Design of a Hard Expectation-Maximization-Based Normalized Matched Filter (EM-NMF) for the Detection of Chemical Warfare Agents Under Background Contamination

Hyeong-Geun Yu, *Student Member, IEEE*, Jai-Hoon Lee, Dong-Jo Park, *Member, IEEE*, Dong Eui Chang, and Hyunwoo Nam

***Abstract*—We propose a hard expectation-maximization-based normalized matched filter (EM-NMF) for the detection of chemical warfare agent (CWA) clouds under background contamination. The NMF, which is one of the most powerful detectors, requires background statistics calculated from a background training dataset. However, in practice, because the training dataset is likely to contain CWA-on background pixels, the performance of the NMF is severely degraded. This phenomenon is referred to as background contamination. To address this issue, we propose an algorithm that estimates the posterior probability of each pixel belonging to either the background or the CWA class. The optimal posterior probabilities are obtained by maximizing the log-likelihood of a contaminated dataset using the EM algorithm. Based on the posterior probability, we extract CWA-free background pixels from the contaminated dataset and design a hard EM-NMF with extracted CWA-free background pixels. We demonstrate that the proposed algorithm is an effective solution for background contamination, via experimental results conducted with actual CWA data measured by a Bruker HI–90 instrument in an outdoor setting as well as synthetic CWA data.**

***Index Terms*—Chemical warfare agent (CWA) detection, expectation maximization, hyperspectral imaging system, normalized matched filter.**

I. Introduction

THE hyperspectral imaging system (HIS) has emerged as a key technology for the detection of chemical warfare agent (CWA) clouds [1]–[18]. A passive HIS sensor is composed of a spectrometer and a focal plane array detector. The passive HIS sensor can measure the radiance spectrum in each pixel for the instantaneous corresponding field of view (IFOV) at a standoff distance without any additional light sources and generate cube data with a spatial resolution of $m \times n$ pixels and a spectral resolution of p channels. Therefore, the HIS sensor is able to detect and visualize the CWA cloud in the atmosphere from a distance. Many CWA cloud detection algorithms have been proposed based on probabilistic theories [2]–[7], the spectral unmixing [8], [9], or machine learning techniques [10]–[12]. Among these algorithms, a normalized matched filter (NMF) algorithm, which is also referred to as the adaptive cosine estimator, is a uniformly most powerful invariant test, which has the best detection probability out of all tests for a given false alarm probability [13], [14].

On the other hand, estimates of background statistics are required to construct an NMF. The background statistics are calculated from a background training dataset that consists of CWA-free background pixels measured before the CWA cloud occurs. Since it is impossible always to have information about the background in advance, the training dataset should be composed of unknown measured pixels. Then, it is likely to contain CWA-on background pixels. If the training dataset contains CWA-on background pixels, the background statistics are distorted by CWA-on background pixels. This phenomenon is referred to as *background contamination*. Distorted background statistics degrade the detection performance of the matched filter (MF) [15]–[17]. It is obvious that the detection performance of the NMF also deteriorates due to background contamination because both the NMF and the MF are designed in a similar way using background statistics.

To mitigate the issue of background contamination, Niu *et al.* [17] proposed an algorithm, which alleviates the distortion for the covariance matrix by adding a diagonal matrix to the contaminated background covariance matrix. However, since the mean vector of the background is still contaminated, there is a limit to eliminating background contamination. Kim *et al.* [18] proposed an iterative filtering algorithm, which separates CWA-free background pixels by applying the MF iteratively. However, to resolve the background contamination problem, the iterative filtering algorithm requires information about the degree of contamination, which indicates a ratio of CWA-on background pixels in the contaminated dataset. Since it is difficult to predict

Manuscript received February 13, 2020; revised May 3, 2020; accepted May 19, 2020. Date of publication May 29, 2020; date of current version June 16, 2020. This work was supported by the Agency for Defense Development of the South Korea. *(Corresponding author: Dong-Jo Park.)*

Hyeong-Geun Yu, Jai-Hoon Lee, Dong-Jo Park, and Dong Eui Chang are with the School of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: elloss@kaist.ac.kr; leejhwogns0313@gmail.com; djpark@kaist.ac.kr; dechang@kaist.ac.kr).

Hyunwoo Nam is with the CRB Defense Technology Directorate, Agency for Defense Development, Daejeon 31486, South Korea (e-mail: hyunwoonam@add.re.kr).

Digital Object Identifier 10.1109/JSTARS.2020.2998451

the degree of contamination in advance, the iterative filtering algorithm is not practical. To extract CWA-free background pixels from the contaminated background dataset without information about the degree of contamination, we focus on the posterior probability of each pixel belonging to either a CWA-free background pixel group or a CWA-on background pixel group.

In this article, we propose an algorithm that estimates the posterior probability of each pixel and extracts CWA-free background pixels from a contaminated dataset based on the posterior probability. The optimal posterior probability of each pixel is attained by maximizing the log-likelihood probability of the contaminated dataset with the expectation and maximization (EM) algorithm. The likelihood probability of the contaminated dataset can be modeled as a product of the likelihood probabilities of all pixels in the contaminated dataset. For the fast and accurate convergence of the EM algorithm, we propose an initialization method, which is based on rough classification using a robust NMF. Then, we extract CWA-free background pixels from the contaminated dataset by thresholding the posterior probability of each pixel and calculate undistorted background statistics, which can be used to construct a hard EM-NMF.

We conduct experiments with synthetic CWA and actual CWA data measured by a Bruker hyperspectral imaging system (HI-90, Bruker Corporation, Germany) in outdoor settings. The synthetic CWA data are generated by embedding CWA signals into actual background data measured by the HI–90. The actual CWA data are measured in a scenario where SF_6 gas is sprayed outdoors. The experimental results show that the proposed EM algorithm correctly calculates the posterior probabilities without information about the degree of contamination, which results in the robust detection performance of the hard EM-NMF under background contamination, and the proposed initialization method provides good convergence characteristics for the EM algorithm.

The remainder of this article is organized as follows. In Section II, the NMF is summarized. We also explain that the detection performance of the NMF deteriorates due to background contamination. We introduce related works to address background contamination in Section III. In Section IV, we present the proposed algorithm that finds the optimal posterior probability of each pixel using an EM algorithm, extract CWA-free background pixels from the contaminated dataset, and construct a hard EM-NMF with extracted CWA-free background pixels. We also discuss a suitable initialization method for the EM algorithm. In Section V, experimental results are assessed with synthetic data and actual data and demonstrate that the hard EM-NMF is robust against background contamination. The final conclusions are drawn in Section VI.

II. Normalized Matched Filter

A. NMF for Remote CWA Cloud Detection

To describe the basic characteristics of the spectra measured by a passive HIS sensor, a three-layer model is widely used [1]. The three-layer model is composed of an atmosphere layer, a CWA cloud layer, and a background layer, as depicted in Fig. 1. From the three-layer model, light radiated from the background reaches an HIS sensor through the CWA cloud and the atmosphere. To express the light measured by the HIS sensor simply, four assumptions are required, as follows. First, the atmosphere and CWA cloud are free of aerosols and scattering by aerosols can be ignored throughout, second, the CWA cloud and atmosphere can be homogeneous in terms of both the temperature and composition, third, the thickness of the CWA cloud layer and the distance between the CWA cloud and background are short such that atmospheric transmittance can be neglected in those layers, and finally reflections between all layers can be ignored.

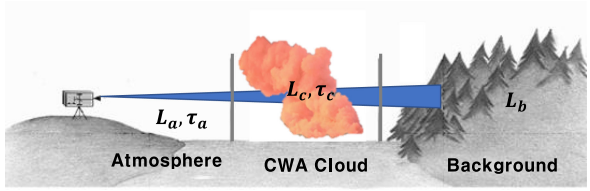

Fig. 1. Three-layer model of the radiance spectra measured by a passive HIS sensor.

The measured radiance spectra in the absence of the CWA cloud $L_{\text{free}}(\nu)$ and that in the presence of the CWA cloud $L_{\text{ON}}(\nu)$ are then correspondingly represented as follows:

$$L_{\text{free}}(\nu) = L_a(\nu) + \tau_a(\nu) L_b(\nu) + n(\nu) \tag{1}$$

$$L_{\text{ON}}(\nu) = L_a(\nu) + \tau_a(\nu)\tau_c(\nu) L_b(\nu) + \tau_a(\nu)\left[1 - \tau_c(\nu)\right] B(\nu, T_c) + n(\nu) \tag{2}$$

where ν is the wavenumber, $L_a(\nu)$ is the atmospheric path radiance, $L_b(\nu)$ is the background radiance, $B(\nu, T_c)$ is the blackbody radiance at temperature T_c, $\tau_a(\nu)$ is the atmospheric transmittance, $\tau_c(\nu)$ is the transmittance of the CWA cloud, and $n(\nu)$ is a noise signal caused by the HIS sensor. Subsequently, $L_{\text{ON}}(\nu)$ is expressed using $L_{\text{free}}(\nu)$ as

$$L_{\text{ON}}(\nu) = L_{\text{free}}(\nu) + \tau_a(\nu)\left[1 - \tau_c(\nu)\right]\left[B(\nu, T_c) - L_b(\nu)\right]. \tag{3}$$

From Beer's law [19], $\tau_c(\nu)$ is obtained as

$$\tau_c(\nu) = \exp\left(-\sum_{r=1}^{N_r} \gamma_r \alpha_r(\nu)\right) \tag{4}$$

where N_r denotes the number of CWAs in the cloud, $\alpha_r(\nu)$ stands for the rth CWA-absorption-coefficient spectrum and γ_r represents the concentration path length for the rth CWA. The term $\gamma_r \alpha_r(\nu)$ is known as the optical depth.

According to (3), $L_{\text{ON}}(\nu)$ is a nonlinear function due to the two terms $1 - \tau_c(\nu)$ and $B(\nu, T_c) - L_b(\nu)$. Applying two reasonable approximations, we approximate (3) with a linear function [1]. First, using Taylor's approximation, we approximate $1 - \tau_c(\nu)$ as follows:

$$1 - \tau_c(\nu) = 1 - \exp\left(\sum_{r=1}^{N_r} -\gamma_r \alpha_r(\nu)\right) \simeq \sum_{r=1}^{N_r} \gamma_r \alpha_r(\nu). \tag{5}$$

The second approximation is the linearization of the term $B(\nu, T_c) - L_b(\nu)$. Many background emissivity functions $\epsilon_b(\nu)$, which do not have sharp spectral peaks, are close to a

uniform function whose value is near to 1. Then, it leads to the approximation $\epsilon_b(\nu) \approx 1$. Assuming $\epsilon_b(\nu) \approx 1$, the background radiance $L_b(\nu)$ is almost same as the blackbody radiance $B(\nu, T_b)$ at the background temperature T_b. If the difference between the background temperature T_b and the CWA cloud temperature T_c is less than $10\,K$, i.e., $\Delta T = |T_c - T_b| < 10\,K$, we can approximate the blackbody radiance difference $B(\nu, T_c) - B(\nu, T_b)$ as a linear function for the temperature difference ΔT by applying the linear approximation of the Planck function. Therefore, $B(\nu, T_c) - L_b(\nu)$ can be represented as follows:

$$B(\nu, T_c) - L_b(\nu) \simeq C_B \Delta T \tag{6}$$

where $C_B = 5.5 \times 10^{-4}$ $\mathrm{W \cdot m^{-2} \cdot str^{-1}/cm^{-1}/K}$ represents a constant, which is independent of wavenumber and the temperature. Adopting (5) and (6), (3) can be linearized as

$$L_{\text{ON}}(\nu) = L_{\text{free}}(\nu) + \tau_a(\nu) C_B \Delta T \sum_{r=1}^{N_r} \gamma_r \alpha_r(\nu). \tag{7}$$

The measured spectra are sampled by the HIS sensor at a set of bands $[\nu_1, \nu_2, \ldots, \nu_p]$ to produce a measurement vector $\mathbf{x} = [L_{\text{ON}}(\nu_1), \ldots, L_{\text{ON}}(\nu_p)]^T \in \mathbb{R}^{p \times 1}$, where p is the number of channels. The measured radiance spectrum vector $\mathbf{x}$ of a CWA-on background pixel is represented as

$$\mathbf{x} = \sum_{r=1}^{N_r} \mathbf{s}_r g_r + \mathbf{v} = \mathbf{S}\mathbf{g} + \mathbf{v} \tag{8}$$

where $\mathbf{s}_r = [\tau_a(\nu_1) \cdot \alpha_r(\nu_1), \ldots, \tau_a(\nu_p) \cdot \alpha_r(\nu_p)]^T \in \mathbb{R}^{p \times 1}$ is the at-sensor CWA signature vector of the rth CWA of interest, the intensity of the rth CWA signal is $g_r = C_B \Delta T \gamma_r$, for $r = 1, \ldots, N_r$, and $\mathbf{v} = [L_{\text{free}}(\nu_1), \ldots, L_{\text{free}}(\nu_p)]^T \in \mathbb{R}^{p \times 1}$ is the background clutter. The at-sensor CWA signature matrix $\mathbf{S}$ is $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_{N_r}] \in \mathbb{R}^{p \times N_r}$, and the intensity vector $\mathbf{g}$ is $\mathbf{g} = [g_1, \ldots, g_{N_r}]^T \in \mathbb{R}^{N_r \times 1}$. Equation (8) is referred to as a linear mixing model (LMM). Although the LMM is a less accurate model than the line-by-line radiative transfer (LBL–RT) model, the LMM have been widely used in CWA detection studies [1], [4], [5], since the LMM represents the difference of the measured spectrum directly depending on the presence or absence of the CWA cloud and is a linear model to which effective detection algorithms can be applied.

The background clutter $\mathbf{v}$ is the measured spectrum $L_{\text{free}}(\nu)$ of a CWA-free background pixel. In (1), the measured radiance spectrum $L_{\text{free}}(\nu)$ of the CWA-free background pixel is not just the background radiance spectrum $L_b(\nu)$, but the radiance spectrum composed of the atmospheric path radiance $L_a(\nu)$, the background radiance $L_b(\nu)$, which is modulated by the atmospheric transmittance $\tau_a(\nu)$, and noise $n(\nu)$. Each background clutter $\mathbf{v}$ has a certain amount of random variation about a nominal mean according to changes in the physical composition, temperature, atmospheric transmission, and sensor noise of the corresponding pixel. The background clutter $\mathbf{v}$ follows the Gaussian distribution with the mean vector $\mathbf{m} \in \mathbb{R}^{p \times 1}$ and the covariance matrix $\sigma^2\mathbf{C}$, i.e., $\mathbf{v} \sim N(\mathbf{m}, \sigma^2\mathbf{C})$, where σ^2 is the correction factor for the covariance matrix [1]. The covariance matrix $\mathbf{C} \in \mathbb{R}^{p \times p}$ contains not only the variation according to changes in physical composition, temperature, atmospheric transmission, and sensor noise of each pixel for the IFOV but also the correlation between background clutters at the different wavenumbers.

The NMF estimates the intensity of the CWA signature $\hat{\mathbf{g}}$, the background clutter $\hat{\mathbf{v}}$, and the correction factor $\hat{\sigma}^2$ from the measured spectrum vector $\mathbf{x}$ using the background statistics as follows:

$$\hat{\mathbf{g}} = (\mathbf{S}^T\mathbf{C}^{-1}\mathbf{S})\mathbf{S}^T\mathbf{C}^{-1}(\mathbf{x} - \mathbf{m}) \tag{9}$$

$$\hat{\mathbf{v}} = \mathbf{x} - \mathbf{S}\hat{\mathbf{g}} \tag{10}$$

$$\hat{\sigma}^2 = (\mathbf{x} - \mathbf{m})^T\mathbf{C}^{-1}(\mathbf{x} - \mathbf{m}). \tag{11}$$

The test statistic of the NMF is obtained by substituting (9) and (11) into the likelihood ratio of $\mathbf{x}$ [14], as follows:

$$T(\mathbf{x}) = \frac{\left[(\mathbf{x} - \mathbf{m})^T\mathbf{C}^{-1}\mathbf{S}\right]\left[\mathbf{S}^T\mathbf{C}^{-1}\mathbf{S}\right]^{-1}\left[\mathbf{S}^T\mathbf{C}^{-1}(\mathbf{x} - \mathbf{m})\right]}{(\mathbf{x} - \mathbf{m})^T\mathbf{C}^{-1}(\mathbf{x} - \mathbf{m})}$$

$$= \frac{(\tilde{\mathbf{x}}^T\tilde{\mathbf{S}})(\tilde{\mathbf{S}}^T\tilde{\mathbf{S}})^{-1}(\tilde{\mathbf{S}}^T\tilde{\mathbf{x}})}{(\tilde{\mathbf{x}}^T\tilde{\mathbf{x}})} = \cos^2\phi \tag{12}$$

where $\tilde{\mathbf{x}} = \mathbf{C}^{-1/2}(\mathbf{x} - \mathbf{m})$ is the whitened measured vector and $\tilde{\mathbf{S}} = \mathbf{C}^{-1/2}\mathbf{S}$ is the whitened at-sensor signature matrix. The test statistic is interpreted as the square value of the cosine of the angle ϕ between the whitened measured spectrum vector $\tilde{\mathbf{x}}$ and the subspace formed by the whitened at-sensor signature matrix $\tilde{\mathbf{S}}$. If $T(\mathbf{x})$ is larger than a detection threshold λ, the pixel corresponding to the spectrum $\mathbf{x}$ is determined as the CWA-on background pixel. Otherwise, it is determined as the CWA-free background pixel.

According to (12), these background statistics are estimated from a background training dataset that consists of sample CWA-free background pixels. In general, accurate estimation of the inverse of the covariance matrix is difficult due to several reasons, e.g., sample overfiting, ill-conditioned problems, high-dimensional data, the small number of samples, etc. In such cases, the regularization of the sample covariance matrix is useful to build the robust NMF. There are several regularization methods, i.e., the Hoffbeck and Landgrebe regularization method [20], [21], the truncated singular value decomposition regularization method [22], and the diagonal loading technique (DLT) [1], [23].

In the hyperspectral image (HSI) data, the ratio of the largest eigenvalue to the smallest eigenvalue of the covariance matrix is extremely large. Since the covariance matrix is close to a singular matrix, it is very hard to obtain an accurate inverse covariance matrix [1]. This phenomenon is referred to as the ill-conditioned problem. Among regularization methods, the DLT, which sets eigenvalues of the covariance matrix above a certain value, is effective in solving the ill-conditioned problem and robust against the distortion of the background covariance matrix. In the DLT, the background covariance matrix $\mathbf{C}$ is approximated from the sample background covariance matrix as

$$\mathbf{C} = \sum_{k=1}^{p} \lambda_k \mathbf{q}_k \mathbf{q}_k^T + \delta\mathbf{I} \tag{13}$$

where λ_k and $\mathbf{q}_k$, $k = 1, 2, \ldots, p$, are the eigenvalues and eigenvectors for the sample background covariance matrix, respectively, and $\delta > 0$ is the loading parameter. Then, the inverse of the background covariance matrix $\mathbf{C}^{-1}$ with the DLT is obtained as following [4]:

$$\mathbf{C}^{-1} = \frac{1}{\delta}\left(\mathbf{I} - \sum_{k=1}^{p} \frac{\lambda_k}{\delta + \lambda_k}\mathbf{q}_k\mathbf{q}_k^T\right). \tag{14}$$

B. NMF Under Background Contamination

The previous studies proved that the signal to clutter ratio, which represents the performance of the MF, is reduced [15]–[17] when background contamination occurs. In this article, we briefly show that the detection performance of the NMF with the DLT deteriorates as the concentration and the proportion of CWA-on background pixels in the contaminated dataset increase. Let H_0 and H_1 correspond to hypotheses of the absence and presence of the target CWA, respectively. Let $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ denote the contaminated training dataset, which contains both CWA-free background pixels and CWA-on background pixels. Here, N is the number of pixels in the contaminated dataset, and the degree of contamination, which indicates a ratio of CWA-on background pixels in the contaminated dataset, is a%. Then, the numbers of CWA-free background pixels and CWA-on background pixels are $(1-a)N$ and aN, respectively.

The jth pixel $\mathbf{x}_j$ in the contaminated dataset can be represented with the following two hypotheses:

$$\begin{aligned} H_0 &: \mathbf{x}_j = \mathbf{v}_j \\ H_1 &: \mathbf{x}_j = \mathbf{S}\mathbf{g}_j + \mathbf{v}_j \end{aligned} \tag{15}$$

where $\mathbf{S}$ is the at-sensor CWA signature matrix, and $\mathbf{g}_j$ is the intensity vector of the CWA signature. Let $\mathbf{m}_g$ denote the mean vector of $\mathbf{g}_j$. The background clutter $\mathbf{v}_j$ follows the multivariate Gaussian distribution with the mean vector $\mathbf{m}_0$ and the covariance matrix $\sigma^2\mathbf{C}_0$, i.e., $\mathbf{v}_j \sim N(\mathbf{m}_0, \sigma^2\mathbf{C}_0)$.

Then, the mean vector $\mathbf{m}_{\text{con}}$ of the contaminated dataset is expressed as follows:

$$\mathbf{m}_{\text{con}} = \frac{1}{N}\left\{\sum_{i=1}^{(1-a)N} \mathbf{x}_{i,0} + \sum_{k=1}^{aN} \mathbf{x}_{k,1}\right\} = \mathbf{m}_0 + a\mathbf{S}\mathbf{m}_g \tag{16}$$

where $\mathbf{x}_{i,0}$ and $\mathbf{x}_{k,1}$ are the CWA-free background pixels and the CWA-on background pixels in the contaminated dataset, respectively. The sample covariance matrix $\bar{\mathbf{C}}_{\text{con}}$ of the contaminated dataset is also obtained as

$$\begin{aligned} \bar{\mathbf{C}}_{\text{con}} &= \frac{1}{N}\sum_{j=1}^{N}(\mathbf{x}_j - \mathbf{m}_{\text{con}})(\mathbf{x}_j - \mathbf{m}_{\text{con}})^T \\ &= \frac{1}{N}\sum_{i=1}^{(1-a)N}(\mathbf{v}_i - \mathbf{m}_0 - a\mathbf{S}\mathbf{m}_g)(\mathbf{v}_i - \mathbf{m}_0 - a\mathbf{S}\mathbf{m}_g)^T \\ &\quad + \frac{1}{N}\sum_{k=1}^{aN}(\mathbf{v}_k + \mathbf{S}\mathbf{g}_k - \mathbf{m}_0 - a\mathbf{S}\mathbf{m}_g) \\ &\quad \cdot (\mathbf{v}_k + \mathbf{S}\mathbf{g}_k - \mathbf{m}_0 - a\mathbf{S}\mathbf{m}_g)^T. \end{aligned} \tag{17}$$

Assuming that $\mathbf{v}_k$ and $\mathbf{g}_k$ are independent, the sample covariance matrix $\bar{\mathbf{C}}_{\text{con}}$ of the contaminated dataset is

$$\bar{\mathbf{C}}_{\text{con}} = \mathbf{C}_0 + a^2\mathbf{S}\left(a\mathbf{\Sigma}_g + (1-a)\mathbf{m}_g\mathbf{m}_g^T\right)\mathbf{S}^T \tag{18}$$

where $\mathbf{\Sigma}_g = \sum_{k=1}^{Na}(\mathbf{g}_k - a\mathbf{m}_g)(\mathbf{g}_k - a\mathbf{m}_g)^T$. After the DLT with the loading parameter is applied, the contaminated covariance matrix $\mathbf{C}_{\text{con}}$ can be represented as

$$\mathbf{C}_{\text{con}} = \mathbf{C}_0 + a^2\mathbf{S}\left(a\mathbf{\Sigma}_g + (1-a)\mathbf{m}_g\mathbf{m}_g^T\right)\mathbf{S}^T + \delta\mathbf{I}. \tag{19}$$

We refer to an NMF designed with the distorted background statistics as the contaminated NMF.

Let $\mathbf{x}_0 = \mathbf{v}$ and $\mathbf{x}_1 = \mathbf{v} + \mathbf{S}\mathbf{m}_g$ be a CWA-free background pixel and a CWA-on background pixel, respectively. We define T_{con} as the test statistic of the contaminated NMF. Then, the test statistics of the contaminated NMF for the CWA-free background pixel and the CWA-on background pixel are $T_{\text{con}}(\mathbf{x}_0) = \cos^2\phi_0$ and $T_{\text{con}}(\mathbf{x}_1) = \cos^2\phi_1$, respectively. Here, ϕ_0 denotes the angle between $\tilde{\mathbf{x}}_0$ and $\tilde{\mathbf{S}}$, and ϕ_1 represents the angle between $\tilde{\mathbf{x}}_1$ and $\tilde{\mathbf{S}}$. Using (16) and (19), $\tilde{\mathbf{x}}_0$, $\tilde{\mathbf{x}}_1$, and $\tilde{\mathbf{S}}$ can be represented as

$$\begin{aligned} \tilde{\mathbf{x}}_0 &= \left[\mathbf{C}_0 + a^2\mathbf{S}\left(a\mathbf{\Sigma}_g + (1-a)\mathbf{m}_g\mathbf{m}_g^T\right)\mathbf{S}^T + \delta\mathbf{I}\right]^{-1/2} \\ &\quad \cdot [\mathbf{v} - \mathbf{m}_0 - a\mathbf{S}\mathbf{m}_g] \end{aligned} \tag{20}$$

$$\begin{aligned} \tilde{\mathbf{x}}_1 &= \left[\mathbf{C}_0 + a^2\mathbf{S}\left(a\mathbf{\Sigma}_g + (1-a)\mathbf{m}_g\mathbf{m}_g^T\right)\mathbf{S}^T + \delta\mathbf{I}\right]^{-1/2} \\ &\quad \cdot [\mathbf{v} - \mathbf{m}_0 + (1-a)\mathbf{S}\mathbf{m}_g] \end{aligned} \tag{21}$$

$$\tilde{\mathbf{S}} = \left[\mathbf{C}_0 + a^2\mathbf{S}\left(a\mathbf{\Sigma}_g + (1-a)\mathbf{m}_g\mathbf{m}_g^T\right)\mathbf{S}^T + \delta\mathbf{I}\right]^{-1/2}\mathbf{S}. \tag{22}$$

In case that the DLT is not used, i.e., $\delta = 0$, the term $a^2\mathbf{S}(a\mathbf{\Sigma}_g + (1-a)\mathbf{m}_g\mathbf{m}_g^T)\mathbf{S}^T$ whitens CWA signatures in $\tilde{\mathbf{x}}_1$ and $\tilde{\mathbf{S}}$ even though the degree of contamination is low. Since the magnitudes of CWA signature in $\tilde{\mathbf{x}}_0$ and $\tilde{\mathbf{x}}_1$ are ignored, ϕ_0 and ϕ_1 are about the same. Then, the NMF cannot distinguish between CWA-free background pixels and CWA-on background pixels at all, and the detection probability is the same as the false alarm probability.

In case that the DLT is applied to the contaminated NMF, since the term $\delta\mathbf{I}$ mitigates whitening of CWA signatures, the difference between ϕ_0 and ϕ_1 are mainly determined by the second terms in (20) – (22). As the degree of contamination a increases, $T_{\text{con}}(\mathbf{x}_0)$ is elevated and it causes a higher false alarm probability. On the other hand, $T_{\text{con}}(\mathbf{x}_1)$ is degraded, and the detection probability is lowered. If the degree of contamination is over 50%, the $T_{\text{con}}(\mathbf{x}_1)$ is smaller than $T_{\text{con}}(\mathbf{x}_0)$, and the NMF classifies CWA-free background pixels as CWA-on background pixels and vice versa. As with the degree of contamination, the larger concentration of CWA-on background pixels in the contaminated dataset, the worse the detection performance of the NMF. No matter how the concentration of CWA-on background pixels increases, however, $T_{\text{con}}(\mathbf{x}_1)$ is larger than $T_{\text{con}}(\mathbf{x}_0)$, and the reverse classification of the NMF does not occur.

III. Related Works on Background Contamination

In this section, we briefly review related works to overcome the background contamination problem.

A. Robust NMF (R-NMF)

As with the NMF, the performance of the MF is degraded when background contamination occurs. To deal with this problem, Niu *et al.* [17] proposed the robust MF (R-MF), which applies the DLT. In the R-MF, the optimal loading parameter $\hat{\delta}$, which minimizes distortion for the background covariance matrix, is determined by solving the following equation:

$$\sum_{k=1}^{p} \frac{|\tilde{s}_k|^2}{(1+\delta^{-1}\lambda_k)} = \kappa \tag{23}$$

where $\tilde{s}_k = \mathbf{q}_k^T \mathbf{s}$, λ_k, and $\mathbf{q}_k$ are the eigenvalues and eigenvectors noted in (13), κ is a positive number satisfying $\|\mathbf{s}-\mathbf{s}_0\|^2 \leq \kappa$, and $\mathbf{s}_0$ is the absorption coefficient vector of a CWA. The MF and the NMF are in a similar structure and the same background statistics are used to construct both the MF and the NMF. Therefore, we define R-NMF as an NMF, which uses the DLT with the optimal loading parameter $\hat{\delta}$. The optimal loading parameter $\hat{\delta}$ mitigates the distortion for the background covariance. However, it is not a fundamental solution because it does not eliminate the distortion of the background mean at all.

B. Iterative NMF (I-NMF)

Kim *et al.* [18] proposed an iterative MF (I-MF), which iteratively applies an MF to exclude the CWA-on background pixels from the contaminated training dataset. Like the R-NMF, we define an I-NMF that iteratively applies an NMF. The I-NMF can only be designed if the degree of contamination a is known in advance. The process of the I-NMF algorithm is summarized as follows. First, the initial I-NMF detector $T_{\text{I-NMF}}^{(0)}$ is designed using the contaminated dataset $\mathbf{X}$ and (12). Next, the initial NMF is applied to all pixels in the contaminated training dataset. Subsequently, the contaminated training dataset is divided into the CWA-free background pixel set, $\mathbf{X}_0^{(i+1)}$ and the CWA-on background pixel set $\mathbf{X}_1^{(i+1)}$:

$$\mathbf{X}_0^{(i+1)} = \left\{ \mathbf{x} \,\middle|\, \mathbf{x} \in \mathbf{X},\ \ T_{\text{I-NMF}}^{(i)}(\mathbf{x}) < \eta^{(i)} \right\} \tag{24}$$

$$\mathbf{X}_1^{(i+1)} = \left\{ \mathbf{x} \,\middle|\, \mathbf{x} \in \mathbf{X},\ \ T_{\text{I-NMF}}^{(i)}(\mathbf{x}) \geq \eta^{(i)} \right\} \tag{25}$$

where $\eta^{(i)}$ denotes the detection threshold at the ith iteration. The detection threshold $\eta^{(i)}$ is determined as the test statistic value of the ith NMF $T_{\text{I-NMF}}^{(i)}(\mathbf{x})$ for the pixel whose test statistic value is the upper a%.

Next, $T_{\text{I-NMF}}^{(i)}(\mathbf{x})$ is constructed using pixels in the $(i+1)$st CWA-free background pixel set $\mathbf{X}_0^{(i+1)}$. This process is repeated until the mean value of test statistics converges. The convergence condition is expressed as

$$\left| M^{(i+1)} - M^{(i)} \right| < \beta \tag{26}$$

where $M^{(i)} = \frac{1}{N}\sum_{j=1}^{N} T_{\text{I-NMF}}^{(i)}(\mathbf{x}_j)$ is the mean value of the test statistics at the ith iteration and β represents a convergence threshold. Although the I-NMF eliminates the contaminated pixels to some extent, it is considered impractical since it is difficult to predict the degree of contamination in advance.

IV. Hard EM-NMF

Our proposed algorithm consists of two procedures. First, we estimate the posterior probability of each pixel in the contaminated dataset. The optimal posterior probability, which maximizes the log-likelihood of the contaminated dataset, is obtained using the EM algorithm. To achieve a fast and accurate convergence of the EM algorithm, we also propose an initialization method for the EM algorithm. Second, we extract CWA-free background pixels by thresholding the posterior probabilities and design an hard EM-NMF using the extracted CWA-free background pixels.

A. Estimation of the Posterior Probability

1) Log-Likelihood of the Contaminated Dataset: From the signal model (15) describing the jth pixel $\mathbf{x}_j$ in the contaminated dataset $\mathbf{X}$, $\mathbf{x}_j|H_i$ is a Gaussian random vector with mean $\mathbf{m}_i$ and covariance $\sigma_i^2\mathbf{C}_i$. Because the variation of the CWA signals is less severe than that of the background clutter, we assume that the covariance matrix of the CWA-free background pixels is identical to that of the CWA-on background pixels, i.e., $\sigma_1^2\mathbf{C}_1 = \sigma_0^2\mathbf{C}_0 = \mathbf{C}$. The probability density function (PDF) of $\mathbf{x}_j$ under H_i, $i = 0, 1$ can be expressed as

$$P(\mathbf{x}_j|\,H_0) = \rho \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{m}_0)^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{m}_0)\right] \tag{27}$$

$$P(\mathbf{x}_j|\,H_1) = \rho \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{m}_1)^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{m}_1)\right] \tag{28}$$

where $\rho = 1/\sqrt{(2\pi)^p|\mathbf{C}|}$ and p is the number of spectral channels of the HSI image.

We define $P(H_i|\mathbf{x}_j)$ as the posterior probability of the pixel $\mathbf{x}_j$ for a hypothesis H_i for $i = 0, 1$, which represents the probability that the pixel $\mathbf{x}_j$ belongs to a background hypothesis H_0 or a CWA hypothesis H_1, respectively. From Bayes' rule [26], the posterior probability $P(H_i|\mathbf{x}_j)$ is represented as follows:

$$P(H_i|\mathbf{x}_j) = \frac{P(H_i)P(\mathbf{x}_j|\mathbf{m}_i,\mathbf{C})}{\sum_{k=0}^{1}P(H_k)P(\mathbf{x}_j|\mathbf{m}_k,\mathbf{C})} \tag{29}$$

where $P(H_i)$ denotes a mixing coefficient (called as *a priori* probability) corresponding to H_i, which satisfies $P(H_0) + P(H_1) = 1$. Here, $P(H_i)$ for $i = 0, 1$ refers to the ratio of background or CWA-on background pixels in the contaminated dataset, respectively. To find $P(H_i|\mathbf{x}_j)$, it is necessary to estimate $\mathbf{m}_i$, $P(H_i)$ and $\mathbf{C}$. Let Θ denote a model parameter set defined as $\Theta = \{P(H_0), P(H_1), \mathbf{m}_0, \mathbf{m}_1, \mathbf{C}\}$. We calculate the maximum-likelihood estimate (MLE) of the model parameter set, $\hat{\Theta}$, that maximizes the likelihood probability of the contaminated dataset $P(\mathbf{X}|\Theta)$.

Assuming that pixels in $\mathbf{X}$ are independent and identically distributed, the PDF of the contaminated dataset $\mathbf{X}$ is expressed as

$$P(\mathbf{X}|\Theta) = \prod_{j=1}^{N} P(\mathbf{x}_j|\Theta) \tag{30}$$

where N is the number of pixels in the contaminated dataset and $P(\mathbf{x}_j|\Theta)$ represents a mixture PDF of $\mathbf{x}_j$. It is expressed as follows:

$$P(\mathbf{x}_j|\Theta) = \sum_{i=0}^{1} P(H_i)P(\mathbf{x}_j|H_i, \mathbf{m}_i, \mathbf{C}). \tag{31}$$

According to (30) and (27), $P(\mathbf{X}|\Theta)$ is represented as the product of many exponential terms. These terms make it hard to obtain $\hat{\Theta}$ given that solving $\nabla_\Theta P(\mathbf{X}|\Theta) = \mathbf{0}$ is very difficult. Therefore, we take the logarithm of $P(\mathbf{X}|\Theta)$. Then, $Q = \ln P(\mathbf{X}|\Theta)$ is represented as the summation of the polynomial terms

$$Q = \sum_{j=1}^{N} \ln\left(\sum_{i=0}^{1} P(H_i)P(\mathbf{x}_j|H_i, \mathbf{m}_i, \mathbf{C})\right). \tag{32}$$

2) Optimal Solution for Log-Likelihood Maximization: In order to obtain $\hat{\Theta}$, we need to solve $\nabla_\Theta Q = \mathbf{0}$. First, by solving $\partial Q/\partial \mathbf{m}_i = \mathbf{0}$, the optimal mean vector, $\hat{\mathbf{m}}_i$, is determined by

$$\hat{\mathbf{m}}_i = \frac{1}{N_i}\sum_{j=1}^{N} P\left(H_i|\mathbf{x}_j\right)\mathbf{x}_j \tag{33}$$

where $P(H_i|\mathbf{x}_j)$ is represented in (29), and N_i denotes the effective number of pixels belonging to H_i

$$N_i = \sum_{j=1}^{N} P\left(H_i|\mathbf{x}_j\right). \tag{34}$$

Second, the optimal covariance matrix $\hat{\mathbf{C}}$ is found by solving $\partial Q/\partial \mathbf{C}^{-1} = 0$

$$\hat{\mathbf{C}} = \sum_{j=1}^{N}\sum_{i=0}^{1} P\left(H_i\,|\mathbf{x}_j\right)\left(\mathbf{x}_j - \mathbf{m}_i\right)\left(\mathbf{x}_j - \mathbf{m}_i\right)^T. \tag{35}$$

From (33) and (35), the optimal statistics $(\hat{\mathbf{m}}_0, \hat{\mathbf{m}}_1, \hat{\mathbf{C}})$ in $\hat{\Theta}$ are interpreted as the weighted mean vector and covariance matrix with $P(H_i|\mathbf{x}_j, \hat{\Theta})$ as weights. The posterior probability $P(H_1|\mathbf{x}_j, \hat{\Theta})$ for the CWA hypothesis is matched with the label of the pixel with a value of 0 for a CWA-free background pixel or 1 for a CWA-on background pixel. Therefore, $P(H_1|\mathbf{x}_j, \hat{\Theta})$ is regarded as a soft label.

Finally, the optimal mixing coefficient $\hat{P}(H_i)$ can be attained by solving the following optimization problem:

$$\begin{aligned} \max_{P(H_i)} \quad & \sum_{j=1}^{N} \ln \sum_{i=0}^{1} P(H_i)P(\mathbf{x}_j|H_i, \mathbf{m}_i, \mathbf{C}) \\ \text{subject to} \quad & P(H_0) + P(H_1) = 1 \\ & P(H_i) \geq 0. \end{aligned} \tag{36}$$

We solve the abovementioned problem using the Karush–Kuhn–Tucker (KKT) condition, a necessary condition for a solution of the optimization problem to be optimal [24]. First, we set the Lagrangian function for (36) as follows:

$$\begin{aligned} L\left(P\left(H_i\right), \upsilon\right) = & \sum_{j=1}^{N} \ln \sum_{i=0}^{1} P(H_i)P(\mathbf{x}_j\,|H_i, \mathbf{m}_i, \mathbf{C}\,) \\ & - \upsilon\left(\sum_{i=0}^{1} P(H_i) - 1\right). \end{aligned} \tag{37}$$

According to the KKT condition, the optimal solutions of (37), $(\hat{P}(H_i), \hat{\upsilon})$, satisfy $\partial L/\partial P(H_i) = 0$ and $\partial L/\partial \upsilon = 0$ as

$$\begin{aligned} \frac{\partial L}{\partial P(H_i)} &= \sum_{j=1}^{N} \frac{P(\mathbf{x}_j\,|\mathbf{m}_i, \mathbf{C}\,)}{\sum_{k=0}^{1} P(H_k)P(\mathbf{x}_j\,|\mathbf{m}_k, \mathbf{C}\,)} - \upsilon \\ &= 0 \end{aligned} \tag{38}$$

$$\frac{\partial L}{\partial \upsilon} = P(H_0) + P(H_1) - 1 = 0. \tag{39}$$

By multiplying (38) by $P(H_i)$ and substituting it into (39), $\hat{P}(H_i)$ is derived as

$$\hat{P}(H_i) = \frac{1}{N}\sum_{j=1}^{N} P\left(H_i\,|\mathbf{x}_j\,\right) = \frac{N_i}{N}. \tag{40}$$

3) Optimization With the EM Algorithm: According to (33), (35), and (40), $\hat{\Theta}$ cannot be expressed in a closed form solution. Therefore, we exploit an expectation maximization algorithm to calculate $\hat{\Theta}$. The EM algorithm is an iterative means of finding the MLE of the parameters in a statistical model [27]. The EM algorithm alternates between an expectation step (E-step) and a maximization step (M-step). In the E-step, using the mth iteration's model parameter set $\Theta^{(m)} = \{P^{(m)}(H_0), P^{(m)}(H_1), \mathbf{m}_0^{(m)}, \mathbf{m}_1^{(m)}, \mathbf{C}^{(m)}\}$, we calculate the mth iteration's posterior probability $P(H_i|\mathbf{x}_j, \Theta^{(m)})$ for H_i $(i = 0, 1)$, as follows:

$$P(H_i|\mathbf{x}_j, \Theta^{(m)}) = \frac{P^{(m)}(H_i)P(\mathbf{x}_j|\mathbf{m}_i^{(m)}, \mathbf{C}^{(m)})}{\sum_{k=0}^{1} P^{(m)}(H_k)P(\mathbf{x}_j|\mathbf{m}_k^{(m)}, \mathbf{C}^{(m)})}. \tag{41}$$

In the M-step, we update $\Theta^{(m)}$ to $\Theta^{(m+1)}$, which maximizes the log-likelihood probability Q as

$$P^{(m+1)}(H_i) = \frac{N_i^{(m)}}{N} \tag{42}$$

$$\mathbf{m}_i^{(m+1)} = \frac{1}{N_i^{(m)}}\sum_{j=1}^{N} P(H_i|\mathbf{x}_j, \Theta^{(m)})\mathbf{x}_j \tag{43}$$

$$\begin{aligned} \mathbf{C}^{(m+1)} = & \frac{1}{N}\sum_{j=1}^{N}\sum_{k=0}^{1} P(H_i|\mathbf{x}_j, \Theta^{(m)}) \\ & \cdot (\mathbf{x}_j - \mathbf{m}_i^{(m)})(\mathbf{x}_j - \mathbf{m}_i^{(m)})^T \end{aligned} \tag{44}$$

$$N_i^{(m)} = \sum_{j=1}^{N} P(H_i|\mathbf{x}_j, \Theta^{(m)}). \tag{45}$$

We then acquire $Q^{(m+1)}$ by substituting $\Theta^{(m+1)}$ into (32). The E-step and the M-step are iteratively performed until the convergence criterion is achieved, which is $|Q^{(m+1)} - Q^{(m)}| < \beta$, where β is the convergence threshold.

4) Initialization Method for the EM Algorithm: The convergence result of the EM algorithm is sensitive to an initial model parameter set $\Theta^{(0)}$. During the process of allocating $\Theta^{(0)}$, pixels in $\mathbf{X}$ are assigned to two-pixel sets, i.e., the CWA-free background pixel set $\mathbf{X}_0^{(0)}$ and the CWA-on background pixel set $\mathbf{X}_1^{(0)}$. Then, $\Theta^{(0)}$ is calculated from the assigned pixels.

We propose an initialization method that obtains $\Theta^{(0)}$ from pixels classified by the R-NMF. In most cases, the proposed initialization method provides fast and accurate convergence because the R-NMF can roughly distinguish between CWA-free background pixels and CWA-on background pixels.

Let $T_{\text{R-NMF}}(\mathbf{x})$ denote a test statistic of the R-NMF. Using the initial R-NMF threshold η, we classify pixels in $\mathbf{X}$ as

$$\mathbf{X}_0^{(0)} = \{\mathbf{x} \,|\mathbf{x} \in \mathbf{X}, T_{\text{R-NMF}}(\mathbf{x}) < \eta\} \tag{46}$$

$$\mathbf{X}_1^{(0)} = \{\mathbf{x} \,|\mathbf{x} \in \mathbf{X}, T_{\text{R-NMF}}(\mathbf{x}) \geq \eta\} \tag{47}$$

where $\mathbf{X}_i^{(0)}$ denotes a set of pixels assigned to H_i, $i = 0, 1$ to calculate $\Theta^{(0)}$. Subsequently, $\Theta^{(0)}$ is determined from the pixels in $\mathbf{X}_0^{(0)}$ and $\mathbf{X}_1^{(0)}$. Because $T_{\text{R-NMF}}(\mathbf{x})$ varies with the degree of contamination and the concentration of the contaminated CWA-on background pixels, it is recommended to set $\eta = E(T_{\text{R-NMF}}(\mathbf{X}))$ rather than to a specific value, where $E(\cdot)$ is the expectation of $(\cdot)$.

B. Design of Hard EM-NMF

1) Hypothesis Labeling Check: If the degree of contamination is over 50%, the test statistic of the R-NMF, $T_{\text{R-NMF}}(\mathbf{x}_0)$ for the CWA-free background pixels are higher than that for CWA-on background pixels $T_{\text{R-NMF}}(\mathbf{x}_1)$. When assigning the initial model parameter set, CWA and CWA-free background pixels are mainly allocated to the CWA hypothesis and background hypothesis, respectively. As a result, the EM algorithm recognizes CWA-free background pixels as the CWA hypothesis H_1 and CWA-on background pixels as the background hypothesis H_0. To check whether pixels belonging the background hypothesis H_0 is the actual CWA-free background pixels or CWA-on background pixels, we use correlation coefficients ρ_i for $i = 0, 1$ between estimated mean vectors $\hat{\mathbf{m}}_i$ and the CWA signature matrix $\mathbf{S}$ as follows:

$$\rho_i = \sqrt{\frac{(\hat{\mathbf{m}}_i^T \mathbf{S})(\mathbf{S}^T \mathbf{S})^{-1}(\mathbf{S}^T \hat{\mathbf{m}}_i)}{\hat{\mathbf{m}}_i^T \hat{\mathbf{m}}_i}}. \tag{48}$$

If $\rho_0 > \rho_1$, the hypothesis labeling is wrong. Then, we exchange posterior probabilities for each hypothesis and recalculate the optimal model parameter set $\hat{\Theta} = \{\hat{P}(H_0), \hat{P}(H_1), \hat{\mathbf{m}}_0, \hat{\mathbf{m}}_1, \hat{\mathbf{C}}\}$.

2) Design of the Hard EM-NMF: The optimal background statistics $(\hat{\mathbf{m}}_0, \hat{\mathbf{C}})$ are background statistics robust to distortion caused by background contamination. We denote an NMF designed with background statistics as a soft EM-NMF. However, there is still room for improvement. If $\mathbf{x}_j$ is a CWA-on background pixel, the background statistics are affected by $P(H_0|\mathbf{x}_j, \hat{\Theta})$ of the CWA-on background pixel, as observed in (33) and (35). For better exclusion of the effect of contamination, it is necessary to extract the CWA-free background pixels from the contaminated dataset and calculate the statistics of the extracted CWA-free background pixels.

Algorithm 1: Design of the Hard EM-NMF From a Contaminated Dataset $\mathbf{X}$.

```
Using (12) and (13), construct an R-NMF
Find Θ^(0) from (46) and (47)
while |Q^(m) − Q^(m−1)| < β do
  m = m + 1
  Obtain P(H_i|x_j, Θ^(m)) for i = 0, 1 (E-Step)
  Update Θ^(m) using (42), (43) and (44) (M-step)
  Find Q^(m)
end while
Find ρ_0 and ρ_1 using (48).
If ρ_0 > ρ_1, exchange P(H_0|x_j) and P(H_1|x_j).
Calculate Θ* using (33), (35) and (40)
Extract X_0 using (49)
Design the hard EM-NMF using (12), (13) and X_0
```

We extract CWA-free background pixels $\mathbf{X}_0$ from the contaminated dataset based on $P(H_1|\mathbf{x}_j, \hat{\Theta})$ as follows:

$$\mathbf{X}_0 = \left\{\mathbf{x} \,\middle|\, \mathbf{x} \in \mathbf{X}, P(H_1|\mathbf{x}_j, \hat{\Theta}) < \zeta\right\} \tag{49}$$

where ζ is the classification threshold. In setting the classification threshold ζ, it is crucial to minimize side effects due to misclassification. Through several experiments, we determine the optimal classification threshold as $\zeta = 0.1$. Then, we calculate the mean vector and covariance matrix of $\mathbf{X}_0$ and design an NMF exploiting (12) and (13). We denote an NMF designed using $\mathbf{X}_0$ as a hard EM-NMF. The overall design process of the hard EM-NMF is described in Algorithm 1.

V. Experimental Results

In this section, we describe experiments that were conducted to test the performance of the proposed algorithm on both synthetic and real CWA data. HSIs used in the experiments were collected by a Bruker hyperspectral imaging system (HI-90, Bruker Corporation, Germany), which provides data with a spectral resolution of 3.2 cm^{-1} from 903 to 1264 cm^{-1} with 128 channels and a spatial resolution of 122×122 pixels [28]. The at-sensor CWA signature matrix was composed of sulfur hexafluoride (SF_6), Freon, tabun (GA), sarin (GB), mustard gas (HD), methanol (MeOH), and triethyl phosphate (TEP). Each at-sensor CWA signature vector $\mathbf{s}_r$, for $r = 1, \ldots, 8$, is obtained by interpolating the standard absorption coefficient spectrum vector α_r of the rth CWA material, which is sourced from the National Institute of Standard and Technology (NIST) [29], to suit the resolution of an HI-90 equipment and is shown in Fig. 2.

A. Synthetic Data Experiments

In the synthetic data experiment, the synthetic CWA data were generated by embedding the CWA signals into the background spectra based on (7). The TEP gas was selected as the synthetic CWA. The charge-coupled device (CCD) image of the HSI used in the synthetic data experiment is shown in Fig. 3. There were 14 884 pixels in the contaminated dataset where synthetic CWA-on background pixels formed a circle at the

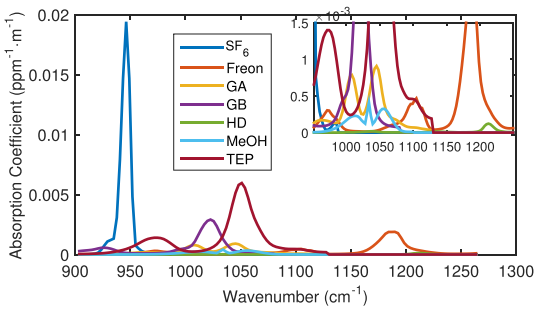

Fig. 2. Absorption coefficient spectra of target CWAs sourced from the NIST and the PNNL over the HI-90 wavelength range. The right upper figure shows the absorption coefficient spectra of the GA, MeOH, and HD, which have relatively small absorption coefficients.

Fig. 3. CCD image of the HSI used in the synthetic data experiment.

center of the CCD image. The degree of contamination was set to $a = 40\%$. The red pixels in Fig. 3 indicate the synthetic TEP-on background pixels. The TEP-free background pixels used to generate the synthetic CWA-on background pixels were taken with a background of grass and a building in Yuseong-gu, Daejeon, South Korea, where the latitude is 36° N, the longitude is 127° E and the altitude is 212 m. These data were measured at 10 P.M. on April 11, 2017. The climate data of the experiment are as follow. The atmospheric temperature was 284.5 K, the background temperature was 286.3 K. The air humidity was 68.6 %, and the Zenith angle was 89°. It was sunny and there were no shadows.

The transmittance of the atmosphere τ_a was ascertained from the moderate resolution atmospheric transmission 4 (MODTRAN 4) [30]. We set the temperature difference to $\Delta T = 3$ K. The concentration path length for the TEP gas was assumed to be Gaussian-distributed: $\gamma \sim N(150,\, 75^2)$ ppm·m. The spectra of the TEP-free background pixels and synthetic TEP-on background pixels are shown in Fig. 4. There is only a slight difference between the TEP-free background spectra and the CWA-on background spectra in the 1000−1100 cm^{-1} band, which encompasses the absorption property of the TEP gas.

1) Estimation of the Posterior Probability: First, we performed an experiment to show how the EM algorithm given in Section IV-A3 estimates the posterior probability of each pixel. Fig. 5(a) presents the initial label image used to obtain

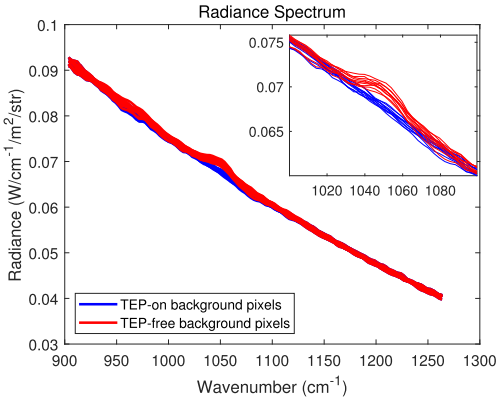

Fig. 4. Spectra of the synthetic TEP-on background pixels and TEP-free background pixels in the synthetic data experiment. The right upper figure shows the enlarged spectra in the 1,000 − 1,100 cm^{-1} band where the TEP gas has absorption properties.

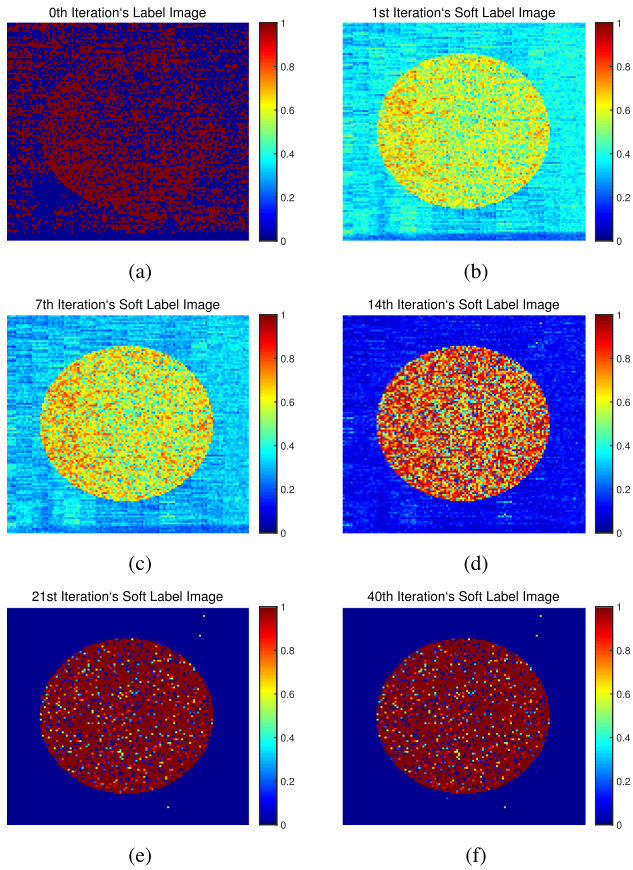

Fig. 5. Soft label images of the EM algorithm at the (a) initial, (b) 1st, (c) 7th, (d) 14th, (e) 21st, and (f) 40th.

the initial model parameter set $\Theta^{(0)}$. Fig. 5(b)–(f) illustrates images of the soft label $P(H_1|\mathbf{x}_j, \Theta^{(m)})$ at the first, seventh, fourteenth, twenty-first, and fortieth iteration, respectively. Label values for TEP-free background pixels gradually approach zero. Label values of some less-concentration TEP-on background pixels, which has little TEP signatures and are nearly identical to TEP-free background pixels, also reach to zero.

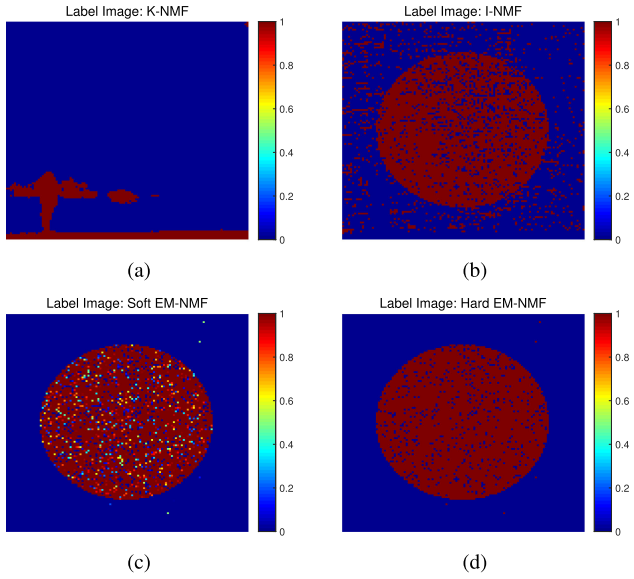

Fig. 6. Label images of (a) K-NMF, (b) I-NMF, (c) soft EM-NMF, and (d) hard EM-NMF for the contaminated dataset.

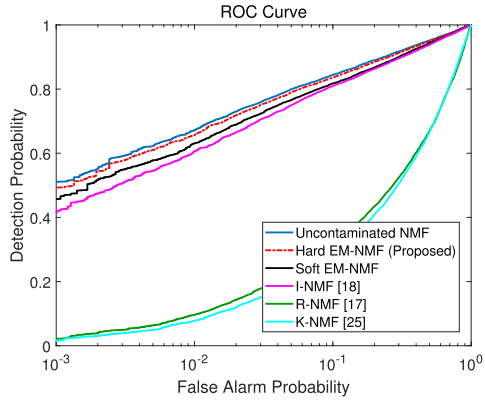

Fig. 7. ROC curves of several NMF algorithms.

In the case of low-concentration TEP-on background pixels that are ambiguous to be classified as TEP-on background pixels, the label values converge to $0.1 - 0.9$. The label values for the evident TEP-on background pixels become 1. As the number of iteration increases, the more precise the label values, the more accurate the model parameter set can be estimated, allowing the calculation of more precise label values. After the twenty-first iteration, there are no changes in the label images, demonstrating the convergence of the EM algorithm.

2) Comparison of Background Extraction Performances: To compare the background extraction performance, we obtain the label images of several NMFs, as shown in Fig. 6. The K-NMF is an NMF made with TEP-free background pixels extracted by the K-means clustering algorithm, which is one of the classification algorithms [25]. Because information about the degree of contamination is required to design the I-NMF, it is unfair to compare the EM-NMFs with the I-NMF. Nevertheless, we compared the EM-NMFs with the I-NMF because the I-NMF is a state-of-the-art algorithm for the background contamination problem. In the case of the soft EM-NMF, the posterior probability $P(H_1|\mathbf{x}_j, \Theta^{(m)})$ of each pixel is used as the label value.

These NMFs were designed using the same optimal loading parameter. In the real world, it is difficult to determine the actual value of κ in (23). Therefore, we found the optimal loading parameter numerically, i.e., $\hat{\lambda} = 5 \cdot 10^{-6}$. The convergence thresholds were set to $\beta = 10^{-3}$. In Fig. 5(f), the label values of TEP-on background pixels exceed 0.1, except for the less-concentration TEP-on background pixels, which hardly contaminate the background statistics. Therefore, we determined the classification threshold for the hard EM-NMF as $\zeta = 0.1$ in the experiments.

Compared with the ground truth depicted in Fig. 3, the K-means clustering does not distinguish TEP-free background pixels from TEP-on background pixels at all because it categorizes pixels based on the mean value of each pixel's radiance spectrum, which is determined by the temperature of the corresponding pixel. Therefore, the effective degree of contamination for the K-NMF is larger than that for the R-NMF. The I-NMF and the EM algorithm divide TEP-free background pixels from TEP-on background pixels correctly, except for the less-concentration TEP-on background pixels. The EM algorithm estimates TEP-free background pixels more accurately than the I-NMF since the EM algorithm calculates the posterior probabilities using not only background statistics but also TEP statistics, unlike the I-NMF that estimates TEP-free background pixels using the only background statistics. Comparing Fig. 6(c) with (d), we confirm that low-concentration TEP-on background pixels whose labels are ambiguous (i.e., $0.1 < P(H_i|\mathbf{x}_j, \hat{\Theta}) < 0.9$) in the soft EM-NMF are classified as TEP-on background pixels in the hard EM-NMF.

3) Detection Performance Comparison: We also use the receiver operating characteristic (ROC) curve, which directly shows the relationship between the probability of false alarm and that of detection, as a metric for evaluating the performance capabilities of the algorithms. We performed Monte Carlo simulations in which 14 884 TEP-free background pixels and 14 884 synthetic TEP-on background pixels whose concentration path lengths follow $\gamma \sim N(150,\ 75^2)$ ppm·m were used. Fig. 7 plots ROC curves of the uncontaminated NMF, soft and hard EM-NMFs, I-NMF, K-NMF, and R-NMF. The uncontaminated NMF represent an NMF designed with only TEP-free background pixels in the contaminated dataset.

The detection performance of the R-NMF is degraded since the background mean vector is still contaminated. The K-NMF has a worse detection performance than the R-NMF because the effective degree of contamination for the K-NMF is bigger than that of the R-NMF. The I-NMF and EM-NMFs show robust detection performances under the background contamination condition. According to Fig. 7, the hard EM-NMF is better than the I-NMF in terms of the detection performance although the degree of contamination is unknown. The hard EM-NMF has better performance than the soft EM-NMF since the background statistics of the hard EM-NMF is not distorted by low-concentration TEP-on background pixels. It also performs nearly equal to the uncontaminated NMF.

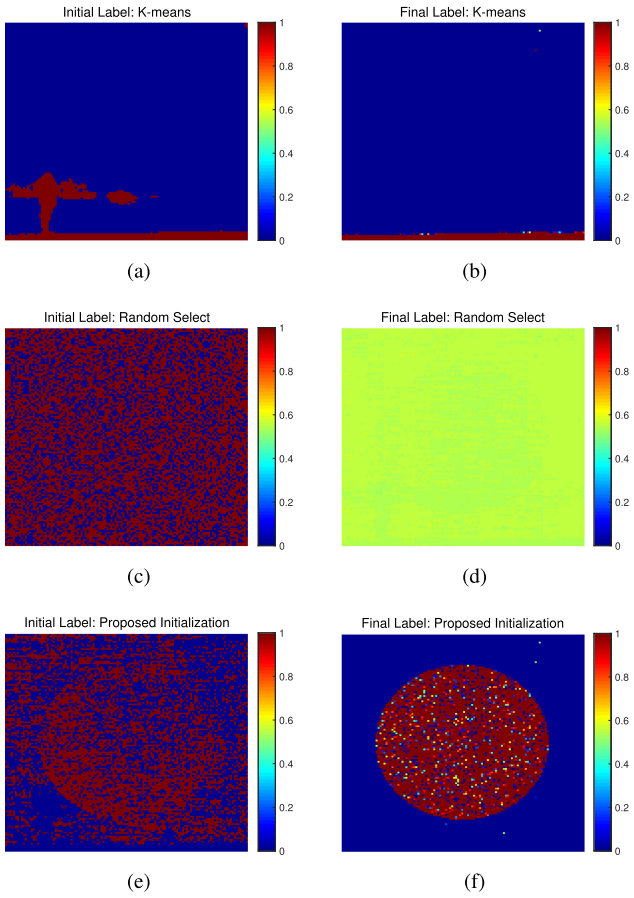

Fig. 8. Label images before (left images) and after (right images) the EM algorithm according to the initialization method: (a), (b) the K-means clustering, (c), (d) the random selection, and (e), (f) the proposed initialization method.

4) Comparison of Initialization Methods: Next, we performed experiments to compare our initialization method proposed in Section IV-A4 with two other initialization methods: the K-means clustering and the random selection. In the random selection, randomly selected half of the contaminated dataset, are categorized as CWA-on background pixels. In Fig. 8, images on the left are label images obtained from the initialization methods, and those on the right are the soft label images resulted from the EM algorithm. The K-means clustering categorizes pixels based on the pixels' temperature, and the random selection method classifies pixels without any criteria, as depicted in Fig. 8(a) and (c), respectively. Then, the EM algorithms with the K-means clustering or the random selection method do not estimate accurate soft label values, as shown in Fig. 8(b) and (d). However, the EM algorithm with the proposed initialization method estimates the soft label values well since the R-NMF distinguishes between TEP-free background pixels and CWA-on background pixels roughly.

5) Detection Performance According to Degree of Contamination: We conducted experiments to compare detection performances based on the degree of contamination. The AUC score defined as an area under a ROC curve is used as a metric for evaluating the performance capabilities of the algorithms. We changed the degree of contamination of the contaminated

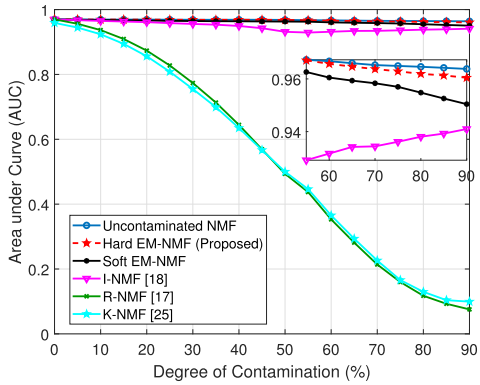

Fig. 9. AUC scores according to the degree of contamination from 0% to 90%. The right upper figure illustrates the AUC scores of the uncontaminated NMF, soft and hard EM-NMFs, and I-NMF according to the degree of contamination from 55% to 90%.

dataset a% from 0% to 90% at intervals of 5%, and the concentration path-lengths of synthetic TEP-on background pixels in the contaminated dataset were set to $\gamma \sim N(150,\ 75^2)$ ppm·m. To obtain the AUC score, we performed a Monte Carlo simulation with 14 884 TEP-free background pixels and 14 884 TEP-on background pixels whose concentration path length is $\gamma \sim N(150, 75^2)$ ppm · m. Fig. 9 depicts AUC scores of the R-NMF, K-NMF, I-NMF, and EM-NMFs.

The detection performance of the uncontaminated NMF decreases in the high degree of contamination because the covariance matrix is calculated with a few TEP-free background pixels and becomes slightly inaccurate. As the degree of contamination increases, the detection performances of the R-NMF and the K-NMF diminish. When the degree of contamination is over 50%, AUC scores of the R-NMF and the K-NMF are under 0.5, which means that the R-NMF and the K-NMF recognize TEP-free background pixels as target CWA-on background pixels, as explained in Section II-B. Even though the degree of contamination exceeds 50% and the initial model parameter set of the EM algorithm is assigned in reverse, the EM algorithm estimates posterior probabilities correctly via the hypothesis labeling check process described in Section IV-B2. The hard EM-NMF has robust detection performance regardless of the degree of contamination.

6) Detection Performance According to Concentration: Next, we also obtained AUC scores based on the concentration of CWA-on background pixels. In the experiment, the degree of contamination a of the contaminated dataset was 40%, and the concentration path-lengths of the synthetic TEP-on background pixels in the contaminated dataset was set to $\gamma \sim N(c,\ (c/2)^2)$ ppm·m. The mean concentration path-length c was changed from 75 ppm·m to 325 ppm·m at intervals of 25 ppm · m. We acquired the AUC scores as conducting a Monte Carlo simulation whose condition is the same as the experiment in Section V-A5. Fig. 10 shows the corresponding AUC scores. As the concentration of synthetic TEP gas increases, the AUC scores of the K-NMF and the R-NMF are degraded, but not less than 0.5. On the other hand, the detection performances of the I-NMF and EM-NMFs increase since it is easier for the I-NMF

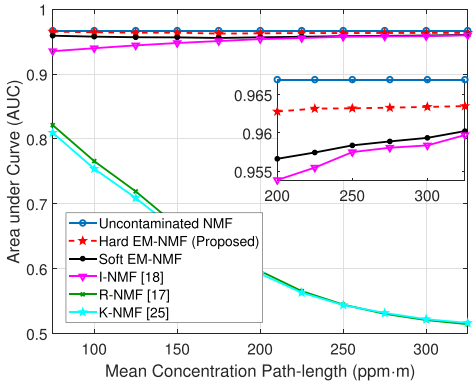

Fig. 10. AUC scores according to the mean concentration path-length of synthetic TEP-on background pixels from 75% to 325%. The right upper figure shows the AUC scores of the uncontaminated NMF, soft and hard EM-NMFs, and I-NMF according to the mean concentration path-length of TEP pixels from 200% to 325%.

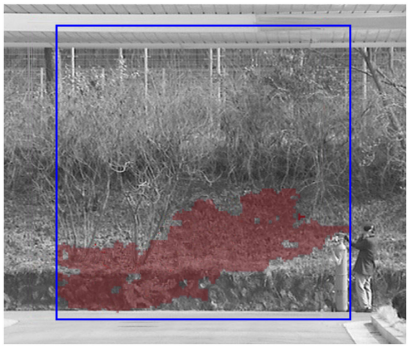

Fig. 11. CCD image of the HSI used in the sprayed gas experiment.

and the EM algorithm to distinguish TEP-on background pixels and TEP-free background pixels. Like other experiments, the hard EM-NMF has better performance than the I-NMF without knowing the degree of contamination. It performs nearly equally to the uncontaminated NMF.

B. Actual Data Experiment

In the experiments with real data, we obtained measurement data for a scenario. The scenario is that SF_6 gas was sprayed into the air with a grass field as the background. The HSI for the scenario was measured a few seconds after spraying of the SF_6 gas. The measurement environment is the same as the synthetic data experiments and the temperature of the SF_6 cloud is 283 K. A blue box in Fig. 11 presents a CCD image of the measured HSI. The SF_6 gas area was obtained by applying an uncontaminated NMF that is designed using the HSI data measured before the SF_6 spray and eliminating some outlier pixels. Red pixels in Fig. 11 indicate the area where the SF_6 gas was present. The degree of contamination, which is the ratio of the red pixel area to the total area, was calculated to be 21.2%. Fig. 12 plots the spectra of the SF_6-on background pixels and the SF_6-free background pixels. The absorption property for the SF_6-on background pixels exists in the 900–1000 cm^{-1} range.

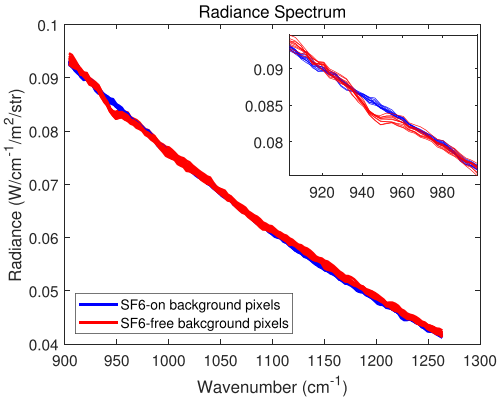

Fig. 12. Spectra of SF_6-on background pixels and SF_6-free background pixels in the sprayed gas experiment. The right upper figure shows the enlarged spectra in the 900−1000 cm^{-1} band where the SF_6 gas has absorption properties.

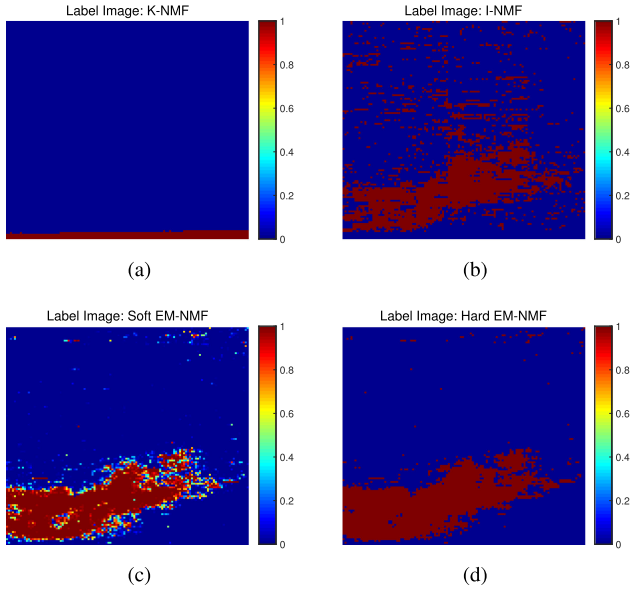

Fig. 13. Label images of (a) the K-NMF, (b) the I-NMF, (c) the soft EM-NMF, and (d) the hard EM-NMF for the measured HSI in the sprayed gas experiment.

Fig. 13 shows label images of the K-NMF, I-NMF, and soft and hard EM-NMFs. The K-means clustering does not extract SF_6-on background pixels at all. There are many false alarm pixels in the I-NMF since background statistics are only exploited in each iteration. Then, SF_6-on background pixels (especially diffused low-concentration pixels) are classified as SF_6-free background pixels by the number of these false alarm pixels. On the other hand, EM-NMFs obtain more accurate SF_6-on background pixels than the I-NMF because both background statistics and SF_6 statistics are used in each iteration. In the soft EM-NMF, the labels of low-concentration SF_6 pixels are in the range of $0.2 - 0.8$. These low-concentration pixels are labeled as SF_6-on background pixels in the hard EM-NMF.

We applied these NMF algorithms to the measured HSI and obtained ROC curves. Fig. 14 shows the ROC curves of several

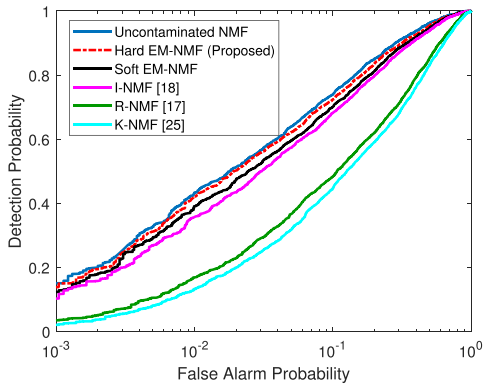

Fig. 14. ROC curves of several NMF algorithms in the sprayed gas experiment.

NMF algorithms. The R-NMF and the K-NMF have poor detection performances. Compared with the uncontaminated NMF, the hard EM-NMF performs as well as the uncontaminated NMF. The hard EM-NMF has slightly better detection performance than the I-NMF, which is in agreement with the result pertaining to the label image analysis results in Fig. 13. In the soft EM-NMF, low-concentration SF_6 pixels distort background statistics as much as $P(H_i|\mathbf{x}_j, \hat{\Theta})$. Because the hard EM-NMF regards these low-concentration SF_6 pixels as SF_6 pixels, the hard EM-NMF has a more robust performance than the soft EM-NMF under the background contamination condition.

C. Summary of Experiments

In the synthetic data experiments, we showed that the EM algorithm calculated posterior probabilities when the proposed initialization method, which provided rough classification results, is used. We also compared the proposed hard EM-NMF with the uncontaminated NMF, R-NMF [17], K-NMF [23] [25], I-NMF [18], and the soft EM-NMF through label images, ROC curves, and AUC scores. The R-NMF had poor detection performance because CWA-on background pixels distort the background mean vector. The K-NMF also showed low detection performance because K-means clustering, which categorized pixels based on the mean value of each pixel's radiance spectrum, did not distinguish pure background pixels from the CWA-on background pixels at all. The I-NMF, soft and hard EM-NMF demonstrated robust detection performance outcomes in condition with background contamination.

The soft and hard EM-NMFs, which exploited both the statistics of estimated CWA-free background pixels and CWA-on background pixels, shoed better detection performance than the I-NMF, which used only the statistics of the estimated CWA-free background pixels, even though we did not know information about the degree of contamination, which is impractical. The detection performance of the proposed hard EM-NMF was slightly better than that of the soft EM-NMF because the estimated background statistics of the hard EM-NMF was not distorted by low-concentration CWA-on background pixels. The hard EM-NMF exhibited a more robust detection performance than the K-NMF, I-NMF and soft EM-NMF regardless of the degree of contamination or the concentrations of the CWA-on background pixels. Through the sprayed gas experiment, we demonstrated that the hard EM-NMF detector could match the performance as the uncontaminated NMF in real environments.

VI. Conclusion

A NMF, which is one the most powerful detectors, is vulnerable to background contamination, which refers to a situation where a background training set contains CWA-on background pixels. In order to mitigate background contamination, we utilized the optimal posterior probabilities by maximizing the log-likelihood of the contaminated dataset using the EM algorithm. We presented a hard EM-NMF constructed with CWA-free background pixels extracted from the contaminated dataset based on the the optimal posterior probability of each pixel. For the fast and accurate convergence characteristics for the EM algorithm, we also proposed an initialization method using a robust NMF. Experimental results confirmed that the proposed hard EM-NMF has a more robust detection performance than other NMFs even without information about the degree of contamination. Therefore, the hard EM-NMF is a useful solution to the problem where the NMF not being feasible for use when background statistics are unavailable.

References

[1] D. Manolakis, S. E. Golowich, and R. S. DiPietro, "Long-wave infrared hyperspectral remote sensing of chemical clouds: A focus on signal processing approaches," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 120–141, Jul. 2014.

[2] E. Truslow, D. Manolakis, M. Pieper, T. Cooley, and M. Brueggeman, "Performance prediction of matched filter and adaptive cosine estimator hyperspectral target detectors," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2337–2350, Jun. 2014.

[3] D. Manolakis, C. Sirasusa, and G. Shaw, "Hyperspectral subpixel target detection using the linear mixing model," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1392–1409, Jul. 2001.

[4] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, "Hyperspectral detection algorithms: Use covariances or subspaces?" in *Proc. SPIE*, vol. 7457, Aug. 2009, Art. no. 74570Q.

[5] Y. Wang, G. Chen, and M. Maggioni, "High-dimensional data modeling techniques for detection of chemical plumes and anomalies in hyperspectral imgages and movies," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4316–4324, Sep. 2016.

[6] J. C. Harsanyi, and C.-I Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.

[7] C.-I Chang, "Orthogonal subspace projection (OSP) revisited: A comprehensive study and analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 502–518, Mar. 2005.

[8] R. Harig and G. Matz, "Toxic cloud imaging by infrared spectrometry: A scanning FTIR system for identification and visulization," *Field Anal. Chem. Technol.*, vol. 5, no. 1–2, pp. 75–90, 2001.

[9] J. H. Lee, J. M. Lee, and Y. I. Kang, "Field identification and spatial determination of hazardous chemicals by fourier transform infrared imaging," *Instrum. Sci. Technol.*, vol. 44, no. 5, pp. 504–520, 2016.

[10] H. G. Yu, J. H. Lee, Y. C. Kim, and D. J. Park, "Intelligent detection algorithm of hazardous gases for FTIR-based hyperspectral imaging system using SVM classifier," in *Proc. SPIE*, vol. 10198, May 2017, Art. no. 1019808.

[11] H. W. Nam, J. S. Kim, H. J. Kim, J. H. Lee, Y. I. Kang, and B. H. Park, "Development of a radiative transfer model for the determination of toxic gases by Fourier transform-infrared spectroscopy with a support vector machine algorithm," *Instrum. Sci. Technol.*, vol. 47, pp. 1–14, 2018.

[12] Y. C. Kim, H. G. Yu, J. H. Lee, D. J. Park, and H. W. Nam, "Hazardous gases detection for FTIR-based hyperspectral imaging system using DNN and CNN," in *Proc. SPIE*, vol. 10433, Oct. 2017, Art. no. 1043317.

[13] L. L. Scharf and L. T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Proc. 30th Asilomar Conf. Signals, Syst., Comput.*, 1996, pp. 1114–1117.
[14] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic," *IEEE Trans. Signal Process.* vol. 53, no. 2, pp. 427–438, Feb. 2005.
[15] J. Theiler, and B. Foy, "Effect of signal contamination in matched-filter detection of the signal a cluttered background," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 98–102, Jan. 2006.
[16] J. Theiler, B. Foy, and A. M. Fraser, "Nonlinear signal contamination effects for gaseous plume detection in hyperspectral imagery," in *Proc. SPIE*, vol. 6233, May 2006, Art. no. 62331U.
[17] S. Niu, S. E. Golowich, V. K. Ingle, and D. Manolakis, "Implications and mitigation of model mismatch and covariance contamination for hyperspectral chemical agent detection," *Opt. Eng.*, vol. 52, no. 2, pp. 1–19, Feb. 2013.
[18] K. E. Kim, S. S. Lee, and H. S. Baik, "Iterative matched filtering for detection of non-rare target materials in hyperspectral imagery," in *Proc. SPIE.*, vol. 10004, Oct. 2016, Art. no. 100040E.
[19] G. E. Thomas and K. Stamnes, *Radiative Transfer in the Atmosphere and Ocean*. Cambridge, U.K.: Cambridge Univ. Press, 2002.
[20] J. P. Hoffbeck and D. A. Landgrebe, "Covariance matrix estimation and classification with limited training data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 7, pp. 763–767, Jul. 1996.
[21] J. Theiler, "Incredible shrinking covariance estimator," in *Proc. SPIE*, May 2012, vol. 8391, Art. no. 83910P.
[22] P. C. Hansen, "The truncatedsvd as a method for regularization," *BIT Numer. Math.*, vol. 27, no. 4, pp. 534–553, Jul. 1987.
[23] N. M. Nasrabadi, "Regularization for spectral matched filter and RX anomaly detector," in *Proc. SPIE*, Apr. 2008, vol. 6966, Art. no. 696604.
[24] E. K. Chong, and S. H. Zak, *An Introduction to Optimization*. 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
[25] P. S. Bradley, U. M. Fayyad, "Refining initial points for k-means clustering," in *Proc. Int. Conf. Mach. Learn.*, Jul. 1998, vol. 98, pp. 91–99.
[26] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed., New York, NY, USA: Wiley, 2001.
[27] R. Hogg, J. McKean and A. Craig, *Introduction to Mathematical Statistics*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005, pp. 359–364.
[28] S. Sabbah, R. Harig, P. Rusch, J. Eichmann, A. Keens, and J. H. Gerhard, "Remote sensing of gases by hyperspectral imaging: System performance and measurements," *Opt. Eng.*, vol. 51, no. 11, Nov. 2012, Art. no. 111717.
[29] P. M. Chu, F. R. Guenther, G. C. Rhoderick, and W. J. Lafferty, "The NIST quantitative infrared database," *J. Res. Nat. Inst. Standards Technol.*, vol. 104, no. 1, pp. 59–81, 1999.
[30] A. Berk *et al.*, "MODTRAN4 radiative transfer modeling for atmospheric correction," in *Proc. SPIE*, vol. 3756, pp. 348–354, Oct. 1999.

Hyeong-Geun Yu (Student Member, IEEE) received the B.S. degree in electronic and communication engineering from Hanyang University, Seoul, South Korea, in 2013, and the M.S. degree in electrical engineering in 2014 from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, where he is currently working toward the Ph.D. degree.

His research interests include hyperspectral image signal processing, statistical signal processing, cognitive radio, and detection theory.

Jai-Hoon Lee received the B.S. degree in electrical and electronic engineering from Yonsei University, Seoul, South Korea, in 2007, and the M.S. and Ph.D. degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2010 and 2019, respectively.

His research interests include hyperspectral image signal processing, communication theory, and energy harvesting.

Dong-Jo Park (Member, IEEE) received the B.S. degree in electrical engineering from Seoul National University, Seoul, South Korea, in 1976 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Los Angeles, CA, USA, in 1981 and 1984, respectively.

From 1984 to 1985, he was a Senior Researcher with the Electronics and Telecommunications Research Institute, Daejeon, South Korea. He is currently an Emeritus Professor with the School of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. His research interests include multimedia signal processing, hyperspectral signal processing, target identification in millimeter-wave and long-wave infrared bands, convex optimization, and machine learning.

Dong Eui Chang received the B.S. degree in control and instrumentation engineering and the M.S. degree in electrical engineering from Seoul National University, Seoul, South Korea, in 1994 and 1997, respectively, and the Ph.D. degree in control and dynamical systems from the California Institute of Technology, Pasadena, CA, USA, in 2002.

He is an Associate Professor with the School of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. His current research interests include control, signal processing, robotics, and deep learning.

Hyunwoo Nam received the Ph.D. degree in electrical engineering from the Department of Electrical Engineering, Columbia University, New York, NY, USA, in October 2016.

He was with the Internet Real-Time Lab under the supervision of Prof. H. Schulzrinne. During Ph.D. years, his collaboration with Bell Labs and Verizon focused on capacity planning for wireless networks, analysis of adaptive bitrate streaming technologies, SDN and NFV for intelligent content delivery over wireless and remote gas sensing system. He is a Senior Researcher with Agency for Defense Development, South Korea. His research interests include the analysis of video streaming and intelligent content delivery over wireless networks.