modeling," *IEEE Transactions on Medical imaging*, vol.12, no.3, 1993, pp.566-574.

[3] Y. E. Shimabukuro and J. A. Smith, "The least-squares mixing models to generate fraction images derived from remote sensing multispectral data," *IEEE, Trans. Geosci and Remote Sens.*, vol. 29, no. 1, 1991, pp. 16-20.

[4] Choi H. S., Haynor D. R., and Kim Y., "Partial volume tissue classification of multichannel magnetic resonance images-a mixel model," *IEEE Transactions on Medical imaging*, vol.10, no.3, 1991, pp.395-407.

[5] Hasan R., Mohammad R. S., "Sub-pixel classification of MODIS images," *Proceedings of the 6th WSEAS International Conference on Non-Linear Analysis Non-linear Systems and Chaos*, France, 2007, pp.103-108.

[6] Hasan R., Rahmatolah F., Majid R., "Multi temporal disaggregation of MODIS images using non-linear analysis," *5th WSEAS Int. Conf. on ENVIRONMENT, ECOSYSTEMS and DEVELOPMENT*, Spain, 2007, pp.222-227.

[7] Radu Muthac, Marc M. Van Hulle, "Statistics of future extraction by topographic independent component analysis from nature images," *Control & Signal Processing and 2002 WSEAS Int. Conf. on e-activities*, 2002.

[8] P. Bosdogianni, M. Petrou, and H. Kittler, "Mixture models with higher order moments," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 2, pp.341-353, 1997.

[9] Byung Eun Lee, Thanh Binh Nguyen, Sun Tae Chung, "An efficient cast shadow removal for motion segmentation," *Proceedings of the 9th WSEAS International Conference on SIGNAL PROCESSING, COMPUTATIONAL GEOmetry and ARTIFICIAL VISION*, 2009, pp. 83-87.

[10] Y. Kobayashi and Minoru Inamura, "An influence due to mixel in category classification using EM algorithm and its improvement," *Inform. Process. Soc. of Japan*, vol.37, no.1, 1996 (in Japanese).

[11] Hansen L. P., "Large sample properties of generalized method of moment estimators," *Econometrica*, vol.50, no.4, 1982, pp.1029-1054.

[12] Hamilton J. D., "*Time Series Analysis*," Promceton University Press, 1994.

[13] William H. Greene, *Econometric analysis sixth edition*, Princeton UNIV Pr, 2007.

[14] M. Evans, N. Hasings, and B. Peacock, *Statistical Distributions second edition*, New York: A Wiley-Interscience Publication, 1993.G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.

Mitsuhiro Tomosada was born in Japan on October 6, 1969. He graduated from Tsukuba University. Now, he belongs to central research institute of electric power industry. His employment experience included researcher at the Tsukuba University and at the statistical mathematics. His special fields of interest include measurement, applied statistics, satellite remote sensing.

Hiroe Tsubaki is a professor and the director of newly established risk research center in the Institute of Statistical Mathematics. Professor Tsubaki

was formerly a lecturer of statistics at Keio University from 1987 to 1997. Before teaching at Keio he taught in the Department of Mathematical Engineering at the University of Tokyo from 1982 to 1987. He earned his Bachelor, Master and Doctor of Engineering from the University of Tokyo in 1979, 1982 and 1988, respectively. He is also an expert member of several Japanese government committees as Japanese Industrial Standard Committee and Pharmaceutical Affairs and Food Sanitation Council.