Haze removal using deep convolutional neural network for Korea Multi-Purpose Satellite-3A (KOMPSAT-3A) multispectral remote sensing imagery

Soohwan Yu [a], Doochun Seo [b], Joonki Paik [a,c,*]

[a] *Department of Image, Chung-Ang University, Seoul, 06974, South Korea*
[b] *Department of Satellite Data Cal/Val Team, Korea Aerospace Research Institute, Daejeon, 34133, South Korea*
[c] *Department of Artificial Intelligence, Chung-Ang University, Seoul, 06974, South Korea*

ARTICLE INFO

ABSTRACT

This paper presents a convolutional neural network to automatically remove the haze distribution using a single multispectral remote sensing image in the raw file format. To train the proposed dehazing network, we synthesized multispectral hazy images using the haze thickness map (HTM) and relative scattering model representing the wavelength-dependent scattering property of the haze distribution. Since the raw multispectral hazy images have a low dynamic range, we cannot accurately estimate the haze distribution directly from them. To differently impose a proper amount of attention to hazy and haze-free regions, we used the HTM from the contrast-enhanced version of the input hazy image. The proposed dehazing network consists of four subnetworks: (i) shallow feature extraction network (SFEN), (ii) cascaded residual dense block network (CRDBN), (iii) multiscale feature extraction network (MFEN), and (iv) refinement network (RN). The densely connected convolutional layers and local residual learning allow the residual dense block (RDB) to extract the abundant local features, and the cascaded architecture further improves the propagation of the local information and gradients. The MFEN is used to extract multiscale local features representing the hierarchical information for the haze distribution and haze-free region. Experimental results demonstrated that the proposed method can achieve improved dehazing performance on Korea Multi-Purpose Satellite-3A (KOMPSAT-3A) multispectral remote sensing imagery without undesired artifacts. In the sense of quantitative assessment, the proposed method produced improved peak signal-to-noise ratio (PSNR) by 10%, structural similarity index measure (SSIM) by 1%, and spectral angle mapper (SAM) by 19% compared with the existing best method.

1. Introduction

As of 2021, more than 80 public domain government agency satellites are active in orbit. Among them, notable Earth observation (EO) satellites include the Landsat series by the National Aeronautics and Space Administration (NASA) (Teluguntla et al., 2018), the Sentinel series by the European Space Agency (ESA) (Guzinski and Nieto, 2019), the Gaofen series by the China National Space Administration (CNSA) (Xie et al., 2020), and the Korea Multi-Purpose Satellite (KOMPSAT) series by the Korea Aerospace Research Institute (KARI) (Jeong et al., 2016), to name a few. Despite their versatility, these satellites have common challenging factors of image quality such as unpredictable atmospheric conditions as well as sensor resolution and precision of registration in spectral, spatial, and temporal domains.

Specifically, atmospheric particles such as water vapor, gas, and particulates scatter the incident sunlight and its reflected version from the earth's surface. Physical properties of the atmospheric particles determine the amount of atmospheric scattering (Kaufman and Sendra, 1988). The atmospheric scattering degrades the visibility of remote sensing imagery with low dynamic range and spectral distortion, and finally results in poor performance of the post-processing step such as object detection and segmentation (Rottensteiner et al., 2014), terrain change detection (Hebel et al., 2013), and terrain classification (Pingel et al., 2013). To solve this problem, various haze removal methods have been developed in the remote sensing field over the past few decades. Unlike normal RGB color images, dehazing of multispectral satellite images may utilize an extended range of wavelengths to obtain a better performance.

The simplest approach, called dark object subtraction (DOS), aims to subtract the additive path radiance by the atmospheric scattering from the total transmitted radiance at the sensor. DOS-based methods assume that a few black pixels such as shadows having 0% reflectance exist within a scene (Vincent, 1972; Chavez, 1988). Specifically, these

---

methods regard the lowest radiance on the dark surface as the haze thickness within a scene, and the atmospheric correction is performed by subtracting the constant or different digital number (DN) values from each spectral band (Chavez, 1988). The early DOS-based atmospheric correction method aims to remove homogeneous haze components. However, these methods cannot avoid over-dehazing and spectral distortion since the dark region may not exist in the acquired scene.

To extract more dark pixels, Kaufman *et al.* proposed a dense dark vegetation (DDV)-based automatic haze thickness detection and correction method (Kaufman and Sendra, 1988). The long wavelength band is less affected than visible spectrum bands by the haze particle and the vegetation index is calculated using red and near-infrared bands (Kaufman et al., 1997). For this reason, Kaufman's method finds the vegetation pixels using the relationship of surface reflectance of visible spectrum bands (blue and red bands) and near- and short-wave infrared bands (Kaufman and Sendra, 1988). Richter *et al.* proposed an improved version of Kaufman's DDV-based atmospheric correction method using the surface reflectance of red and near infra-red (NIR) bands (Richter et al., 2006). DDV-based methods regard the dark vegetation pixels as the dark surface, but this method is sensitive to the assumption of a fraction of the dark vegetation pixels within a scene.

Makarau *et al.* improved DOS-based haze removal method by considering local dark objects (Makarau et al., 2014). Given an input hazy image, Makarau's method estimates the initial haze thickness map (HTM) by searching the local minimum value on the extrapolated spectrum band, which was synthesized using the short wavelength bands of aerosol and blue bands. The haze removal was performed by subtracting the haze thickness map multiplied by the linear regression coefficient, which represents the linear relationship between the initial and band specific HTMs from each spectral band. This method can perform context-aware haze removal while DOS-based methods deal with only a homogeneous haze distribution. However, the regression coefficient should be estimated according to the acquired remote sensing image since it is sensitive to the amount of haze particles within a scene.

Since the dehazing process is highly non-linear and data-dependent, a deep learning-based approach is a promising solution. Recent development in convolutional neural network (CNN)-based model significantly improved the dehazing performance.

Li *et al.* proposed a CNN-based dehazing model called All-in-One Dahzing Network (AOD-Net), which directly generates the haze-removed image through a light-weight end-to-end CNN model instead of separately estimating the transmission map and the atmospheric light (Li et al., 2017). Chen *et al.* proposed Gated Context Aggregation Network (GCA-Net) for dehazing and deraining. The GCA-Net adopted smoothed dilation to remove the artifacts of dilated convolution, and used a gated sub-network to fuse the features from different levels (Chen et al., 2019). Qin *et al.* proposed Feature Fusion Attention Network (FFA-Net) for single image dehazing (Qin et al., 2019). The FFA-Net proposed a feature attention (FA) module that combines channel and pixel attention mechanism, and adopted local residual learning and feature attention as the basic block to bypass less important information such as haze of low-frequency regions. The adaptive feature weights used in FFA-Net gives more weight to important features. As a result the FFA-Net produces the best dehazing result in the sense of image quality and color fidelity among existing methods.

In this paper, we propose a deep convolutional neural network to remove the haze distribution of the multispectral remote sensing image in the raw file format. The contributions of the proposed dehazing method are summarized as follows:

- We present a novel dataset generation method that synthesizes the multispectral hazy image using the scattering coefficient and estimated HTM from the real hazy image to impose different scattering properties according to the wavelength.

- The proposed dehazing network uses the dense connection and residual architecture to fully utilize the local and global information. In addition, the network takes as input the estimated HTM from the contrast-enhanced version of the input multispectral hazy image to impose different attention to the hazy and haze-free regions in the low dynamic range of the raw file format.
- Experiments were conducted on multiscpectral remote sensing images acquired by Korea Multi-Purpose Satellite-3 A (KOMPSAT-3 A), and demonstrated that the estimation of haze distribution instead of directly estimating the dehazed result can achieve better dehazing performance for the multispectral hazy image in the raw file format without undesired artifacts.

Although a clear, haze-free image is unknown in a general situation, a satellite repeatedly acquires the same scene with and without haze. For this reason, existing haze removal methods use synthetic hazy images to assess the haze removal performance in the literature. Therefore, we compared the haze removal performance of the proposed method using the synthetic hazy images in terms of qualitative and quantitative assessments.

The purpose of this research is mainly to eliminate the interference caused by haze in daily changes and information acquisition, not in three-dimensional geometric analysis such as the shape of the terrain. This research was conducted under the guidance of Korea Aerospace Research Institute (KARI), the supporting institution for this research, in order to reflect the professional user requirements such as color preservation and avoiding brightness saturation.

Although remote sensing image should include both spatially varying haze distribution and haze-free region in the real-world case, the correspondingly acquired image in the raw file format looks like that the haze is uniformly distributed due to narrow dynamic range. In addition, when the contrast is enhanced, the hazy remote sensing image is significantly degraded by amplified haze distribution with a severe spectral distortion because the haze distribution usually has high intensity values, which makes the haze removal more difficult. For this reason, this paper presents a deep CNN-based haze removal method to automatically remove the semi-transparent and spatially varying haze distribution using a single multispectral hazy image, while preserving the surface information of the haze-free region.

The remainder of the paper is organized as follows. Section 2 briefly describes the related work and background theory. Sections 3 and 4 respectively present the proposed multispectral hazy image synthesis method and dehazing network. In Section 5, we demonstrate the performance of the proposed dehazing method in the sense of both objective and subjective assessments. Section 6 analyzes the performance of the proposed dehazing method for different combinations of the network input and modules, and Section 7 concludes this paper.

2. Background theory and related work

2.1. Related work

In this subsection, we briefly review existing dehazing methods using: (i) haze optimized transform (HOT), (ii) combined frequency and wavelet, (iii) atmospheric scattering model, and (iv) deep learning.

Zhang *et al.* first introduced HOT to detect the spatial haze distribution using a two-dimensional (2D) spectral space of red and blue bands which are highly correlated under a haze-free condition (Zhang et al., 2002). They demonstrated that a DOS-based haze removal method with the HOT better removed the haze distribution of Landsat images. However, this method requires manually pre-selected haze-free pixels before estimating the HOT, and cannot avoid over-dehazing by spurious components in the HOT. Moro *et al.* minimized the spurious HOT using the mask for a water body and man-made features (Moro and Halounova, 2007). Jiang *et al.* proposed a semi-automatic HOT estimation method by selecting the haze-free pixels based on R-squared and

constant threshold values in a non-overlapping local window (Jiang et al., 2016). Sun *et al.* proposed a fully automated HOT estimation method using an iterative upper trimming linear regression (Sun et al., 2017).

Many researchers have incorporated the atmospheric scattering model into the dehazing of a remote sensing image. The model describes that an observed hazy image is acquired by the sum of an airlight and attenuation irradiances (Narasimhan and Nayar, 2003). The airlight irradiance represents the scattered light integrated into the sensor along the path between the object and the camera. The attenuation irradiance is the transmitted light received by the sensor except for the scattered light by the atmosphere under the haze condition.

Motivated by the DOS-based haze removal approach, He *et al.* proposed the atmospheric scattering model-based haze removal method using the dark channel of an input hazy image (He et al., 2011). This method statistically analyzed the histograms of haze-free non-sky natural images, and derived the dark channel prior model, which assumes that, in a haze-free image, at least one channel among the red, green, and blue channels tends to have a low intensity value close to zero. Based on the dark channel prior and atmospheric scattering model, the haze removal methods for remotely sensed images have been proposed.

Shao *et al.* proposed the linear regression-based transmission estimation method using the haze relevant features of luminance, saturation, and saliency components (Shao et al., 2019). Pan *et al.* analyzed the histograms of haze-free remote sensing and common natural images and derived that the remote sensing images have an offset with a certain degree (Pan et al., 2015). For this reason, this method introduced the additional translation term to the atmospheric scattering model to make the dark channel of the remote sensing image have a low DN close to zero. Li *et al.* described that the haze component is related to the uneven illumination and spatially lower frequency property (Li et al., 2019). They adopted the homomorphic filtering to suppress the haze component and enhanced the detail component, simultaneously. In addition, the sphere model was used to estimate the improved transmission map, which is robust to the additive noise component.

Huang's method estimated the detail component in the frequency domain, and removed the haze component in the spatial domain using multiscale retinex and histogram equalization (Huang et al., 2019). These intermediate results were combined in the wavelet domain. Du *et al.* used the wavelet decomposition to detect the haze distribution using the difference between the low-frequency wavelet bands of the hazy and haze-free images (Du et al., 2002). The haze removal was performed by subtracting the estimated haze distribution from the approximation level of a hazy image. However, this method requires the hazy-free image for the same scene to extract the haze distribution. Ju *et al.* removed the haze component using the estimated scene albedo by adjusting the gamma, but haze removal performance depends on the size of a patch to estimate the dark and bright channels (Ju et al., 2019).

To remove thin cloud, Liu *et al.* proposed the improved multi-scale retinex method to enhance the visibility of the aerial images acquired under the hazy condition (Liu et al., 2017).

In recent years, many deep learning-based haze removal methods have been proposed in the literature. Li *et al.* proposed an encoder–decoder-based network with symmetric concatenation and residual network architecture for real Landsat 8 images (Li et al., 2019). Li *et al.* derived the physical model to decompose the transmission, reflection, and constraint layers of the cloud image and used a generative adversarial network (GAN)-based thin cloud removal method for Sentinel-2 A multispectral imagery (Li et al., 2020).

Gu *et al.* proposed an encoder–decoder-based dense network using the haze density map (Gu et al., 2019; Pan et al., 2016). However, this method is only suitable for the RGB hazy image since the haze density map is estimated by using red, green, and blue bands. Qin *et al.* obtained the dehazed result using the weighted summation of

intermediate haze removed results via multiple residual networks (Qin et al., 2018). Li *et al.* unified the airlight and transmission map of the atmospheric scattering model to a new variable, which is estimated using a convolutional neural network (Li et al., 2017).

Chen *et al.* used a smoothed dilated convolution to extract more contextual information with the instance normalization (Chen et al., 2019; Ulyanov et al., 2017; Wang and Ji, 2018). This method fused the intermediate feature maps to impose the different weights from low- to high-level features. Zhang *et al.* estimated the airlight, transmission map, and dehazed result using U-net and a pyramid-based neural network in a step-by-step manner (Zhang and Patel, 2018; Ronneberger et al., 2015). Qin *et al.* proposed the attention module to differently focus on the uneven haze distribution by imposing different weights to the channel- and pixel-wise feature information (Qin et al., 2019).

2.2. Haze image degradation model

The total radiance at the sensor under cloudless, haze-free conditions can be expressed as (Kaufman and Sendra, 1988)

$$L_i^{\text{sensor}}(x,y) = L_i^D(x,y) + L_i^S(x,y) + L_i^A(x,y), \tag{1}$$

where $L_i^{\text{sensor}}(x,y)$ represents the total radiance of the light integrated into the satellite imaging sensor at a pixel coordinate (x,y), $L_i^D(x,y)$ the path radiance of the sunlight directly transmitted to the sensor by the atmospheric scattering without the surface reflection, $L_i^S(x,y)$ the radiance of the reflected light from the surface and integrated to the sensor through the atmosphere, $L_i^A(x,y)$ the path radiance by the atmospheric scattering of the reflected light from the surface, and i the band number.

In the haze condition, (1) can be rewritten by the sum of the surface reflectance and the scattered path radiance by the haze component (Vincent, 1972; Chavez, 1988; Makarau et al., 2014) as

$$L_i^{\text{sensor}}(x,y) = L_i^S(x,y) + L_i^{haze}(x,y), \tag{2}$$

where $L_i^{haze}(x,y)$ represents the total path radiance of the scattered sunlight and reflected light from the surface by the atmospheric particles. Using the linear relationship between the DN and radiance, (2) can be rewritten as

$$g_i(x,y) = f_i(x,y) + H_i(x,y), \tag{3}$$

where $g_i(x,y)$ represent the DN values of the observed multispectral hazy image, $f_i(x,y)$ the DN values of the latent haze-free image to be estimated, and $H_i(x,y)$ the DN values of an additive haze component for each spectral band.

2.3. Atmospheric scattering

The atmospheric scattering is dependent on the relationship between the size of constituent molecules in the atmosphere, and the wavelength of incident sunlight (Olsen, 2007). To the best of the authors' knowledge, Rayleigh and Mie scattering models are most widely used to describe atmospheric scattering for different atmospheric conditions in the literature. Rayleigh and Mie scattering models describe that the amount of scattering is inversely proportional to the power of the wavelength λ as (Chavez, 1988; Qin et al., 2018)

$$\beta(\lambda) \propto \frac{1}{\lambda^\gamma}, \tag{4}$$

where $\beta(\lambda)$ represents the amount of scattering of wavelength λ, and γ a positive parameter, which determines the atmosphere conditions, and varies in $[0, 4]$. Chavez derived the relative scattering model of varying γ for the different atmosphere conditions (Chavez, 1988). If γ is close to 0, the atmosphere is heavily hazy or cloudy. For the haze-free atmosphere, γ is set to 4 in the Rayleigh scattering model. For the haze condition, since the haze distribution varies by the concentration and size of haze particles, γ is in the range of $[0.5, 0.7]$ by the Mie scattering model.

Table 1

A set of Landsat 8 OLI images used for a real hazy patch extraction.

ID
LC08_L1TP_013032_20180608_20180615_01_T1
LC08_L1TP_013032_20180624_20180704_01_T1
LC08_L1TP_013032_20180726_20180731_01_T1
LC08_L1TP_013032_20190526_20190605_01_T1
LC08_L1TP_013032_20190627_20190705_01_T1
LC08_L1TP_013032_20190915_20190925_01_T1
LC08_L1TP_014031_20180717_20180730_01_T1
LC08_L1TP_014031_20181208_20181226_01_T1
LC08_L1TP_014032_20180919_20180928_01_T1
LC08_L1TP_018045_20181001_20181010_01_T1
LC08_L1TP_018046_20190716_20190721_01_T1
LC08_L1TP_018046_20190918_20190926_01_T1
LC08_L1TP_198026_20190118_20190201_01_T1
LC08_L1TP_198026_20190510_20190521_01_T1
LC08_L1TP_199026_20190415_20190423_01_T1
LC08_L1TP_199026_20190501_20190508_01_T1
LC08_L1TP_200026_20190406_20190422_01_T1
LC08_L1TP_200026_20190422_20190507_01_T1
LC08_L1TP_200026_20190524_20190604_01_T1

Table 2

Spectral bands of Landsat 8 OLI and KOMPSAT-3A.

Spectral bands	Wavelength range (µm)	
	Landsat 8 OLI	KOMPSAT-3A
Band 1 - Coastal/Aerosol	0.433–0.454	·
Band 2 - Blue	0.450–0.515	0.450–0.520
Band 3 - Green	0.525–0.600	0.520–0.600
Band 4 - Red	0.630–0.680	0.630–0.690
Band 5 - Near infrared	0.845–0.885	0.760–0.900
Band 6 - Short wavelength infrared	1.560–1.660	·
Band 7 - Short wavelength infrared	2.100–2.300	·
Band 8 - Panchromatic	0.500–0.680	0.450–0.900
Band 9 - Cirrus	1.360–1.390	·

Landsat 8 OLI images. The haze-free patches were sampled from the KOMPSAT-3 A images without overlapping in the spectral bands 2 to 5.

Based on the degradation model in (3), the multispectral hazy image is synthesized using the wavelength-dependent haze distribution from real hazy patches. The estimated haze distribution of each spectral band is added to the haze-free patch. Given the real hazy patch, the haze distribution of each spectral band is estimated using Makarau's method in three steps: (i) spectrum extrapolation using two shortest wavelength bands, (ii) initial HTM estimation, and (iii) wavelength-dependent HTM generation (Makarau et al., 2014).

The spectrum extrapolation step uses real hazy patches in the shortest wavelength bands expressed as

$$g^{\text{EXT}}(x, y) = g^{B1}(x, y) + (1 - \alpha)g^{B2}(x, y), \tag{5}$$

where $g^{B1}(x, y)$ and $g^{B2}(x, y)$ respectively represent the aerosol/coastal and blue bands of a real hazy patch extracted from Landsat 8 OLI images, $g^{\text{EXT}}(x, y)$ the extrapolated band which contains more haze distribution than the coastal/aerosol band, and we experimentally used $\alpha = 0.95$.

Given the extrapolated band $g^{\text{EXT}}(x, y)$, the HTM, which is regarded as a haze distribution, is estimated as a local minimum of the extrapolated band as

$$g^{\text{HTM}}(x, y) = \min_{(m,n) \in \Omega(x,y)} g^{\text{EXT}}(m, n), \tag{6}$$

where $g^{\text{HTM}}(x, y)$ represents the estimated HTM, $\Omega(x, y)$ the specified window centered at (x, y), and (m, n) the coordinate in $\Omega(x, y)$. The HTM is estimated by searching the minimum value in the local region without overlapping the specified window. The estimated HTM is enlarged using the bilinear interpolation to have the same size as the input hazy image.

Makarau *et al.* estimated the wavelength-dependent HTM for each spectral band by multiplying different linear coefficients to the HTM in (6) (Makarau et al., 2014). However, Makarau's method requires the computation of a linear coefficient for different hazy scenes due to different haze distribution. Qin *et al.* synthesized the wavelength-dependent hazy image using the ratio between the center wavelengths of the coastal/aerosol and other spectral bands and the transmission map (Qin et al., 2018). If the remote sensing imagery does not have the coastal/aerosol band, the transmission map might be inaccurately estimated, and the resulting image is degraded by the radiometric distortion.

To solve that problem, the proposed method uses the scattering coefficient for each spectral band as a linear coefficient to be multiplied by the estimated HTM in (6). Based on the relative scattering model in Chavez (1988), the scattering coefficient for the ith spectral band is computed as

$$k_i = \left(\frac{\lambda_1}{\lambda_i}\right)^{\gamma}, \tag{7}$$

Specifically, when the size of molecules in the atmosphere is much smaller than the wavelength, the scattering of incident sunlight is highly dependent on the wavelength by Rayleigh scattering model. It means that the short wavelength light is scattered more than the longer wavelength light since the size of oxygen and nitrogen molecules is much smaller than the wavelength for the haze-free condition.

On the other hand, Mie scattering model is used to describe the atmospheric scattering in the haze condition. Since the size and concentration of molecules, such as smoke, dust, and water vapor, tend to become larger, the atmospheric scattering is more dependent on the size of haze particles than the wavelength. For these reasons, the different amounts of haze distribution should be added to each spectral band in the multispectral hazy image synthesis. If the same amount of haze distribution is added to the multispectral haze-free image, the correspondingly dehazed image might be degraded by the over-dehazing in the longer wavelength such as red and NIR bands.

3. Multispectral hazy image synthesis

In this section, we present a wavelength-dependent haze simulation method to synthesize artificial multispectral hazy images in detail. Haze distribution in a real multispectral hazy image is determined by the wavelength of each spectral band and the size of constituent molecules in the atmosphere (Olsen, 2007). In other words, the wavelength-dependent hazy image synthesis is required to obtain better dehazing performance.

3.1. Estimation of real haze distribution

To synthesize multispectral hazy images for training datasets, we collected real hazy and haze-free images acquired by Landsat 8 Operational Land Imager (OLI) (USGS, 2022) and KOMPSAT-3 A. The list of Landsat images used for real hazy patches is given in Table 1, and spectral bands of Landsat 8 OLI and KOMPSAT-3 A are given in Table 2.

The size of training data should be large enough to avoid overfitting the model. Although the primary objective of this work is to process KOMPSAT-3 A images, it should work for general satellite images. For this reason, we selected OLI images together with KOMSAT-3 A images to synthesize multispectral hazy images.

An observed real hazy image contains not only a haze distribution but also a haze-free region that should be preserved without over-dehazing or radiometric distortion in the learning-based dehazing process. For that reason, we selected the real image patches including both hazy and haze-free regions in the spectral bands 1 to 5 of
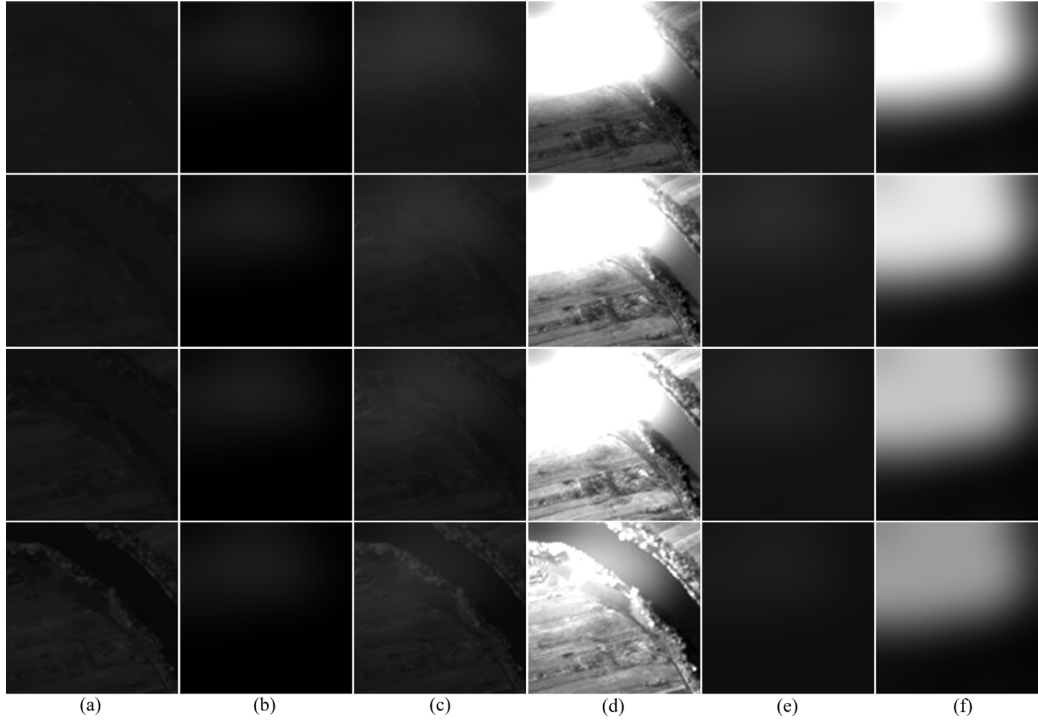
Fig. 1. Four spectral bands of a synthetic hazy image using the real hazy and haze-free scene of Landsat 8 OLI and KOMPSAT-3 A images, respectively: (a) a haze-free scene of KOMPSAT-3 A image, (b) a real haze distribution of Landsat 8 OLI image, (c) a synthetic multispectral hazy image, (d) contrast-enhanced version of (c), and (e)–(f) HTMs estimated from (c) and (d), respectively. At each column, RGB and NIR spectral bands are shown from top to bottom.

where k_i and λ_i respectively represent the scattering coefficient and the center wavelength of the ith spectral band. Given the scattering coefficient k_i, the synthetic multispectral hazy image is defined as

$$g_i(x, y) = f_i(x, y) + k_i g^{\text{HTM}}(x, y), \tag{8}$$

where the second term in the righthand side represents the haze distribution for the ith spectral band.

Figs. 1(a) and 1(b) show the spectral bands of haze-free KOMPSAT-3 A image in the raw file format and real haze distribution estimated from Landsat 8 OLI image, respectively. Fig. 1(c) shows the synthetic multispectral hazy image using (8). As shown in the figure, the amount of haze distribution decreases as the wavelength of spectral bands increases, which proves that the proposed method can synthesize the wavelength-dependent hazy image for different hazy scenes.

4. Proposed multispectral dehazing method

4.1. Pre-processing

This subsection describes the pre-processing step to generate the network input and the proposed dehazing network for a multispectral hazy image in the raw file format.

4.1.1. Contrast enhancement

Since a raw multispectral hazy image has an extremely low dynamic range, it is difficult to accurately separate the hazy and haze-free regions without pre-processing. Fig. 2 shows histograms of spectral bands 2 to 5 of KOMPSAT-3 A imagery. As shown in the figure, we cannot exactly discriminate the hazy and haze-free regions from a low dynamic range image.

To solve this problem, we first estimate the wavelength-dependent HTM of each spectral band from the contrast-enhanced version of the input multispectral hazy image. Next, we take the estimated HTM as the input of the proposed dehazing network to impose more attention on the haze distribution than the haze-free region.

The proposed method enhances the contrast by stretching the histogram, which is a straightforward way to enhance the contrast by linearly spanning the histogram using the upper and lower threshold values as

$$g_{\text{CE}}(x, y) = L \left(\frac{g(x, y) - t_{min}}{t_{max} - t_{min}} \right), \tag{9}$$

where $g(x, y)$ and $g_{\text{CE}}(x, y)$ respectively represent the input and its contrast-enhanced version using the histogram stretching method, t_{min} and t_{max} respectively the lower and upper threshold values, L the maximum level of the bit range. t_{min} and t_{max} are respectively set by the DN values at a certain percentage of the top brightest and darkest pixels of an input image using the cumulative distribution function (CDF). The DN values are clipped into the linearly scaled range in $[0, L]$.

However, if a bright object such as a cloud exists in an input hazy image, the histogram congregates to the bright region, and t_{max} becomes larger. As a result, the histogram stretching method more enhances the brightness and contrast of the cloud regions than the moderate haze distribution. To solve this problem, the proposed method estimates the histogram using the binary mask to exclude the bright objects. The binary mask is estimated using the mean of the averaged spectral band as

$$T(x, y) = \begin{cases} 1, & \text{if } M(x, y) < \mu_M \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

where

$$M(x, y) = \frac{1}{N} \sum_{i \in \{R, G, B, N\}} g_i(x, y), \tag{11}$$

$T(x, y)$ and $M(x, y)$ respectively represent the binary mask and averaged spectral band images, μ_M the mean value of $M(x, y)$, and N the total number of spectral bands of the input multispectral hazy image. The histogram of the i-th spectral band of the input multispectral hazy image is computed as

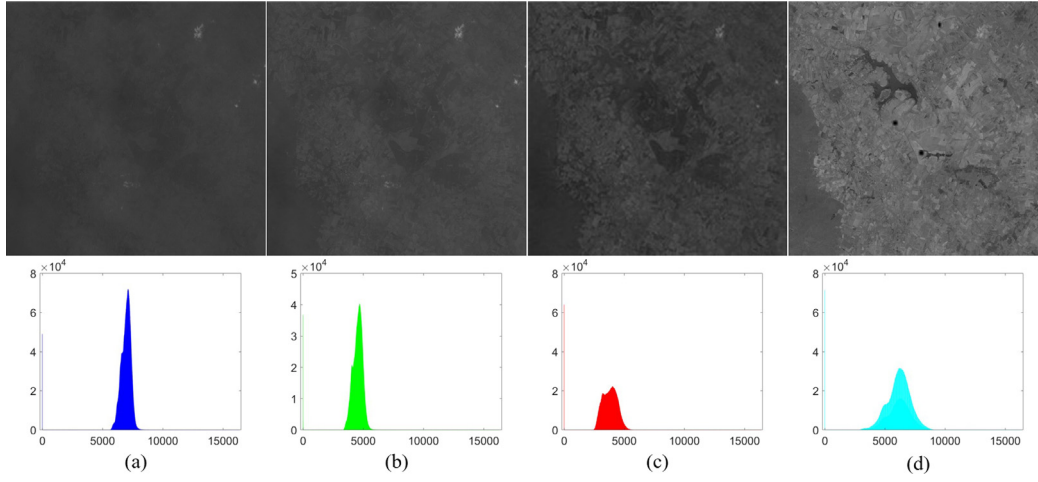$$h_i(k) = \mathbf{card} \left\{ (x, y) | g_i(x, y) = k, T(x, y) = 1 \right\}, \tag{12}$$

Fig. 2. Spectral bands and their histograms of a real multispectral hazy image acquired by KOMPSAT-3 A earth observation satellite: (a) blue, (b) green, (c) red, and (d) NIR bands.
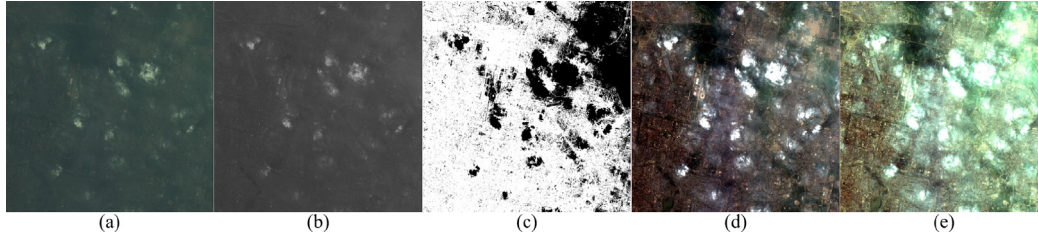

Fig. 3. Comparative results of the contrast enhancement using the histogram stretching method: (a) an input multispectral hazy image acquired by KOMPSAT-3 A, (b) an averaged spectral band $M(x, y)$, (c) a binary mask $T(x, y)$, (d) result of a conventional histogram stretching method, and (e) result of the proposed histogram stretching method. t_{max} and t_{min} were set by the top 0.01% brightest and darkest DN values.
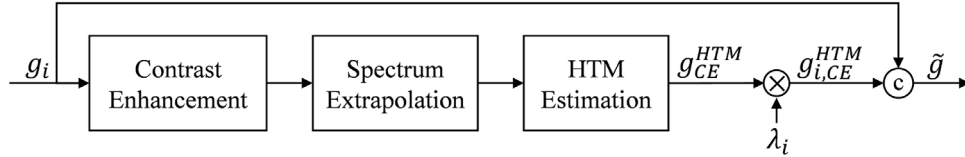

Fig. 4. The block diagram of the network input generation process.

where $\mathbf{card}\{\cdot\}$ represents the cardinality of a set. The dynamic range of multispectral hazy image is enhanced using (9) and the CDF computed from the histogram h_i for $k \in [0, L-1]$. The proposed method can better enhance the contrast than the conventional histogram stretching method for the same t_{min} and t_{max} as shown in Fig. 3.

4.1.2. HTM estimation

As mentioned above, the estimation of the haze distribution may become inaccurate under a low-contrast condition without any pre-processings. To solve this problem, the proposed method estimates the HTM from the contrast-enhanced version of an input multispectral hazy image to give more attention to the haze distribution. Since the KOMPSAT-3 A imagery does not have a coastal/aerosol band as shown in Table 2, the spectrum extrapolation is performed using blue and green bands of the contrast-enhanced input hazy image, respectively. Given the extrapolated band, the proposed method estimates the HTM using (6) as shown in Fig. 4. We use a guided filter to provide the locally constant property of the HTM using the guidance image whose intensity values are all zero (He et al., 2013).

4.1.3. Network input

Let g_{CE}^{HTM} be the HTM of the contrast-enhanced version of the input hazy image. The wavelength-dependent HTM for the i-th spectral band

is estimated by multiplying the scattering coefficient, λ_i, of (7) to the estimated HTM, g_{CE}^{HTM}, as shown in Fig. 4. Fig. 1(d) shows the contrast-enhanced version of a synthetic multispectral hazy image of Fig. 1(c). Figs. 1(e) and 1(f) respectively show the estimated HTMs from Figs. 1(c) and 1(d) for each spectral band. As shown in Fig. 1(f), since the wavelength-dependent HTMs of the contrast-enhanced version have higher DN values than the haze-free region, the proposed dehazing network can give different attention to the hazy and haze-free regions.

The input multispectral hazy image in the raw file format and estimated HTM are concatenated to make $\widetilde{g}$, which is the input of the proposed dehazing network as

$$\widetilde{g} = \left[g_i, g_{i,CE}^{HTM} \right], \tag{13}$$

where g_i represents the ith spectral band of a multispectral hazy image in the raw file format, and $g_{i,CE}^{HTM}$ the wavelength-dependent HTM of the ith spectral band estimated from the contrast-enhanced version of g_i. We will compare the dehazing performance of the proposed method with and without HTM in Section 6.2.

4.2. Network architecture

In this subsection, we describe the proposed dehazing network using the residual learning between an input multispectral hazy image and its
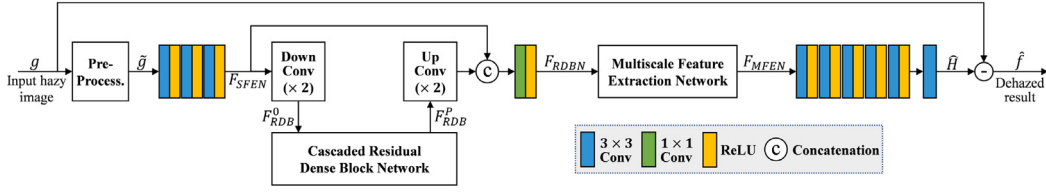
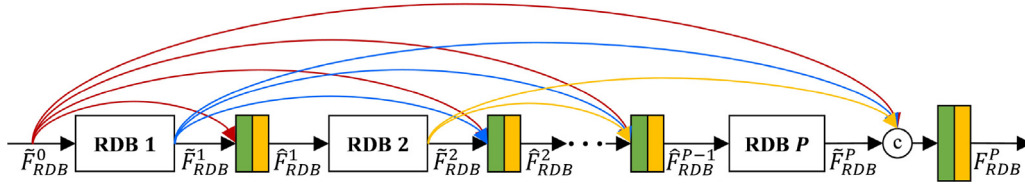**Fig. 5.** The architecture of the proposed dehazing network.

Fig. 6. The inside details of the CRDBN in Fig. 5.

haze distribution. In the proposed method, an observed multispectral hazy image is obtained by adding the haze distribution to the haze-free image. It means that we can predict the latent haze-free image by subtracting the estimated haze distribution from an observed multispectral hazy image.

For the denoising problem, Zhang *et al.* demonstrated that the residual mapping, which learns the relationship between an input noisy image and its noise component, can achieve better performance than estimating the denoised image (Zhang et al., 2017). In the same manner, we adopt the residual learning to estimate the haze distribution $H_i(x, y) \approx k_i g^{\text{HTM}}(x, y)$ in (8) from an input multispectral hazy image instead of estimating the resulting dehazed image.

As shown in Fig. 5, the architecture of the proposed dehazing network consists of four sub-networks: (i) shallow feature extraction network (SFEN), (ii) cascaded residual dense block network (CRDBN), (iii) multiscale feature extraction network (MFEN), and (iv) refinement network (RN). The SFEN extracts the shallow feature maps from the network input $\widetilde{g}$ in (13) as

$$F_{\text{SFEN}} = \mathcal{W}_{\text{SFEN}}(\widetilde{g}), \tag{14}$$

where $\mathcal{W}_{\text{SFEN}}(\cdot)$ represents three 3×3 convolutional layers. The shallow feature F_{SFEN} becomes the input of the cascaded RDBN.

The second sub-network is the CRDBN, which combines the local residual learning and dense skip connections of convolutional layers (Zhang et al., 2018b). Since the densely connected layers become the shortcut connections to propagate the feature information and local gradients, we can achieve better performance in the image restoration problem (Zhang et al., 2018b; Ahn et al., 2018; Huang et al., 2017).

Unlike the original RDBN, the proposed CRDBN does not use the global residual learning between the input and output features. Instead, the proposed dehazing network has the dense connectivity to fully utilize the locally extracted features from RDBs as shown in Fig. 6. Since all the RDBs are densely connected, it can guarantee a further improvement in the flow of the information and local gradient in both forward and backward propagations. In addition, we added a 1×1 convolutional layer before the 3×3 convolutional layer in the RDB to further reduce the network parameters than the original RDBN as shown in Fig. 7.

To estimate the global features representing the hazy and haze-free region, the proposed CRDBN uses the compressed feature map. The input of the first RDB is obtained by compressing the shallow feature using a 3×3 convolution layer with a stride of 2 as

$$\widetilde{F}_{\text{RDB}}^0 = \mathcal{W}_{\text{RDB}}^D(F_{\text{SFEN}}), \tag{15}$$

where $\widetilde{F}_{\text{RDB}}^0$ is the compressed version of F_{SFEN} and becomes the input of the proposed CRDBN, and $\mathcal{W}_{\text{RDB}}^D(\cdot)$ the down-sampling layer to reduce the spatial resolution using the strided convolution. Consequently, the proposed method can utilize the larger receptive field in RDBs.

Let $\widetilde{F}_{\text{RDB}}^p$ be the output feature map of the pth RDB for $p = 1, \ldots, P$, then $\widetilde{F}_{\text{RDB}}^p$ can be formulated as

$$
\begin{aligned}
\widetilde{F}_{\text{RDB}}^p &= \mathcal{W}_{\text{RDB}}^p\left(\widetilde{F}_{\text{RDB}}^0\right) \\
&= \mathcal{W}_{\text{RDB}}^p\left(\mathcal{W}_{\text{RDB}}^{p-1}\left(\cdots\left(\mathcal{W}_{\text{RDB}}^1\left(\widetilde{F}_{\text{RDB}}^0\right)\right)\cdots\right)\right) \\
&= \mathcal{W}_{\text{RDB}}^p\left(\mathcal{W}_{\text{LFF}}^p\left(\left[\hat{F}_{\text{RDB}}^{p-1}, F_{\text{RDB},1}^p, \ldots, F_{\text{RDB},C}^p\right]\right) + \hat{F}_{\text{RDB}}^{p-1}\right),
\end{aligned}
\tag{16}
$$

where

$$\hat{F}_{\text{RDB}}^{p-1} = \mathcal{T}\left(\left[\widetilde{F}_{\text{RDB}}^0, \widetilde{F}_{\text{RDB}}^1, \ldots, \widetilde{F}_{\text{RDB}}^{p-1}\right]\right), \tag{17}$$

$\mathcal{W}_{\text{RDB}}^p$ represents the pth RDB, $\mathcal{W}_{\text{LFF}}^p$ the local feature fusion using 1×1 convolutional layer, $\hat{F}_{\text{RDB}}^{p-1}$ the input feature map of the pth estimated RDB using the output feature maps of the preceding RDBs and $\widetilde{F}_{\text{RDB}}^0$, $\left[\hat{F}_{\text{RDB}}^{p-1}, F_{RDB,1}^p, \ldots, F_{RDB,C}^p\right]$ the concatenation of $\hat{F}_{\text{RDB}}^{p-1}$ and the intermediate feature maps of the sequences of 1×1 and 3×3 convolutional layers. The transition layer $\mathcal{T}$ using a 1×1 convolution reduces the dimensionality of the depth of the concatenated feature map.

Since the proposed CRDBN uses the dense connectivity of all the RDBs instead of global residual learning, the output feature map of the CRDBN is estimated using the concatenation of output feature maps of all the RDBs and the input feature map $\widetilde{F}_{\text{RDB}}^0$. Let F_{RDB}^P be the output feature map of the CRDBN, then F_{RDB}^P can be expressed as

$$F_{\text{RDB}}^P = \mathcal{T}\left(\left[\widetilde{F}_{\text{RDB}}^0, \widetilde{F}_{\text{RDB}}^1, \ldots, \widetilde{F}_{\text{RDB}}^{P-1}, \widetilde{F}_{\text{RDB}}^P\right]\right). \tag{18}$$

To restore the compressed feature map of the CRDBN, the proposed dehazing network enlarges the feature map F_{RDB}^P using the up-sampling layer. In addition, the up-sampled feature map is adaptively fused with the shallow feature map F_{SFEN} using the transition layer $\mathcal{T}$ as

$$F_{\text{RDBN}} = \mathcal{T}\left(\left[F_S, \mathcal{W}_{\text{RDB}}^U\left(F_{\text{RDB}}^P\right)\right]\right), \tag{19}$$

where F_{RDBN} represents the final output feature map of the CRDBN, $\mathcal{W}_{\text{RDB}}^U(\cdot)$ represents the up-sampling layer using bilinear interpolation followed by a 3×3 convolutional layer.

Since an observed multispectral hazy image contains both hazy and haze-free regions, the proposed dehazing network estimates an accurate haze distribution using the multiscale feature extraction network (MFEN). In the earlier work of object detection and semantic segmentation, the multiscale architectures were used to extract scale-invariant features using the convolutional and pooling layers in series (Lin et al., 2017; Zhao et al., 2017; Zhang et al., 2018a). The extracted features hierarchically represent the local and global low-level information for the same object.

Based on this observation, the proposed method uses the multiscale architecture to extract the hierarchical features from the output feature
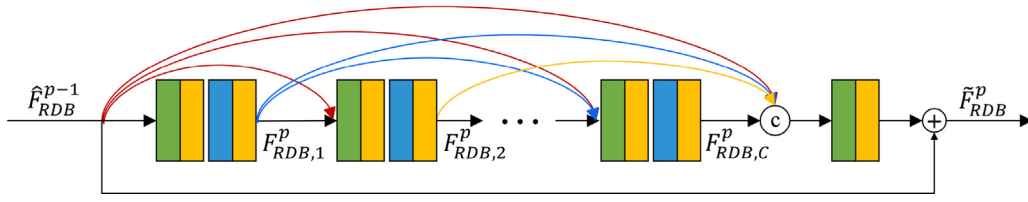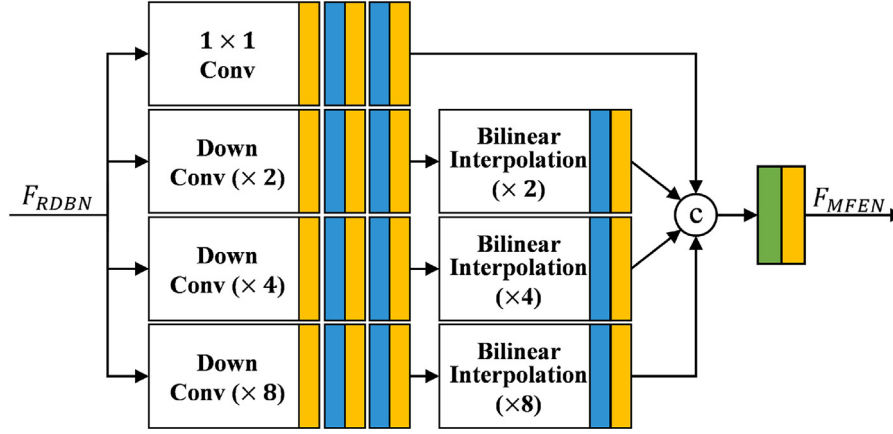
Fig. 7. The inside details of the *p*-th RDB in Fig. 6.

Fig. 8. The architecture of the proposed multiscale feature extraction network (MFEN).

of the CRDBN at different scales having different receptive fields. Given the output feature of the CRDBN, the proposed MFEN can be formulated as

$$F_{\mathrm{MFEN}}^{s} = \mathcal{W}_{\mathrm{MFEN}}^{U,s}\left(\mathcal{W}_{\mathrm{MFEN}}^{D,s}\left(F_{\mathrm{RDBN}}\right)\right), \tag{20}$$

where F_{MFEN}^{s} represents the output feature of the MFEN, for $s \in \{0, 1, 2, 3\}$, $\mathcal{W}_{\mathrm{MFEN}}^{D,s}(\cdot)$ the down-sampling layer compressing the input feature, $\mathcal{W}_{\mathrm{MFEN}}^{U,s}(\cdot)$ the up-sampling layer enlarging the compact and coded feature, and s the scale level. The size of the stride is calculated by 2^{s} and it becomes the scale factor to compress and enlarge the spatial resolution of the feature map at the down- and up-sampling layers, respectively.

$\mathcal{W}_{\mathrm{MFEN}}^{D,s}(\cdot)$ consists of a strided convolution and two 3×3 convolutional layers. $\mathcal{W}_{\mathrm{MFEN}}^{U,s}(\cdot)$ consists of bilinear interpolation followed by a 3×3 convolutional layer. For $s = 0$, the proposed MFEN performs a 1×1 and two 3×3 convolutions without the down- and up-sampling layers. The output feature of the MFEN is obtained by concatenating the hierarchical feature maps followed by a 1×1 convolutional layer as shown in Fig. 8.

Next, the output feature of MFEN is fed into the refinement network (RN) to estimate the haze distribution more accurately as

$$F_{\mathrm{RN}} = \mathcal{W}_{\mathrm{RN}}\left(F_{\mathrm{MFEN}}\right), \tag{21}$$

where F_{RN} represents the output feature of the RN, and $\mathcal{W}_{\mathrm{RN}}$ consists of five 3×3 convolutional layers.

Finally, the output haze distribution is estimated using a 3×3 convolutional layer. Let $\hat{H}_i(x, y)$ be the estimated haze distribution of the ith spectral band. The dehazed result is obtained by subtracting the estimated haze distribution $\hat{H}_i(x, y)$ from an input multispectral hazy image g_i as

$$\hat{f}_i(x, y) = g_i(x, y) - \hat{H}_i(x, y), \tag{22}$$

where $\hat{f}_i(x, y)$ and $g_i(x, y)$ respectively represent the ith spectral band of the dehazed and input multispectral hazy images.

4.3. Loss function

The proposed dehazing network is trained by minimizing the difference between the predicted haze distribution and its latent version. Given a training dataset of N synthetic multispectral hazy images $\{g_i\}$ and latent haze-free images $\{f_i\}$, an ℓ_2 loss function is used to update the trainable network parameters. In addition, we adopted an ℓ_2-regularization term to prevent overfitting in the training process. The objective function of the proposed dehazing method is defined as

$$\ell(\theta) = \frac{1}{N} \sum_{i=1}^{N} \left\| (g_i - f_i) - X\left(g_i; \theta_X\right) \right\|_F^2 + \frac{\lambda}{2} \left\| \theta_X \right\|_2^2, \tag{23}$$

where N is the number of the training dataset, $(g_i - f_i)$ the real haze distribution to be estimated, $X(\cdot)$ the proposed dehazing network, λ the regularization parameter, and θ_X the set of trainable parameters of the proposed dehazing network.

5. Experimental results

5.1. Training dataset

To synthesize the multispectral hazy image, we extracted pairs of the real hazy and haze-free patches from selected Landsat 8 OLI and KOMPSAT-3 A images. Given Landsat 8 OLI images, we manually selected 661 cases of the real haze distributions of size 512×512 including the haze-free region from the coastal/aerosol and blue bands. In addition, we extracted about 10,000 haze-free images of size 512×512 from blue, red, green, and near-infrared bands. The extracted real haze distribution and haze-free images were randomly cropped to the size of 128×128, augmented via vertical and horizontal flips, and then transposed during the training phase. The intensity of selected images was normalized to [0, 1] by dividing the peak intensity value, 2^B, where B represents the bit depth. All the multispectral remote sensing images of KOMPSAT-3 A were provided by the Korea Aerospace Research Institute for only non-commercial use.

However, since Landsat 8 OLI and KOMPSAT-3 A images have 16-bit and 14-bit ranges, the synthetic hazy image only shows the uniform

Table 3
Hyper-parameters used in the experiment.

Hyper-parameters	Values
Adam optimizer	$\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$
Learning rate	0.0001 decayed by 0.99 for every epoch
Number of minibatches	64
Number of epochs	30
Regularization parameter	$\lambda = 0.0001$

haze distribution. To impose the non-uniformity and to match the dynamic range between the haze distribution and haze-free images, we adopted the conventional histogram stretching of (9) to the extracted real haze distribution, where t_{min} and t_{max} were set to 0.0001 and 0.9999, respectively. In addition, we multiplied the scale factor, which was randomly selected in $[0.1, 0.2]$, to match the range of DN values of the extracted real hazy image with that of the haze-free image.

To generate the training dataset, the parameter γ in (7), which determines the amount of haze distribution in each spectral band, was randomly set in $[0.5, 1.5]$ for the multispectral hazy image synthesis. In the pre-processing step, the parameter γ was fixed to 1 to generate the network input. Among the training dataset, 9,000 images were randomly selected for training, and the rest were used for validation.

5.2. Implementation details

In the proposed dehazing network, we used 10 RDBs and three pairs of 1×1 and 3×3 convolutional layers. The growth rate of the RDB was set to 32. In addition, we used 32 feature maps and the rectified linear unit (ReLU) activation function was used after all convolutional layers except the output convolutional layer (Nair and Hinton, 2010). The network parameters were initialized using He's method and updated using Adam optimizer with parameters of $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$, and the learning rate of 0.0001 (Kingma and Ba, 2015; He et al., 2015). The learning rate was decayed by multiplying 0.99 after every epoch. The minibatch size was set to 64, and the number of epochs was set to 30. The regularization parameter λ was set to 0.0001. The proposed dehazing network was implemented with Pytorch framework and trained using two NVIDIA Titan RTX GPUs.

The list of hyper-parameters to train the proposed model is summarized in Table 3.

5.3. Objective assessments

In this subsection, we demonstrate objective performance of the proposed method using PSNR, SSIM, and SAM (Wang et al., 2004; Yuhas et al., 1992). Given a reference image f and a test image $\hat{f}$, both of size $M \times N$, the PSNR between B bits gray-level images f and $\hat{f}$ is defined as

$$\text{PSNR}(f, \hat{f}) = 10 \log_{10} \frac{(2^B - 1)^2}{\text{MSE}(f, \hat{f})}, \tag{24}$$

where

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (f_{ij} - \hat{f}_{ij})^2. \tag{25}$$

The SSIM is used to measure the similarity in terms of luminance, structure, and contrast between an input and reference images. It is defines as

$$\text{SSIM}(f, \hat{f}) = l(f, \hat{f}) c(f, \hat{f}) s(f, \hat{f}), \tag{26}$$

where

$$l(f, \hat{f}) = \frac{2\mu\hat{\mu} + C_1}{\mu^2 + \hat{\mu}^2 + C_1}$$

$$c(f, \hat{f}) = \frac{2\sigma\hat{\sigma} + C_2}{\sigma^2 + \hat{\sigma}^2 + C_2} \tag{27}$$

$$s(f, \hat{f}) = \frac{\sigma_{f\hat{f}} + C_3}{\sigma\hat{\sigma} + C_3},$$

μ and $\hat{\mu}$ respectively represent mean luminance of f and $\hat{f}$, σ and $\hat{\sigma}$ respectively represent standard deviations of f and $\hat{f}$, and $\sigma_{f\hat{f}}^2$ is the covariance between f and $\hat{f}$. C_1, C_2, C_3 are positive constants to avoid a null denominator (Hore and Ziou, 2010).

The SAM measures the angular difference between an input and reference spectral vectors. The SSIM value close to 1 and higher PSNR value indicate better image quality, and the SAM value close to 0 indicates better spectral similarity. Synthetic multispectral hazy images of size 512×512 were used to evaluate the dehazing performance of the proposed method.

The performance of the proposed method was compared with He's method (He et al., 2011), Tarel's method (Tarel and Hautière, 2009), Pan's method (Pan et al., 2015), Zhu's method (Zhu et al., 2015), AODNet (Li et al., 2017), GCANet (Chen et al., 2019), and FFANet (Qin et al., 2019). AODNet, GCANet, and FFANet were trained using the training dataset of synthetic multispectral hazy images generated by the proposed method for red-green-blue (RGB) and near-infrared-green-blue (NGB) bands. The hyperparameters were set as described in the original works of AODNet, GCANet, and FFANet.

Specifically, the network parameters were initialized using He's method (He et al., 2015) for AODNet, GCANet, and FFANet, respectively. The trainable parameters were updated using Adam optimizer, and AODNet was trained without the gradient clipping. The patch size of the synthetic multispectral hazy image was set to 128×128, which is the same as the proposed method. The epochs were set to 150, 50, and 30 for AODNet, GCANet, and FFANet, respectively. The number of feature maps of the input and output convolutional layers was set to the number of spectral bands.

Figs. 9 and 10 show the resulting dehazed images using the proposed and existing dehazing methods for synthetic multispectral hazy images. In those figures, the odd rows show the input hazy and dehazed images in the raw file format. The even rows show the contrast-enhanced version of the dehazed images using the conventional histogram stretching method where t_{min} and t_{max} were respectively set to 0.0001 and 0.9999.

The dehazing performance of atmospheric scattering model-based methods depends on the transmission and atmospheric light estimation. These methods estimate the transmission using the dark channel prior, which assumes that one of the intensity values in red, green, and blue pixels is very close to 0. However, since the multispectral hazy image has the offset with a certain degree as described in Pan et al. (2015), it is difficult to estimate the accurate transmission from the dark channel. Consequently, the resulting dehazed image cannot avoid the over-dehazing and spectral distortion due to an inaccurately estimated transmission as shown in He's and Pan's methods.

Tarel's method based on the visibility restoration function also shows the color and contrast distortion since the contrast enhancement and tone mapping are performed in the dehazing process as shown in Figs. 9(d) and 9(d). Zhu's method reduces the haze distribution using the color attenuation prior (CAP), under the assumption that the saturation and luminance components of a hazy image are inversely proportional to each other as the distance from the camera increases with the loss of color information. Based on this observation, the CAP removes the haze distribution using the depth information estimated from the saturation and luminance components. However, the remote sensing image has a constant depth since the distance between the earth's surface and the satellite is very far. For this reason, the CAP-based method cannot effectively remove the haze distribution as shown in Figs. 9(f) and 10(f).

As a deep learning-based approach, Li's method combined the transmission and atmospheric light as a single parameter to be estimated by CNN. As shown in Figs. 9(g) and 10(g), AODNet provided better dehazing performance than non-deep learning-based methods as described above, but this method shows limited dehazing performance
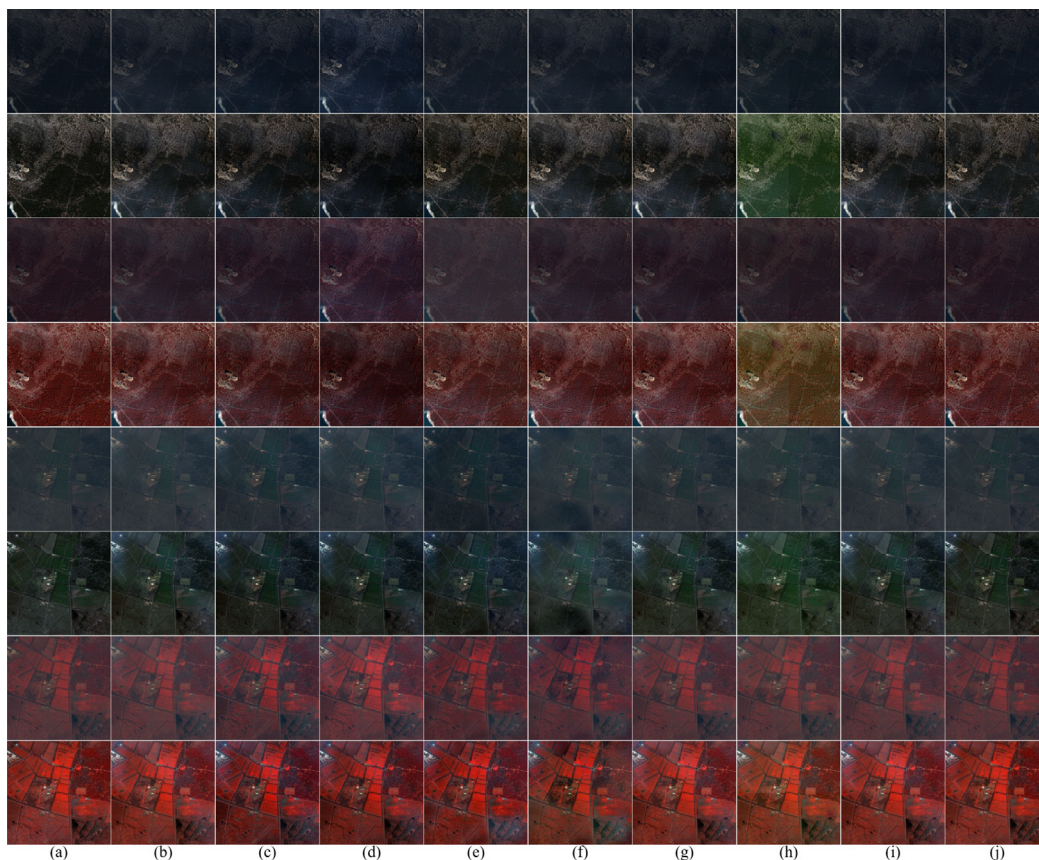
Fig. 9. Comparative results of the dehazing performance on RGB and NGB bands of synthetic multispectral hazy images: (a) haze-free images, (b) synthetic hazy images, (c) He's method (He et al., 2011), (d) Tarel's method (Tarel and Hautière, 2009), (e) Pan's method (Pan et al., 2015), (f) Zhu's method (Zhu et al., 2015), (g) AODNet (Li et al., 2017), (h) GCANet (Chen et al., 2019), (i) FFANet (Qin et al., 2019), and (j) the proposed method ($\gamma = 0.1$, $t_{min} = 0.1$, $t_{max} = 0.9$).

Table 4

Objective assessments of the dehazing performance of the proposed and existing methods on the synthetic multispectral hazy images using PSNR, SSIM, and SAM.

	Input	He's method	Tarel's method	Pan's method	Zhu's method	Li's method	Chen's method	Qin's method	Proposed
PSNR (dB)	28.8945 ± 3.8667	31.0961 ± 3.5645	24.7168 ± 4.6247	29.4150 ± 3.3948	28.6588 ± 4.6985	32.3628 ± 4.5641	32.5849 ± 5.0589	31.2741 ± 4.5387	$\mathbf{35.7425 \pm 4.6089}$
SSIM	0.9615 ± 0.0433	0.9665 ± 0.0251	0.9056 ± 0.0781	0.9339 ± 0.0676	0.9440 ± 0.0731	0.9767 ± 0.0320	0.9745 ± 0.0312	0.9730 ± 0.0345	$\mathbf{0.9866 \pm 0.0198}$
SAM	2.7093 ± 1.3483	3.3188 ± 1.1074	3.8565 ± 2.0142	7.6184 ± 4.3301	4.1397 ± 3.3439	2.0507 ± 1.0479	2.7080 ± 1.1086	2.1963 ± 1.1782	$\mathbf{1.6673 \pm 0.9341}$

for a moderate haze distribution compared to the proposed method. Although the GCANet gives better dehazing performance than AOD-Net, it cannot completely avoid an imbalanced spectral amplification. Furthermore, the resulting images of GCANet show the discontinuity in mean brightness and spectral preservation between the blocks because of the instance normalization as shown in Figs. 9(h) and 10(h).

Qin's method utilizes the attention modules, which impose totally different weights on the features in both channel- and pixel-wise manners, to deal with unevenly distributed haze particles. As shown in Figs. 9(i) and 10(i), FFANet provided improved dehazed results than AODNet and GCANet. However, Qin's method cannot effectively remove the haze distribution compared with the proposed method.

On the other hand, since the wavelength-dependent HTM for each spectral band estimated from the contrast-enhanced input hazy image gives different attention to the hazy and haze-free regions, the resulting images of the proposed method show better dehazing performance than existing dehazing methods without over-dehazing and spectral distortion as shown in Figs. 9(j) and 10(j). In addition, the proposed dehazing method can remove the haze distribution without the discontinuity between blocks for a very high-resolution multispectral hazy image. Fig. 11 shows the boxplots of PSNR, SSIM, and SAM values of the proposed and existing dehazing methods summarized in Table 4.

We also performed statistical analysis of the differences in PSNR, SSIM, and SAM values between the proposed and existing dehazing

methods using a paired *t*-test (McDonald, 2014). In each paired *t*-test, the null hypothesis assumed that PSNR, SSIM, and SAM values have no significant difference for a pair of the proposed and existing dehazing methods. On the other hand, the alternative hypothesis was that PSNR, SSIM, and SAM values between the proposed and existing dehazing methods are significantly different. The null hypothesis is rejected for $p < 0.05$.

Table 5 shows the comparison of the *p*-values of a paired *t*-test for a pair of the proposed and existing dehazing methods. As summarized in Table 5, the low *p*-values over all existing methods indicate that the proposed method provided better dehazing performance for multispectral hazy images in the raw file format. It also implies that the proposed dehazing method achieved significantly higher PSNR, SSIM, and lower SAM values with statistical significance than existing dehazing methods.

5.4. Subjective assessments

In this subsection, we evaluate the dehazing performance of the proposed method on real multispectral hazy images of size 6000×6000 acquired by KOMPSAT-3 A. The experiments were conducted on two NVIDIA Titan RTX GPUs. Since the proposed network cannot infer the dehazed output of size 6000×6000 due to the physical limitation of GPU memory, we obtained the resulting image using block-based
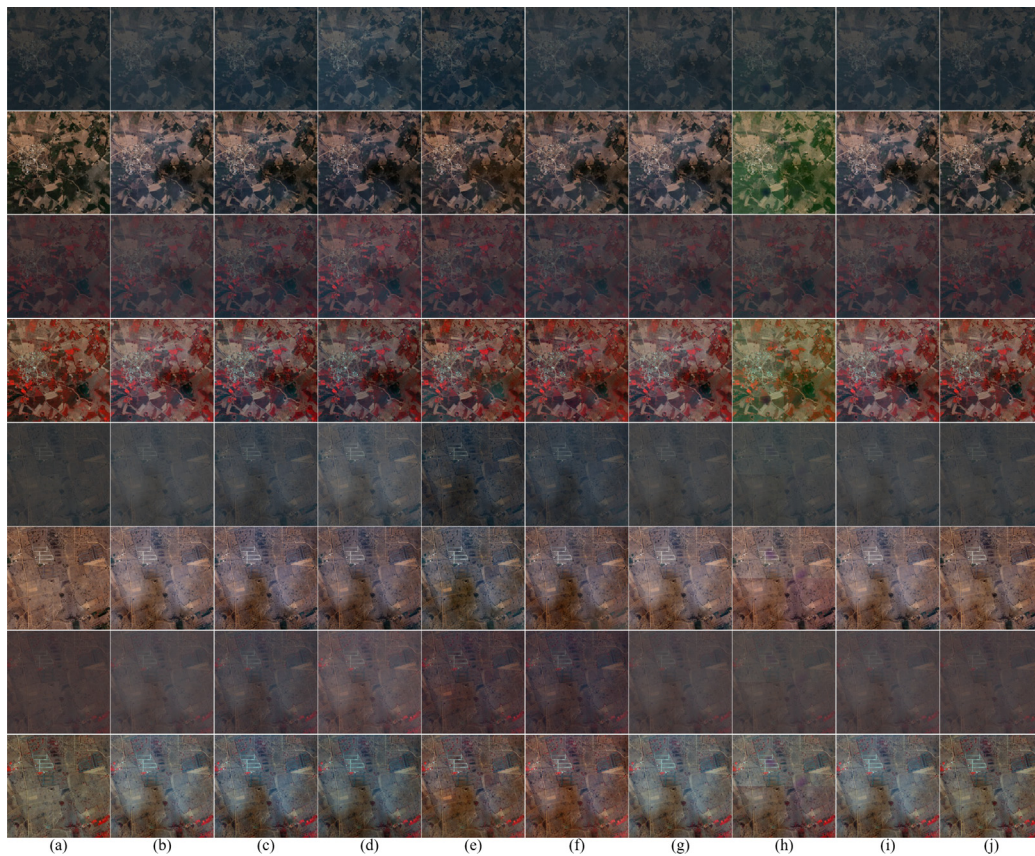
Fig. 10. Comparative results of the dehazing performance on RGB and NGB bands of synthetic multispectral hazy images: (a) haze-free images, (b) synthetic hazy images, (c) He's method (He et al., 2011), (d) Tarel's method (Tarel and Hautière, 2009), (e) Pan's method (Pan et al., 2015), (f) Zhu's method (Zhu et al., 2015), (g) AODNet (Li et al., 2017), (h) GCANet (Chen et al., 2019), (i) FFANet (Qin et al., 2019), and (j) the proposed method ($\gamma = 0.1$, $t_{min} = 0.1$, $t_{max} = 0.9$).
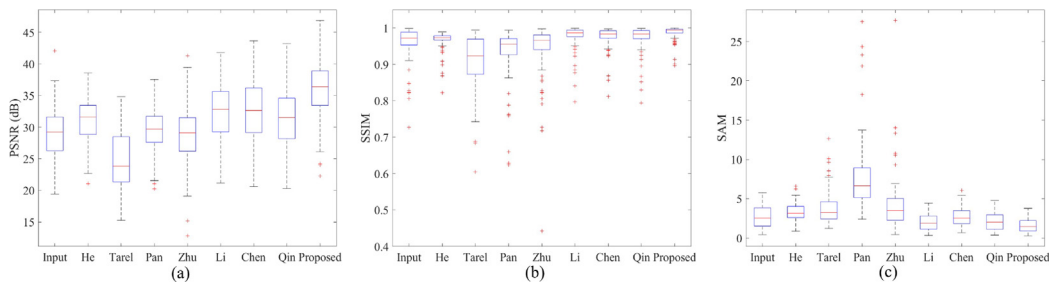

Fig. 11. Boxplots of PSNR, SSIM, and SAM values of the proposed and existing dehazing methods: (a) PSNR, (b) SSIM, and (c) SAM.

Table 5
Comparison of p-values for the differences in PSNR, SSIM, and SAM values using a paired t-test for a pair of the proposed and existing dehazing methods at a significance level of 5%.

	Input	He's method	Tarel's method	Pan's method	Zhu's method	Li's method	Chen's method	Qin's method
PSNR (dB)	2.55×10^{-36}	3.16×10^{-28}	9.72×10^{-39}	1.79×10^{-24}	1.31×10^{-25}	7.81×10^{-20}	1.60×10^{-13}	8.37×10^{-24}
SSIM	3.54×10^{-15}	2.39×10^{-31}	2.48×10^{-20}	1.95×10^{-15}	1.18×10^{-8}	8.98×10^{-9}	9.60×10^{-13}	1.65×10^{-11}
SAM	2.42×10^{-17}	1.19×10^{-28}	4.84×10^{-24}	8.97×10^{-26}	2.16×10^{-10}	2.04×10^{-8}	2.17×10^{-27}	1.13×10^{-9}

processing. We divided the test image into a set of patches of size 1280×1280 with overlapping of size 64 to reduce the artifacts and discontinuities between the adjacent patches. Finally, the dehazed patches were merged into a full-sized image. Given the real test image of size 6000×6000, the proposed method took about 6.8975 s including the pre-processing step.

Figs. 12 and 13 show the resulting dehazed images of the proposed and existing dehazing methods. In Figs. 12 and 13, the odd rows show the input hazy and dehazed images in the raw file format. The even rows show the contrast-enhanced version using conventional histogram

stretching method where t_{min} and t_{max} were respectively set to 0.02 and 0.98.

As shown in Figs. 12(b) and 13(b), He's method shows the halo artifact near edges with weak dehazing performance due to inaccurately estimated transmission and atmospheric light value. Pan's method shows the over-dehazing in the resulting images since this method globally subtracts a constant offset value with a certain degree. For this reason, Pan's method cannot control the over-dehazing and spectral distortion as shown in Figs. 12(c) and 13(c). Zhu's method cannot avoid spectral distortion, either. AODNet shows a weak dehazing
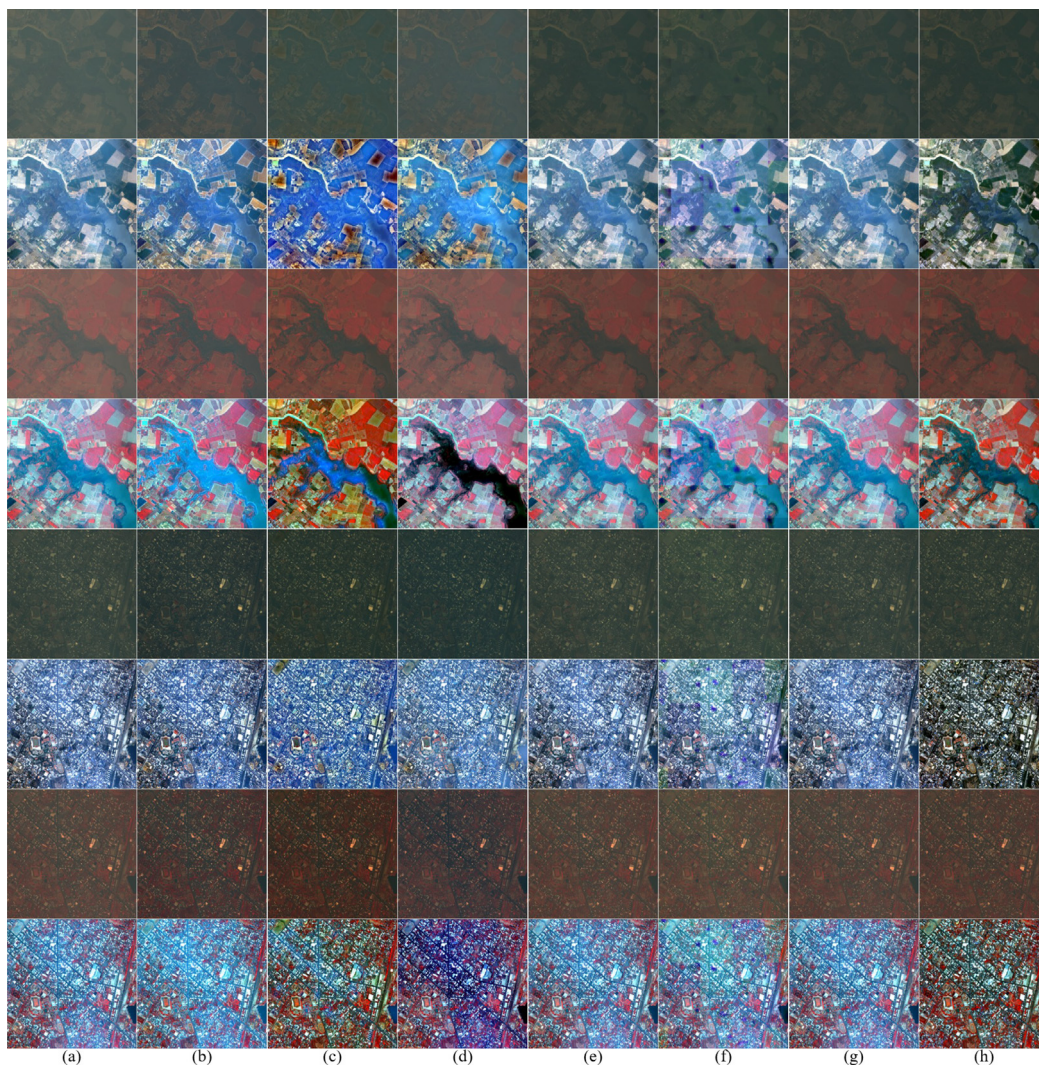
Fig. 12. Comparative results of the dehazing performance on RGB and NGB bands of real multispectral hazy images: (a) input hazy images, (b) He's method (He et al., 2011), (c) Pan's method (Pan et al., 2015), (d) Zhu's method (Zhu et al., 2015), (e) AODNet (Li et al., 2017), (f) GCANet (Chen et al., 2019), (g) FFANet (Qin et al., 2019), and (h) the proposed method ($\gamma = 0.3$, $t_{min} = 0.001$, $t_{max} = 0.999$).

performance for a moderate haze distribution of a real multispectral hazy image as shown in Figs. 12(e) and 13(e).

For the dehazing of very high-resolution multispectral imagery, the input hazy image is divided into blocks of a certain size because of the physical limitation of the GPU's memory. If the network consists of the normalization layer such as instance and batch normalization, each block is normalized using different local mean and standard deviation values.

As a result, the resulting dehazed images of GCANet show the discontinuity between adjacent blocks due to the brightness difference and spectral distortion since different normalization is performed using different local mean and standard deviation values in the instance normalization layer as shown in Figs. 12(f) and 13(f). FFANet provided better dehazing performance than other existing methods, but its performance is limited for a thick haze distribution compared with the proposed dehazing method as shown in Figs. 12(g) and 13(g).

Since the multispectral hazy image in the raw file format has a low dynamic range, we cannot accurately separate the hazy and haze-free regions as shown in Figs. 12(a) and 13(a). Consequently, it results in the over-dehazing in the haze-free region with spectral distortion as shown in the dehazed results of existing methods. The atmospheric scattering drastically changes the incoming light during haze weather, which affects the apparent colors of objects in the image. On the

other hand, since the HTMs estimated from an input hazy image and its contrast version give more attention to the haze distribution, the proposed dehazing method can achieve better dehazing performance than existing dehazing methods as shown in Figs. 12(h) and 13(h).

6. Discussion

6.1. Parameter setting

The proposed method estimates the wavelength-dependent, contrast-enhanced HTM by multiplying the scattering coefficient of (7) for each spectral band. Fig. 15 shows the comparative results of the dehazing performance of the proposed method using different γ with the same t_{min} and t_{max} fixed to 0.0001 and 0.9999, respectively.

As shown in Fig. 15(b), the lower γ provides a strong dehazing performance. On the other hand, the higher γ shows a weak dehazing performance. The scattering coefficient is inversely proportional to the γ-th power of the ratio of the longer wavelength to the shortest wavelength. Consequently, since the lower γ results in a large scattering coefficient, the network input using the contrast-enhanced HTM can give more attention to the heavy and moderate hazy regions.

We further evaluate the dehazing performance of the proposed method using the haze-free image including the snow region. The atmospheric scattering model-based methods remove the haze distribution
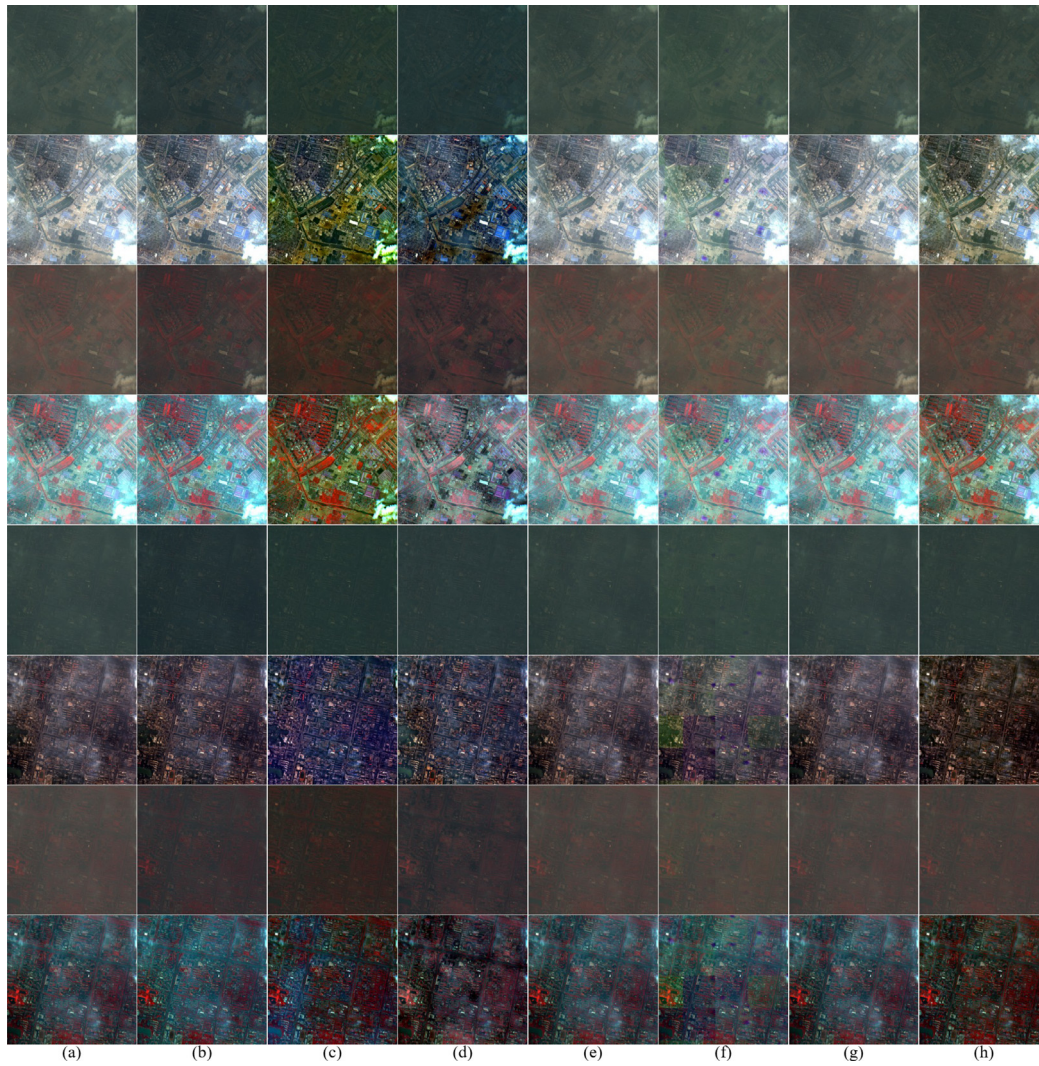
Fig. 13. Comparative results of the dehazing performance on RGB and NGB bands of real multispectral hazy images: (a) input hazy images, (b) He's method (He et al., 2011), (c) Pan's method (Pan et al., 2015), (d) Zhu's method (Zhu et al., 2015), (e) AODNet (Li et al., 2017), (f) GCANet (Chen et al., 2019), (g) FFANet (Qin et al., 2019), and (h) the proposed method ($\gamma = 0.3$, $t_{min} = 0.001$, $t_{max} = 0.999$).

using the transmission map estimated from the dark channel of an input hazy image. Since the dark channel is estimated using the minimum value among the red, green, and blue pixels, it has a higher value in the snow region, and it results in over-dehazing with spectral distortion as shown in Fig. 14.

On the other hand, the proposed method can prevent over-dehazing or spectral distortion under the haze-free condition and snow region. In Figs. 14 and 15, the odd rows show the input and resulting multispectral images in the raw file format and the even rows show the contrast-enhanced versions using conventional histogram stretching, where t_{min} and t_{max} were respectively set to 0.02 and 0.98.

6.2. Study of network input

This subsection analyzes the performance of the proposed dehazing method for different types of the network input. As mentioned in previous subsections, the characteristic of the low contrast of an observed multispectral hazy image makes the estimation of the haze distribution difficult. The proposed method solves this problem using the HTM estimated from the contrast-enhanced input hazy image.

Let g_i be an input multispectral hazy image, g_i^{HTM} be the HTM of g_i, and $g_{i,\text{CE}}^{\text{HTM}}$ be the HTM of the contrast-enhanced version of g_i. We set the four combinations of the network input to evaluate the effect of

the HTMs: (i) g_i, (ii) $[g_i, g_i^{\text{HTM}}]$, (iii) $[g_i, g_i^{\text{HTM}}, g_{i,\text{CE}}^{\text{HTM}}]$, and (iv) $[g_i, g_{i,\text{CE}}^{\text{HTM}}]$. Table 6 shows the quantitative comparison of different combinations. Case IV is a combination of the proposed dehazing network. As summarized in Table 6, the results show that the combinations using either an input hazy image or the concatenation to its HTM provided poor dehazing performance.

Fig. 17 shows the comparative results of the dehazing performance for different combinations of the network input. The odd rows show the multispectral images in the raw file format and the even rows show the contrast-enhanced versions using conventional histogram stretching, where t_{min} and t_{max} were respectively set to 0.02 and 0.98. These combinations cannot effectively remove the haze distribution as shown in Figs. 17(b) and 17(c). On the other hand, the HTM estimated from the contrast-enhanced version plays an important role in removing the haze distribution of the multispectral hazy image in the raw file format as shown in Fig. 17(e).

Fig. 18 shows the intermediate features of the proposed dehazing network for different combinations of the network input using the input hazy image of Fig. 17(a). From top to bottom, each row shows the intermediate features for Cases I to IV, as summarized in Table 6. We reduced the dimensionality of the intermediate features by summing along the depth dimension and adopted the sigmoid function.
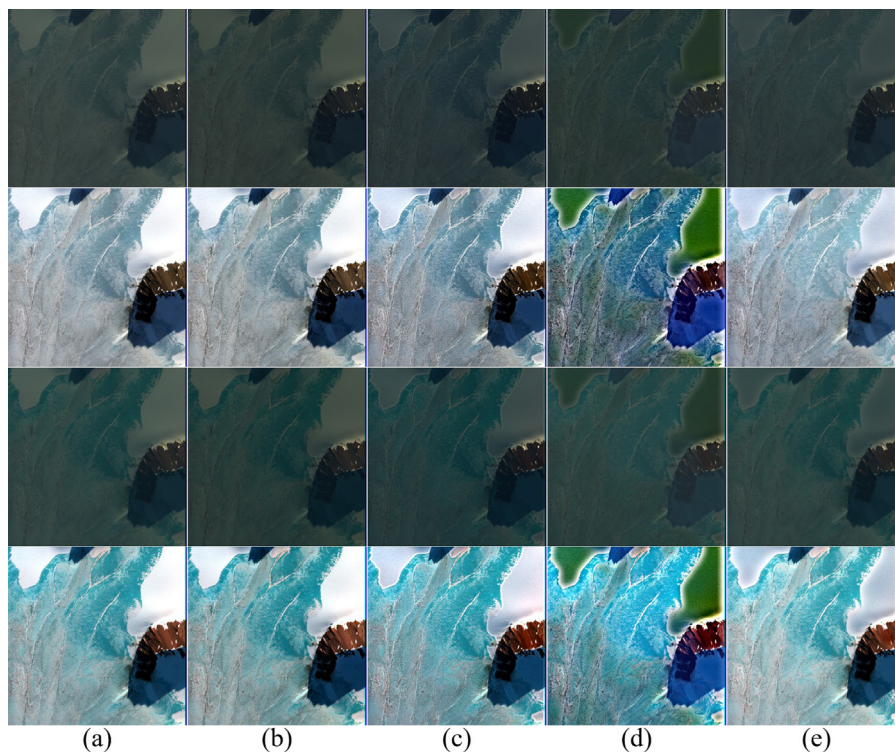
Fig. 14. Comparison of the dehazing performance for the haze-free image including the snow region: (a) input multispectral hazy image, (b) the proposed method ($\gamma = 0.1$), (c) He's method (He et al., 2011), (d) Pan's method (Pan et al., 2015), and (e) Zhu's method (Zhu et al., 2015).
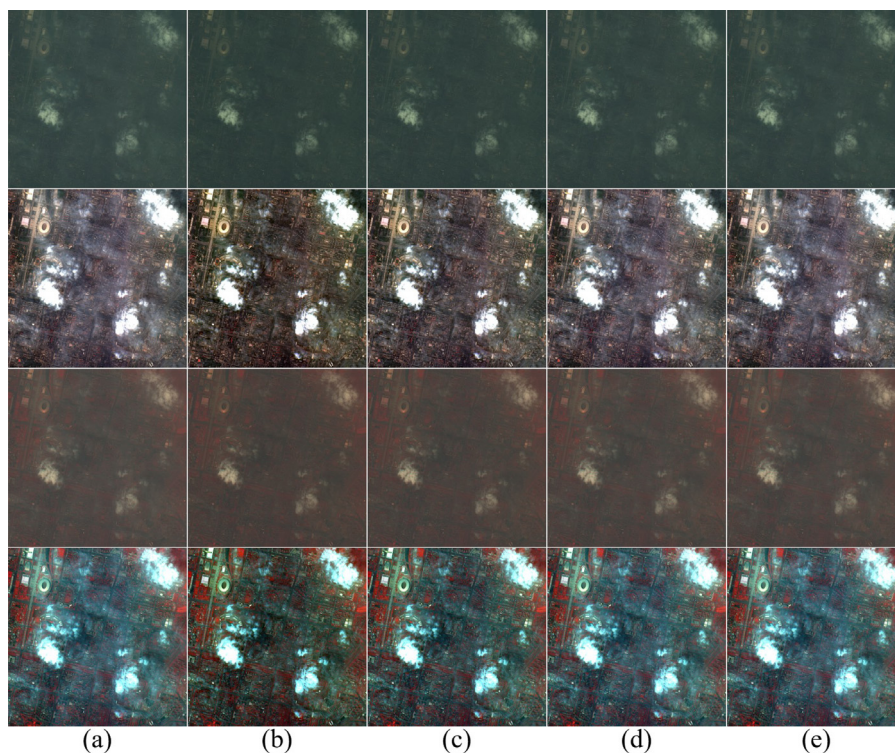

Fig. 15. Comparison of the dehazing performance of the proposed method with different values of γ: (a) the input multispectral hazy image, (b) $\gamma = 0.1$, (c) $\gamma = 0.7$, (d) $\gamma = 1.4$, and (e) $\gamma = 2$.

As shown in the first and second rows of Fig. 18, the network input using either an input hazy image or the concatenation to its HTM cannot extract the accurate haze distribution. Since the multispectral hazy image in the raw file format has a low dynamic range, the output features of the RN for Cases I and II show low activation for the region of haze distribution. For this reason, the output $\hat{H}$ of these combinations shows the uniform haze distribution. It means that the network input using only HTM of an input multispectral hazy image

Table 6

Comparison of the synthetic multispectral hazy image for different combinations of the network input.

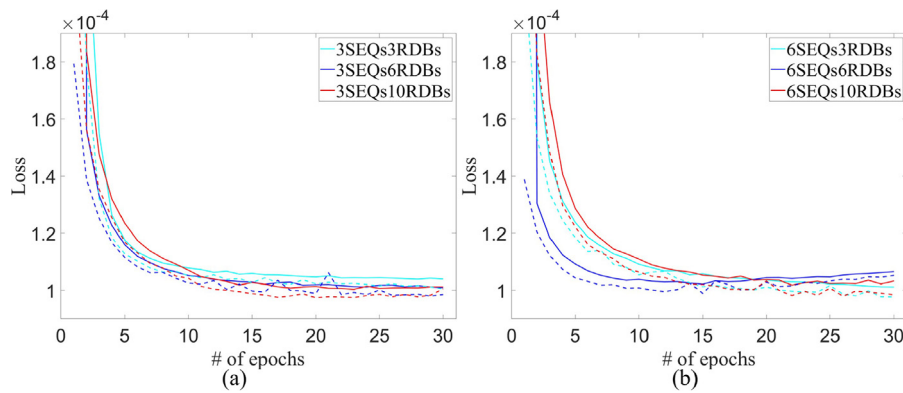|	Case I	Case II	Case II	Case IV
g_i	✓	✓	✓	✓
g_i^{HTM}	✗	✓	✗	✓
$g_{i,CE}^{HTM}$	✗	✗	✓	✓
PSNR (dB)	32.0369 ± 3.8667	32.1673 ± 4.5586	34.0352 ± 5.0125	**35.7425 ± 4.6089**
SSIM	0.9757 ± 0.0326	0.9762 ± 0.0323	0.9820 ± 0.0245	**0.9866 ± 0.0198**
SAM	2.1108 ± 1.1281	2.0794 ± 1.0866	1.9356 ± 1.0947	**1.6673 ± 0.9341**

Fig. 16. Comparison of the learning curves for different combinations of the proposed CRDBN: (a) three sequences of 1×1 and 3×3 convolutional layers with different numbers of RDBs and (b) six sequences of 1×1 and 3×3 convolutional layers with different numbers of RDBs. The dashed lines indicate the learning curves for validation.
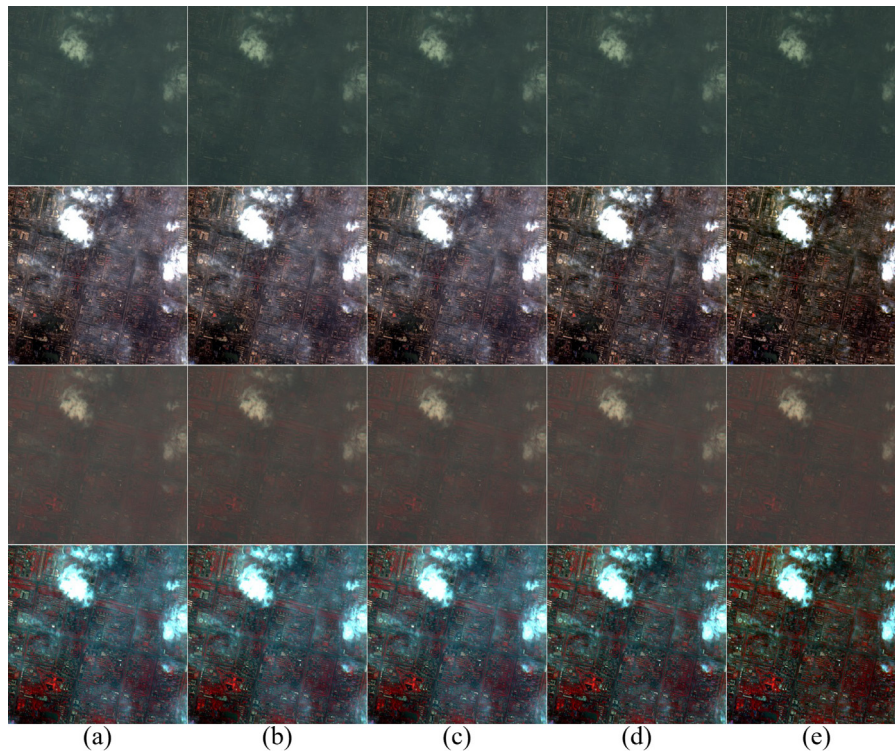

Fig. 17. Dehazing results using different combinations of the network input on RGB and NGB bands: (a) an input hazy image, (b) without HTMs, (c) with HTM estimated from (a), (d) with HTMs estimated from (a) and its contrast-enhanced version, and (d) with HTM estimated from the contrast-enhanced version of (a).

cannot give more attention to the haze distribution. As a result, the resulting images show poor dehazing performance for the moderate and heavy hazy regions as shown in Figs. 17(b) and 17(c).

The network input using the HTMs estimated from both the input multispectral hazy image and its contrast-enhanced version shows higher activation in the output features of the SFEN and RDBN. However, the HTM of an input hazy image cannot estimate the accurate activation for the haze distribution as shown in the second row of Fig. 18. For this reason, the MFEN and RN cannot provide higher activation for the haze distribution. Consequently, this combination cannot improve the dehazing performance since the output feature of the RN also shows the uniform distribution as shown in the third row of Fig. 18.
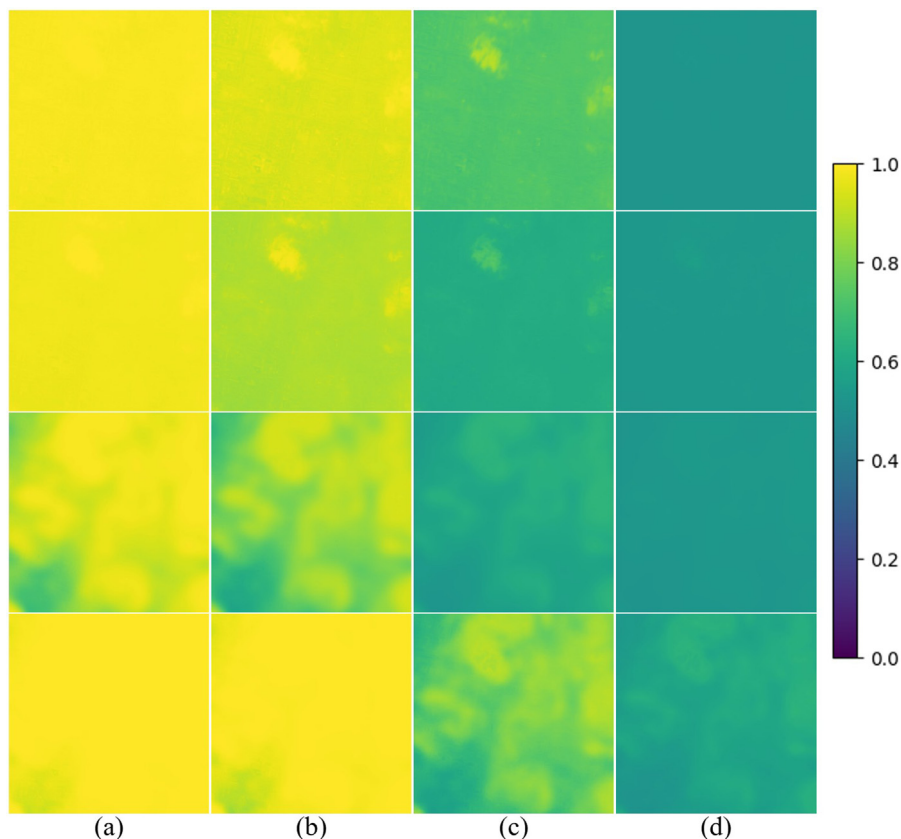
Fig. 18. Intermediate features of the proposed dehazing network for different combinations of the network input in Fig. 15: (a) the shallow feature F_{SFEN}, (b) the output feature of F_{RDBN}, (c) the output feature of F_{MFEN}, and (d) the output feature of F_{RN}. The intermediate features are estimated using the red, green, blue, and near infrared bands of the input hazy image of Fig. 17(a).

Table 7
Comparison of p-values for the differences in PSNR, SSIM, and SAM values using a paired t-test for a pair of Case IV and other combinations at a significance level of 5%.

	Case IV-Input	Case IV-Case I	Case IV-Case II	Case IV-Case III
PSNR (dB)	2.55×10^{-36}	4.02×10^{-21}	3.97×10^{-21}	2.11×10^{-10}
SSIM	3.54×10^{-15}	6.29×10^{-10}	1.55×10^{-9}	2.18×10^{-9}
SAM	2.42×10^{-17}	3.12×10^{-9}	1.81×10^{-8}	1.98×10^{-7}

On the other hand, the proposed combination of the input hazy image and the HTM of the contrast-enhanced input hazy image significantly improved dehazing performance as shown in Fig. 17(e) and Table 6. The output features of the MFEN and RN show higher activation than that of Case III as shown in the fourth row of Fig. 18. It means that the contrast-enhanced HTM can significantly improve the dehazing performance for the spatially varying haze distribution compared to other network input combinations as shown in Fig. 17(e).

Table 7 shows the p-values of paired t-tests for a pair of different network input combinations as previously given in Section 6.2 (McDonald, 2014). For a paired t-test, the null hypothesis assumed that PSNR, SSIM, and SAM values have no significant difference between the proposed and other network input combinations. On the other hand, the alternative hypothesis assumed that PSNR, SSIM, and SAM values are significantly different in the sense of statistical significance. The null hypothesis is rejected for $p < 0.05$.

As summarized in Table 7, the null hypothesis can be rejected at a significance level of 5% over Cases I, II, and III. It implies that the combination using the HTM of the contrast-enhanced input hazy image provided better dehazing performance with higher PSNR, SSIM, and lower SAM values than other network input combinations.

6.3. Study of network architecture

We analyze the dehazing performance for different configurations of the proposed dehazing network. The proposed CRDBN consists of the sequences of RDBs, and each RDB has the sequences of 1×1 and 3×3 convolutional layers as shown in Figs. 6 and 7. The 1×1 convolutional layer is used to reduce the depth dimension of the concatenated features of all preceding convolutional layers.

In this experiment, we used a different number of RDBs and 1×1 and 3×3 convolutional layers. The training configuration is set to the same as the implementation described in Section 5.3. Fig. 16 shows the training loss for different configurations of the CRDBN. As shown in this figure, the ten RDBs with three sequences of 1×1 and 3×3 convolutional layers provided better convergence than other combinations of the RDBN. On the other hand, the training loss of the CRDBN having six RDBs and six sequences of 1×1 and 3×3 convolutional layers increased after 16 epochs.

6.4. Convergence analysis

We compare the convergence of the proposed dehazing network with and without the residual learning strategy for different network input combinations. The combination using only one input hazy image
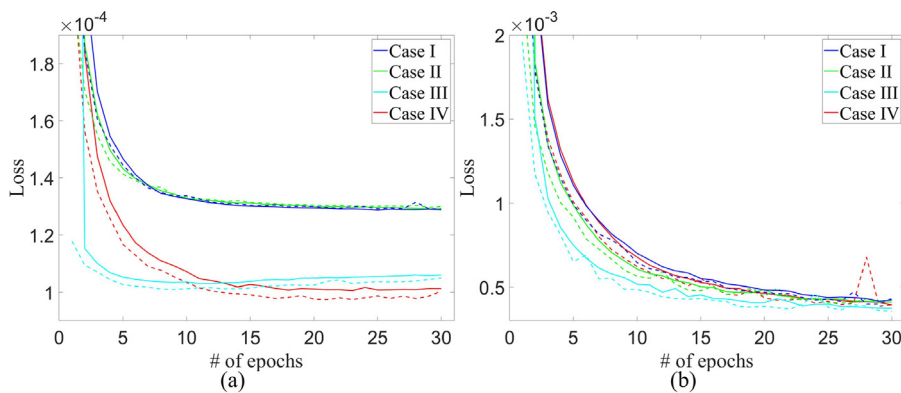
Fig. 19. Comparison of the learning curves for different combinations of the network input in Table 6: (a) with residual learning and (b) without residual learning. The dashed lines indicate the learning curves for validation.

is used as the baseline. The training configuration is set to the same as the implementation details described in Section 5.3. Fig. 19 shows the comparison of the learning curves of training and validation for four combinations with and without residual learning during the training process.

As shown in Fig. 19, the residual mapping estimating the haze distribution from an input hazy image can achieve more stable convergence than the original mapping, which directly estimates the dehazing result. The multispectral hazy image in the raw file format has a low dynamic range, and its latent version also shows a narrow histogram distribution. Learning the mapping from the input hazy image to its latent version is closer to identity mapping. On the other hand, the hidden layers in the residual mapping play an important role in separating the haze distribution from the input hazy image (Zhang et al., 2017). For this reason, residual learning can achieve better dehazing performance compared to the learning between the input hazy and its latent images.

In terms of the network input combinations, an input hazy image or the concatenation to its HTM cannot compete with the combination using the HTM of the contrast-enhanced input hazy image. The combination of the HTMs estimated from both the input and its contrast-enhanced version shows faster convergence than other combinations. However, both training and validation losses increased after 9 epochs, leading to poor dehazing performance. On the other hand, the proposed combination using the HTM of the contrast-enhanced version shows that the training and validation curves converged after 15 epochs. In other words, the HTM of the contrast-enhanced hazy image leads to the best dehazing performance in the sense of objective assessments as summarized in Table 6.

6.5. Ablation study

We further demonstrate the dehazing performance with a different combination of modules of the proposed dehazing method. The ablation investigation of the proposed dehazing network was performed using the same synthetic multispectral hazy images used in Section 5.3.

In Table 8, Case IV is the proposed dehazing network. As summarized in Table 8, the combination of the MFEN and RN provided poor dehazing performance with the lowest PSNR, SSIM, and SAM values than other combinations. The combinations without either the MFEN or RN are inferior to the proposed dehazing network. With three modules, the proposed dehazing network provided the best dehazing performance. We can notice that the CRDBN and MFEN play an essential role in improving the flow of the information and gradient of the proposed dehazing network. The RN also contributes to the accurate estimation of the haze distribution. Fig. 20 shows the learning curves for different combinations of the network module.

To verify the difference in PSNR, SSIM, and SAM values in the statistical manner, we performed a paired t-test for a pair of different
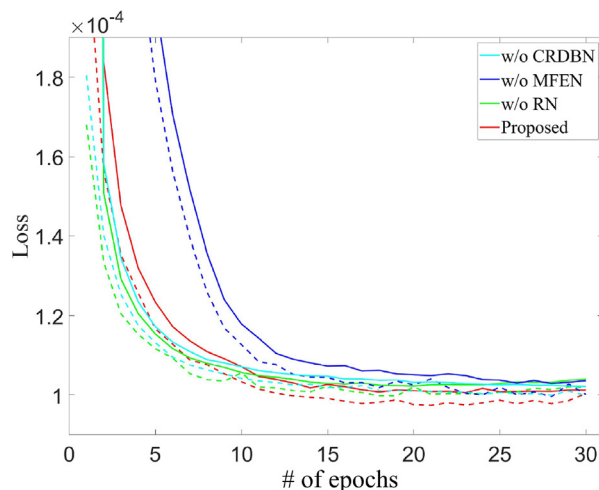

Fig. 20. Learning curves for different combinations of the network module in Table 8. The dashed lines indicate the learning curves for validation.

combinations of the network module of Table 8 (McDonald, 2014). The null hypothesis assumed that PSNR, SSIM, and SAM values have no significant difference between different combinations of the network module. The null hypothesis was rejected for $p < 0.05$. On the other hand, the alternative hypothesis assumed that PSNR, SSIM, and SAM values are significantly different. As summarized in Table 9, the null hypothesis can be rejected at a significance level of 5%. It implies that the difference in PSNR, SSIM, and lower SAM values between the proposed and other network combinations are statistically significant with improved dehazing performance.

6.6. Comparison to open-source image editor

The primary contribution of this work is to remove the spatially variant and semi-transparent haze distribution from a multispectral hazy image having a low dynamic range in the raw file format. However, since the haze removal using open-source image editor (GIMP) is usually based on the local and global contrast enhancement adjusting the brightness level, auto white balance, and unsharp masking methods, the dehazing performance is limited on a multispectral hazy image in the raw file format compared to the proposed method (GIMP, 2022).

The main limitation is that GIMP cannot distinguish the hazy and haze-free regions in the multispectral hazy image having a low dynamic range. As a result, GIMP cannot preserve the original information of an input multispectral hazy image in the raw file format since the local and global contrast enhancement changes the overall contrast and
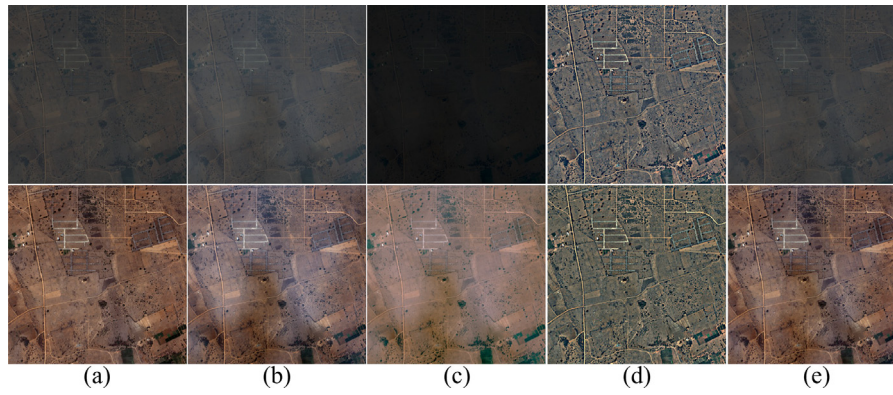
(a)   (b)   (c)   (d)   (e)

Fig. 21. Comparative results of the dehazing performance using GIMP and the proposed method: (a) haze-free images, (b) synthetic hazy images, (c) brightness level and curve adjustment, (d) GIMP multiscale retinex and brightness level adjustment, and (e) the proposed method.

Table 8
Objective assessments of the dehazing performance for different combinations of the network module.

	Case I	Case II	Case II	Case IV
RDBN	×	✓	✓	✓
MFEN	✓	×	✓	✓
RN	×	✓	×	✓
PSNR (dB)	33.4405 ± 5.0815	34.3229 ± 5.1264	34.2228 ± 5.1308	$\mathbf{35.7425 \pm 4.6089}$
SSIM	0.9804 ± 0.0269	0.9823 ± 0.0256	0.9824 ± 0.0248	$\mathbf{0.9866 \pm 0.0198}$
SAM	1.9385 ± 1.0900	1.8804 ± 1.0506	1.8543 ± 1.0635	$\mathbf{1.6673 \pm 0.9341}$

Table 9
Comparison of p-values for the differences in PSNR, SSIM, and SAM values using a paired t-test for a pair of Case IV and other combinations at a significance level of 5%.

	Case IV-Case I	Case IV-Case II	Case IV-Case III
PSNR (dB)	6.44×10^{-12}	8.39×10^{-8}	2.04×10^{-8}
SSIM	6.08×10^{-8}	3.57×10^{-6}	1.26×10^{-6}
SAM	1.63×10^{-5}	3.33×10^{-6}	0.28×10^{-3}

brightness. In the same manner, it generates undesired artifacts in post-processing step and results in inaccurate terrain and object analysis. For this reason, the proposed method uses the contrast-enhanced version of input multispectral hazy image to give different attention on the hazy and haze-free regions.

In addition, the multispectral hazy images are acquired with different amount of haze distribution at different location in the real-world case. It implies that we should apply different enhancement methods in GIMP with different parameters for various hazy scenes. It is critical factor making the haze removal using open-source image editor highly heuristic manner.

Fig. 21 shows the comparative haze removal results using GIMP and the proposed method. The first row shows the input hazy and dehazed images in the raw file format. The second row shows the contrast-enhanced version of the images in the first row using the conventional histogram stretching method where t_{min} and t_{max} were respectively set to 0.0001 and 0.9999. Figs. 21(a) and 21(b) are the multispectral haze-free and hazy images, which are the same as Figs. 10(a) and 10(b). In Fig. 21(c), we adopted the brightness level and curve adjustment to decrease the intensity values of haze distribution, but the resulting image cannot preserve the original information of raw data of Fig. 21(a).

GIMP built-in multiscale retinex method improved the visibility compared to Fig. 21(c) as shown in Fig. 21(d), but the resulting image shows a severe spectral and spatial distortion (Rahman et al., 1996). Fig. 21(e) is the haze removal result of the proposed method, which is the same as Fig. 10(j). As mentioned above, GIMP cannot preserve the raw information of an input multispectral hazy image since the brightness level and curve adjustment change the mean brightness

and overall contrast as shown in Figs. 21(c) and 21(d). It means that GIMP cannot perform the context-aware haze removal compared to the proposed method.

7. Conclusion

In the real-world condition, an observed multispectral hazy image shows the complex haze distribution including the haze-free region with different scattering properties for each spectral band. We synthesized the multispectral hazy image using the extracted haze distribution from a real hazy patch and imposed different atmospheric scattering properties using the relative scattering model. In addition, we used the HTM of each spectral band estimated from the contrast-enhanced input hazy image as the input of the proposed dehazing network to accurately estimate the haze distribution from an input multispectral hazy image having a low dynamic range. In comprehensive experiments, we analyzed that the contrast-enhanced HTM and the residual mapping estimating the haze distribution from an input multispectral hazy image play an important role in improving the dehazing performance in the sense of quantitative and qualitative assessments.

Existing object detection methods have been proposed to detect a certain object in high-quality remote sensing images. In future work, we will extend our research to simultaneously perform the haze removal and object detection from various hazy images having low dynamic range in the raw file format.

CRediT authorship contribution statement

Soohwan Yu: Methodology, Software, Experiment. **Doochun Seo:** Conceptualization, Resources, Data. **Joonki Paik:** Supervision, Writing – original draft, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

References

Ahn, N., Kang, B., Sohn, K., 2018. Fast, accurate, and lightweight super-resolution with cascading residual network. In: European Conference on Computer Vision. ECCV, pp. 256–272.

Chavez, P.S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sens. Environ. 24 (3), 459–479. http://dx.doi.org/10.1016/0034-4257(88)90019-3.

Chen, D., He, M., Fan, Q., Liao, J., Zhang, L., Hou, D., Yuan, L., Hua, G., 2019. Gated context aggregation network for image dehazing and deraining. In: 2019 IEEE Winter Conference on Applications of Computer Vision. WACV, pp. 1375–1383. http://dx.doi.org/10.1109/WACV.2019.00151.

Du, Y., Guindon, B., Cihlar, J., 2002. Haze detection and removal in high resolution satellite image with wavelet analysis. IEEE Trans. Geosci. Remote Sens. 40 (1), 210–217. http://dx.doi.org/10.1109/36.981363.

GIMP, 2022. GNU image manipulation program (GIMP). URL https://www.gimp.org/, (accessed 28 November 2022).

Gu, Z., Zhan, Z., Yuan, Q., Yan, L., 2019. Single remote sensing image dehazing using a prior-based dense attentive network. Remote Sens. 11 (24), http://dx.doi.org/10.3390/rs11243008.

Guzinski, R., Nieto, H., 2019. Evaluating the feasibility of using sentinel-2 and sentinel-3 satellites for high-resolution evapotranspiration estimations. Remote Sens. Environ. 221, 157–172.

He, K., Sun, J., Tang, X., 2011. Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. 33 (12), 2341–2353. http://dx.doi.org/10.1109/TPAMI.2010.168.

He, K., Sun, J., Tang, X., 2013. Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. 35 (6), 1397–1409. http://dx.doi.org/10.1109/TPAMI.2012.213.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In: 2015 IEEE International Conference on Computer Vision. ICCV, pp. 1026–1034. http://dx.doi.org/10.1109/ICCV.2015.123.

Hebel, M., Arens, M., Stilla, U., 2013. Change detection in urban areas by object-based analysis and on-the-fly comparison of multi-view ALS data. ISPRS J. Photogramm. Remote Sens. 86, 52–64.

Hore, A., Ziou, D., 2010. Image quality metrics: PSNR vs. SSIM. In: 2010 20th International Conference on Pattern Recognition. IEEE, pp. 2366–2369.

Huang, S., Li, D., Zhao, W., Liu, Y., 2019. Haze removal algorithm for optical remote sensing image based on multi-scale model and histogram characteristic. IEEE Access 7, 104179–104196. http://dx.doi.org/10.1109/ACCESS.2019.2929591.

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 2261–2269. http://dx.doi.org/10.1109/CVPR.2017.243.

Jeong, J., Kim, J., Kim, T., Rhee, S., 2016. Evaluation of the performance of KOMPSAT-3 stereo images in terms of positioning and the generation of digital surface models. Remote Sens. Lett. 7 (10), 955–964.

Jiang, H., Lu, N., Yao, L., 2016. A high-fidelity haze removal method based on HOT for visible remote sensing images. Remote Sens. 8 (10), 1–18. http://dx.doi.org/10.3390/rs8100844.

Ju, M., Ding, C., Guo, Y.J., Zhang, D., 2019. Remote sensing image haze removal using Gamma-correction-based dehazing model. IEEE Access 7, 5250–5261. http://dx.doi.org/10.1109/ACCESS.2018.2889766.

Kaufman, Y.J., Sendra, C., 1988. Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. Int. J. Remote Sens. 9 (8), 1357–1381. http://dx.doi.org/10.1080/01431168808954942.

Kaufman, Y.J., Wald, A.E., Remer, L.A., Bo-Cai Gao, Rong-Rong Li, Flynn, L., 1997. The MODIS 2.1-/spl mu/m channel-correlation with visible reflectance for use in remote sensing of aerosol. IEEE Trans. Geosci. Remote Sens. 35 (5), 1286–1298. http://dx.doi.org/10.1109/36.628795.

Kingma, D., Ba, J., 2015. Adam: A method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations. ICLR, pp. 1–15.

Li, J., Hu, Q., Ai, M., 2019. Haze and thin cloud removal via sphere model improved dark channel prior. IEEE Geosci. Remote Sens. Lett. 16 (3), 472–476. http://dx.doi.org/10.1109/LGRS.2018.2874084.

Li, W., Li, Y., Chen, D., Chan, J.C.W., 2019. Thin cloud removal with residual symmetrical concatenation network. ISPRS J. Photogramm. Remote Sens. 153, 137–150. http://dx.doi.org/10.1016/j.isprsjprs.2019.05.003, URL https://www.sciencedirect.com/science/article/pii/S092427161930125X.

Li, B., Peng, X., Wang, Z., Xu, J., Feng, D., 2017. AOD-net: All-in-one dehazing network. In: 2017 IEEE International Conference on Computer Vision. ICCV, pp. 4780–4788. http://dx.doi.org/10.1109/ICCV.2017.511.

Li, J., Wu, Z., Hu, Z., Zhang, J., Li, M., Mo, L., Molinier, M., 2020. Thin cloud removal in optical remote sensing images based on generative adversarial networks and physical model of cloud distortion. ISPRS J. Photogramm. Remote Sens. 166, 373–389. http://dx.doi.org/10.1016/j.isprsjprs.2020.06.021, URL https://www.sciencedirect.com/science/article/pii/S0924271620301787.

Lin, T., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017. Feature pyramid networks for object detection. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 936–944. http://dx.doi.org/10.1109/CVPR.2017.106.

Liu, C., Cheng, I., Zhang, Y., Basu, A., 2017. Enhancement of low visibility aerial images using histogram truncation and an explicit retinex representation for balancing contrast and color consistency. ISPRS J. Photogramm. Remote Sens. 128, 16–26.

Makarau, A., Richter, R., Müller, R., Reinartz, P., 2014. Haze detection and removal in remotely sensed multispectral imagery. IEEE Trans. Geosci. Remote Sens. 52 (9), 5895–5905. http://dx.doi.org/10.1109/TGRS.2013.2293662.

McDonald, J.H., 2014. Handbook of Biological Statistics (3rd Ed.). Sparky House Publishing, Baltimore, Maryland.

Moro, G.D., Halounova, L., 2007. Haze removal for high-resolution satellite data: A case study. Int. J. Remote Sens. 28 (10), 2187–2205. http://dx.doi.org/10.1080/01431160600928559.

Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted Boltzmann machines. In: Proceedings of the 27th International Conference on International Conference on Machine Learning. ICML '10, Omni Press, Madison, WI, USA, pp. 807–814.

Narasimhan, S.G., Nayar, S.K., 2003. Contrast restoration of weather degraded images. IEEE Trans. Pattern Anal. Mach. Intell. 25 (6), 713–724. http://dx.doi.org/10.1109/TPAMI.2003.1201821.

Olsen, R.S., 2007. Remote Sensing from Air and Space. SPIE—The International Society for Optical Engineering, Bellingham, Washington 98227-0010 USA.

Pan, X., Xie, F., Jiang, Z., Shi, Z., Luo, X., 2016. No-reference assessment on haze for remote-sensing images. IEEE Geosci. Remote Sens. Lett. 13 (12), 1855–1859. http://dx.doi.org/10.1109/LGRS.2016.2614890.

Pan, X., Xie, F., Jiang, Z., Yin, J., 2015. Haze removal for a single remote sensing image based on deformed haze imaging model. IEEE Signal Process. Lett. 22 (10), 1806–1810. http://dx.doi.org/10.1109/LSP.2015.2432466.

Pingel, T.J., Clarke, K.C., McBride, W.A., 2013. An improved simple morphological filter for the terrain classification of airborne LIDAR data. ISPRS J. Photogramm. Remote Sens. 77, 21–30.

Qin, X., Wang, Z., Bai, Y., Xie, X., Jia, H., 2019. FFA-net: Feature fusion attention network for single image dehazing. arXiv:1911.07559,

Qin, M., Xie, F., Li, W., Shi, Z., Zhang, H., 2018. Dehazing for multispectral remote sensing images based on a convolutional neural network with the residual architecture. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (5), 1645–1655. http://dx.doi.org/10.1109/JSTARS.2018.2812726.

Rahman, Z., Jobson, D.J., Woodell, G.A., 1996. Multi-scale retinex for color image enhancement. In: Proceedings of 3rd IEEE International Conference on Image Processing, Vol. 3. pp. 1003–1006. http://dx.doi.org/10.1109/ICIP.1996.560995.

Richter, R., Schläpfer, D., Müller, A., 2006. An automatic atmospheric correction algorithm for visible/NIR imagery. Int. J. Remote Sens. 27 (10), 2077–2085. http://dx.doi.org/10.1080/01431160500486690.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention. MICCAI, In: LNCS, vol. 9351, Springer, pp. 234–241.

Rottensteiner, F., Sohn, G., Gerke, M., Wegner, J.D., Breitkopf, U., Jung, J., 2014. Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. ISPRS J. Photogramm. Remote Sens. 93, 256–271.

Shao, S., Guo, Y., Zhang, Z., Yuan, H., 2019. Single remote sensing multispectral image dehazing based on a learning framework. Math. Probl. Eng. 2019, 1–8. http://dx.doi.org/10.1155/2019/4131378.

Sun, L., Latifovic, R., Pouliot, D., 2017. Haze removal based on a fully automated and improved haze optimized transformation for landsat imagery over land. Remote Sens. 9 (10), 1–22. http://dx.doi.org/10.3390/rs9100972.

Tarel, J., Hautière, N., 2009. Fast visibility restoration from a single color or gray level image. In: Proceedings of 2009 IEEE 12th International Conference on Computer Vision. pp. 2201–2208. http://dx.doi.org/10.1109/ICCV.2009.5459251.

Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K., Huete, A., 2018. A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on google earth engine cloud computing platform. ISPRS J. Photogramm. Remote Sens. 144, 325–340.

Ulyanov, D., Vedaldi, A., Lempitsky, V., 2017. Instance normalization: The missing ingredient for fast stylization. arXiv:1607.08022.

USGS, 2022. United States geological survey (USGS) earth explorer. URL https://earthexplorer.usgs.gov/, (accessed 28 November 2022).

Vincent, R.K., 1972. An ERTS multispectral scanner experiment for mapping iron compounds. In: The Eighth International Symposium on Remote Sensing of Environment. pp. 1239–1247.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612. http://dx.doi.org/10.1109/TIP.2003.819861.

Wang, Z., Ji, S., 2018. Smoothed dilated convolutions for improved dense prediction. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, pp. 2486–2495.

Xie, J., Liu, R., Mo, F., Tang, H., Jiao, H., Mei, Y., Yang, C., 2020. Pointing bias calibration of gaofen-7 laser altimeter based on single laser footprint image. ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci. 5 (2).

Yuhas, R.H., Goetz, A., Boardman, J., 1992. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In: Proceedings of JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop. pp. 147–149.

Zhang, Y., Guindon, B., Cihlar, J., 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of landsat images. Remote Sens. Environ. 82 (2), 173–187. http://dx.doi.org/10.1016/S0034-4257(02)00034-2.

Zhang, H., Patel, V.M., 2018. Densely connected pyramid dehazing network. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3194–3203. http://dx.doi.org/10.1109/CVPR.2018.00337.

Zhang, H., Sindagi, V., Patel, V.M., 2018a. Multi-scale single image dehazing using perceptual pyramid deep network. In: Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. CVPRW, pp. 1015–101509. http://dx.doi.org/10.1109/CVPRW.2018.00135.

Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y., 2018b. Residual dense network for image super-resolution. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2472–2481. http://dx.doi.org/10.1109/CVPR.2018.00262.

Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., 2017. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. Image Process. 26 (7), 3142–3155. http://dx.doi.org/10.1109/TIP.2017.2662206.

Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017. Pyramid scene parsing network. In: Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 6230–6239. http://dx.doi.org/10.1109/CVPR.2017.660.

Zhu, Q., Mai, J., Shao, L., 2015. A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. 24 (11), 3522–3533. http://dx.doi.org/10.1109/TIP.2015.2446191.

Soohwan Yu

Doochun Seo

Joonki Paik