

Dehazing for Multispectral Remote Sensing Images Based on Convolutional Neural Network with Residual Architecture

Manjun Qin, Fengying Xie*, Wei Li, Zhenwei Shi, *Member, IEEE*, and Haopeng Zhang

Abstract—Multispectral remote sensing images are often contaminated by haze, which causes low image quality. In this paper, a novel dehazing method based on deep convolutional neural network (CNN) with residual structure is proposed for multispectral remote sensing images. Firstly, multiple CNN individuals with residual structure are connected in parallel and each individual is used to learn a regression from the hazy image to the clear image. Then, the outputs of CNN individuals are fused with weight maps to produce the final dehazing result. In the designed network, the CNN individuals, mining multiscale haze features through multiscale convolutions, are trained using different levels of haze samples to achieve different dehazing abilities. In addition, the weight maps change with the haze distribution, and the fusion of the CNN individuals is adaptive. The designed network is end-to-end, and putting a hazy image into it, the clear scene can be restored. To train the network, a wavelength-dependent haze simulation method is proposed to generate labeled data, which can synthesize hazy multispectral images highly close to real conditions. Experimental results show that the proposed method can accurately remove haze in each band of multispectral images under different scenes.

Index Terms—Haze removal, Convolutional Neural Network, haze simulation, multispectral remote sensing images.

I. INTRODUCTION

MULTISPECTRAL remote sensing imagery can provide abundant ground information, which becomes an essential way to explore earth resources and ecological environment. However, multispectral images are often affected by atmospheric conditions such as haze, fog, and cloud. These phenomena reduce the image visibility and lead to the loss of texture details, which bring obstacles for many applications such as land classification and target detection. Therefore, haze removal is necessary to improve the quality of remote sensing images.

Numbers of dehazing methods have been developed for multispectral remote sensing images. Earlier methods usually focused on dehazing for each spectral band of multispectral data. Richter [1] matched the histograms between haze areas and clear parts to correct haze in Landsat TM and SPOT HRV satellite imagery. Feng *et al.* [2] adopted homomorphic filter to remove thin cloud from ASTER data. Du *et al.* [3]

decomposed the image using wavelet transform to remove the spatially varying haze. Liu *et al.* [4] combined the virtual cloud point method with the background suppressed haze thickness index to dehaze for high spatial resolution satellite imagery. Chavez [5] took the band correlation property of atmospheric scattering into account and used dark-object subtraction method to correct uniform haze. Makarau *et al.* [6] further developed Chavez's work [5] through calculating a haze thickness map (HTM) to remove inhomogeneous haze.

Since visible bands are more easily contaminated by haze, some haze removal methods aiming at visible bands were developed. Zhang *et al.* [7] developed a haze optimized transformation method based on the spectral response of visible bands to dehaze. This method was further improved by Moro *et al.* [8] and He *et al.* [9] for more robust and stable performance. Shen *et al.* [10] adopted the homomorphic filter for thin cloud removal. Li *et al.* [11] and Galdran *et al.* [12] used the gradient-based fusion method to sharpen the visible imagery. Lan *et al.* [13] developed a three-stage algorithm for haze removal considering sensor blur and noise. Ni *et al.* [14] used linear intensity transformation and local property analysis to dehaze for images from Google Earth and NASA Earth. Long *et al.* [15] combined the dark channel prior [16] with Gaussian filter to restore the clear image. Pan *et al.* [17] deformed the haze imaging model and combined it with dark channel prior to improve the dehazing performance. Liu *et al.* [18] removed the haze component by calculating a ground radiance suppressed haze thickness map (GRS-HTM), which was more precise than HTM. These methods above did not consider non-visible bands even though they also will be contaminated by haze more or less.

Recently, supervised learning based methods have been developed to solve the dehazing problem for outdoor images. Tang *et al.* [19] extracted a set of haze-relevant features, with which a regression model based on Random Forest was learned to predict the medium transmission. Cai *et al.* [20] designed a Convolutional Neural Network (CNN) model to regress the transmission to dehaze. The real haze in multispectral remote sensing images is spatially varying and wavelength-dependent. However, [19] and [20] synthesized wavelength-independent haze samples to train their dehazing models, and regressed the transmissions under the assumption that the transmission is local constant. Therefore, they cannot effectively remove haze from multispectral images.

In this paper, a novel haze removal method based on CNN is proposed for multispectral remote sensing imagery. In the

*Fengying Xie is with the Image Processing Center, School of Astronautics, Beihang University, Beijing, China (e-mail: xfy_73@buaa.edu.cn).

Manjun Qin, Zhenwei Shi and Haopeng Zhang are also with the Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, China (e-mail: mnjune@163.com).

Wei Li is with Shanghai Institute of Satellite Engineering, Shanghai, People's Republic of China (e-mail: liwei20th@139.com).

A demo has been made available online: <http://xfy.buaa.edu.cn/>

designed network, multiple CNN individuals with residual structure are used to learn the mapping from the hazy image to the clear image for different levels of haze samples. Through the fusion unit, the outputs of CNN individuals are adaptively combined to yield the final restored image. In addition, considering the wavelength correlation property of atmospheric scattering, a wavelength-dependent haze synthesis method based on Rayleigh's law is proposed to generate the labeled data to train the network. The designed network is end-to-end, and the haze in the multispectral images can be effectively and adaptively removed.

In summary, the proposed method has the following three main contributions.

1) An end-to-end haze removal framework based on CNN is proposed for multispectral remote sensing images, in which multiple CNN individuals are connected in parallel to learn the mapping from the hazy image to the clear image, and a fusion unit is used to adaptively combine these individuals' outputs to generate the final restored image.

2) The designed CNN individual employs multiscale convolutions to mine the multiscale features of haze and adopts residual structure to reduce the learning difficulty. Better performance is obtained.

3) A new wavelength-dependent haze simulation method is proposed to generate hazy multispectral images close to real conditions, with which to train the designed network, more accurate dehazing results can be obtained.

The remainder of this paper is organized as follows. In Section II, the atmospheric scattering model is described, which is employed in this work. In Section III, we present the details of the proposed dehazing framework. Section IV introduces the proposed haze simulation method to generate training data. Experimental results and analysis are shown in Section V, and the conclusion is given in Section VI.

II. ATMOSPHERIC SCATTERING MODEL

The atmospheric scattering model was widely used in state-of-the-art haze removal algorithms. It is based on the two scattering phenomena of airlight and attenuation [21], and its definition is:

$$\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})t(\mathbf{x}) + \mathbf{A}(1 - t(\mathbf{x})) \quad (1)$$

where $\mathbf{x}$ denotes the position of the point in the scene, $\mathbf{I}(\mathbf{x})$ is the received hazy image by sensor, $\mathbf{J}(\mathbf{x})$ is the real scene radiance to be recovered, $\mathbf{A}$ is the global atmospheric light, and $t(\mathbf{x})$ is the medium transmission. More precisely, the medium transmission t is described as:

$$t(\mathbf{x}) = e^{-\beta(\mathbf{x}, \lambda)d(\mathbf{x})} \quad (2)$$

where $d(\mathbf{x})$ represents the light path from the scene point to the sensor, which can be viewed as a constant for remote sensing images, and $\beta(\mathbf{x}, \lambda)$ is termed as attenuation coefficient, which is dependent on the atmospheric particles and wavelength λ .

According to Rayleigh's law [22] [23], the scattering coefficient β is inversely proportional to the γ th power of wavelength λ , which is described as:

$$\beta = \frac{\text{constant}}{\lambda^\gamma} \quad (3)$$

where $\gamma \in [0, 4]$ that depends on the size of suspended particles in the atmosphere. For pure air, the size of constituent molecules ($10^{-4}\mu\text{m}$) is much smaller than the wavelength and $\gamma = 4$. For fog and clouds, the size of component water droplets is $1\mu\text{m} - 10\mu\text{m}$, which is larger than the wavelength and $\gamma \approx 0$. For a wide range of atmospheric conditions that are arose from aerosols, their particle sizes range between molecules and water droplets and $0 < \gamma < 4$. More specially, γ varies from 0.5 to 1 for haze conditions [5].

III. THE PROPOSED APPROACH

According to the haze imaging model, namely the atmospheric scattering model (1), once the atmospheric light $\mathbf{A}$ and medium transmission t are estimated, the clear image $\mathbf{J}$ can be easily obtained from:

$$\mathbf{J} = \frac{\mathbf{I} + \mathbf{A}(1 - t)}{t} \quad (4)$$

Since only $\mathbf{I}$ is known, dehazing is an ill-posed problem. Assuming transmission t is local constant, Tang [19] and Cai [20] utilized a supervised learning framework to regress t from the local image patch. Refined by the guided image filtering [24], the obtained transmission t is put into (4) to produce the dehazed image. In this paper, an end-to-end dehazing framework based on CNN with residual structure is designed to directly regress the clear image, which can achieve better dehazing performance.

A. Designed dehazing framework

The architecture of the designed dehazing framework is shown in Fig. 1, where n CNN individuals with residual structure are connected in parallel to learn the regression from the hazy image to the clear image. Training with different levels of haze samples, these CNNs have different dehazing abilities and can generate different outputs. These outputs are multiplied by their corresponding weight maps and fused through a convolutional layer to generate the final clear image.

1) *CNN individual*: In the designed model, n CNN individuals with the same structure are used to learn the mapping from the hazy image to the clear image. A residual network [25] can learn from reference, which is more easily converged than the non-residual network for the regression problem [26]. Since the dehazed image has similar texture and color with the corresponding hazy image, the latter can be viewed as an approximate version of the former. Using the CNN with residual structure to learn the dehazing mapping, the complex texture and color information in the dehazed image is directly provided by the input image, and only the different part (haze component) is learned and represented by convolutional layers. Thus, the learning difficulty is decreased.

The structure of the designed CNN individual is shown in Fig. 2. A 3×3 convolutional layer with 16 filters is first used to map the input image to high dimension. It is then followed by a residual structure with two multiscale convolutional layers, a feature fusion layer and an element-wise subtraction layer. Each multiscale convolutional layer consists of three paralleled convolutions with different kernel sizes $\{1 \times 1, 3 \times 3, 5 \times 5\}$, to

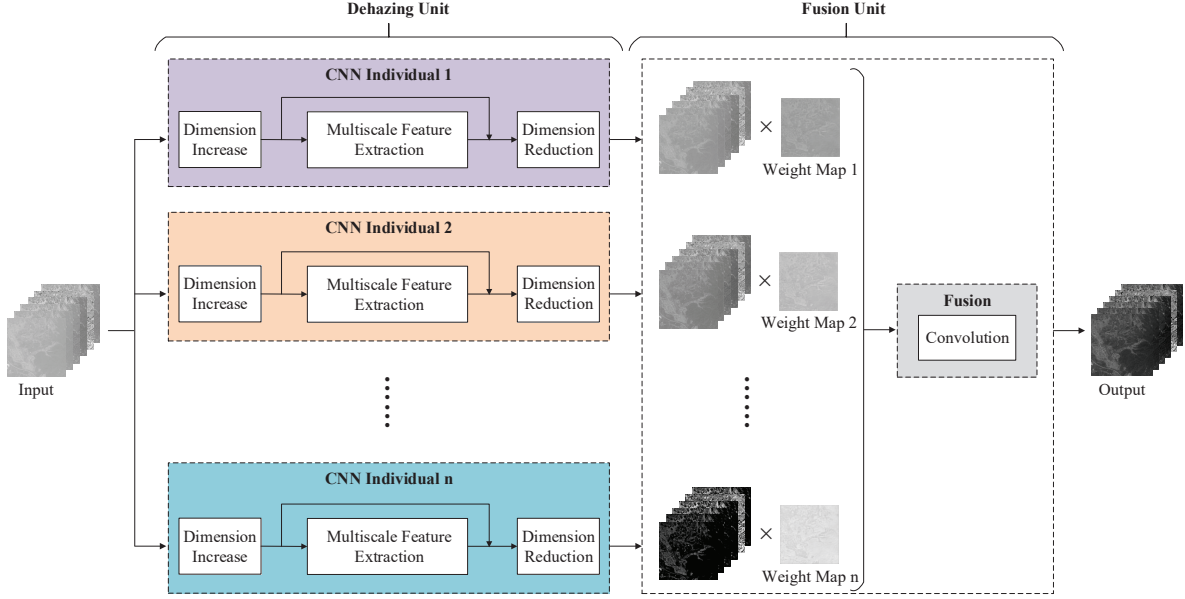


Fig. 1. Architecture of the designed dehazing network.

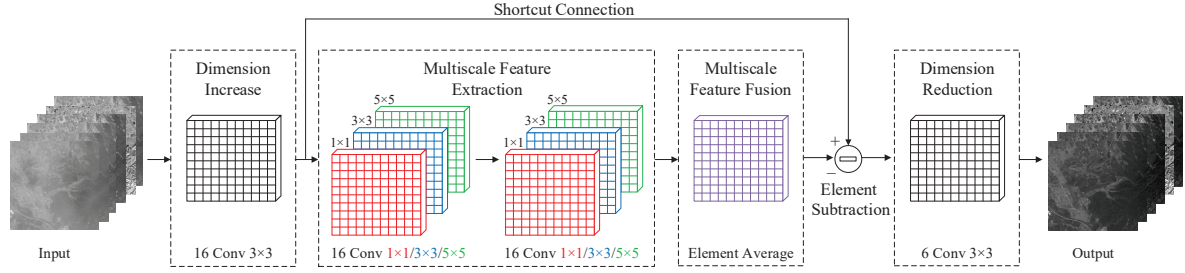


Fig. 2. Structure of the designed CNN individual.

extract multi-scale features of haze. To keep size invariant for the three scales, paddings are set to $\{0, 1, 2\}$ and strides are $\{1, 1, 1\}$. The feature fusion layer is used to fuse the multi-scale feature maps, which calculates the average values among the corresponding feature maps of the three scales, pixel by pixel, see Fig. 3. The fused feature maps represent the haze component. Then, the element-wise subtraction layer subtracts the haze component from the input image to perform haze removal in high dimension. The last 3×3 convolutional layer maps the output of the residual structure to low dimension to generate the final dehazing result.

2) *Adaptive fusion*: In the designed dehazing framework, the results of n CNN individuals are multiplied by their corresponding weight maps, and then fused through a convolutional layer to generate the final dehazing result. The weight map for each CNN individual is obtained through calculating the distance between the individual's inner average haze map and the input image's haze map, as shown in Fig. 4.

HTM method [6] can be used to extract the haze map from a band of the input multispectral remote sensing image, which

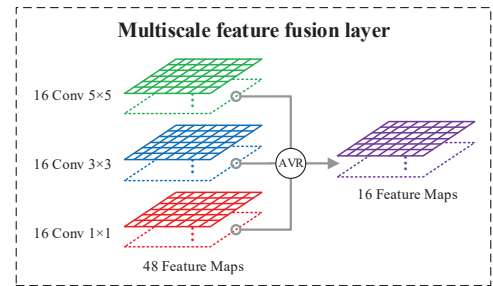


Fig. 3. Illustration of the multiscale feature fusion layer.

searches dark objects within a local window:

$$h(\mathbf{x}) = \min_{\mathbf{y} \in \Omega(\mathbf{x})} I(\mathbf{y}) \quad (5)$$

where $\Omega(\mathbf{x})$ is a window centered at $\mathbf{x}$, I is a band of the multispectral image. In this paper, we use an overlapping window with size 3×3 to search the minimum value in the first band, on which the haze effect is the most serious. Then,

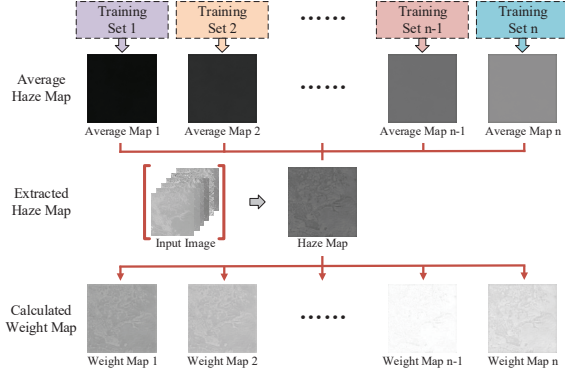


Fig. 4. Generation of the weight maps.

TABLE I
SPECTRAL BANDS OF LANDSAT 8 OLI

Spectral Band	Wavelength Range
Band 1 - Coastal/Aerosol	0.433 – 0.453 μm
Band 2 - Blue	0.450 – 0.515 μm
Band 3 - Green	0.525 – 0.600 μm
Band 4 - Red	0.630 – 0.680 μm
Band 5 - Near Infrared	0.845 – 0.885 μm
Band 6 - Short Wavelength Infrared	1.560 – 1.660 μm
Band 7 - Short Wavelength Infrared	2.100 – 2.300 μm
Band 8 - Panchromatic	0.500 – 0.680 μm
Band 9 - Cirrus	1.360 – 1.390 μm

the guided image filtering [24] is used to refine the haze map to remove the small blocking artifacts.

In the designed framework in Fig. 1, CNN individuals are trained with different levels of haze samples. For individual i , we calculate the average $h(\mathbf{x})$ using (5) on its training set as its inner haze map AM_i , which reflects the strong dehazing ability on this level of haze map. For an input hazy image, we extract its haze map HM using (5). Then, the weight map of individual i is defined as:

$$WM_i = 1 - |HM - AM_i|, \quad i = 1, \dots, n. \quad (6)$$

The weight map of a CNN individual reflects its dehazing performance for an input hazy image. The bigger the weight, the smaller the distance between the individual's inner haze map and the input image's haze map, and the stronger the dehazing ability of this individual on the input image.

Next, n outputs of the CNN individuals are element-wisely multiplied by the calculated weight maps, and then these multiplying results are put into a 1×1 convolutional layer to regress the final clear image, see the fusion unit of Fig. 1.

In our designed framework, when dehazing for an input image, these CNN individuals have different dehazing abilities, and the one with stronger dehazing ability is assigned bigger weight. The haze distribution in the image is random, and the weight maps change with the haze distribution. Therefore, under the guidance of weight maps, the fusion is adaptive.

B. Training and dehazing

Our experimental data is from Landsat 8 OLI (Operational Land Imager) imagery with 12-bit radiometric resolution. The

band information is listed in Table I. According to the atmospheric scattering theory, haze is wavelength-dependent, and the influence of haze decreases gradually with the wavelength increasing. Haze in Landsat 8 OLI images is obvious in the coastal, visible and panchromatic bands, and gets weak in the near-infrared band. For Band 6, Band 7 and Band 9, their wavelengths are more than $1\mu\text{m}$ which can penetrate haze particles ($size < 1\mu\text{m}$), and the influence of haze can be ignored. Therefore, in this paper, haze removal is performed in the 6 bands with wavelength less than $1\mu\text{m}$, including coastal, visible, panchromatic and near-infrared bands, and correspondingly, the input of the designed network has 6 channels. Since the panchromatic band has a different resolution from the other bands, it is resized to the same size with the others. Because deep learning methods usually adopt training images in the range of $[0, 1]$, in this paper, each band is divided by 2^{12} to achieve normalization.

Training a deep CNN requires a big labelled dataset including hazy images and their corresponding ground truth. The collection of them is difficult and time-consuming for multispectral remote sensing images, which need to be taken under the same scene in different time phases. A feasible way is to synthesize hazy images based on the physical haze imaging model [19] [20] [27]. In this paper, a haze synthesis method based on Rayleigh's law is proposed to generate training data, which will be introduced in the next section.

There are two training stages: one is the training of CNN individuals and the other is the training for fusion. In the first stage, we divide the whole training dataset into n groups according to the haze thickness (the training images are generated by simulation, and their haze thickness is known). Using the n groups of images to train the n CNN individuals in Fig. 1 respectively, n dehazing models are obtained. Because the network individuals are trained using different levels of haze, they are strong dehazing networks for the corresponding level of haze, and weak dehazing networks for other levels of haze. The second stage is to train the fusion unit using the whole training dataset. Each image in the whole dataset is put into the n CNN individuals to obtain n dehazing results. And at the same time, the weight maps are calculated. Finally, the n dehazing results are multiplied by their corresponding weight maps and put into the convolutional layer to train the fusion model.

Both training stages use the Euclidean Distance between the output image and the clear image as the loss function. Stochastic gradient descent (SGD) is used to search the optimal parameter. All the convolutional weights are initialized using a zero-mean Gaussian distribution with standard deviation 0.01, and the biases are initialized with constant 1. Training uses a batch size of 10. The weight decay and momentum are set to 0.0001 and 0.9, respectively. The learning rate is initially set to 10^{-7} and decreases by half every 50,000 iterations. The model is trained for 100,000 iterations.

In the test stage, the designed dehazing framework can achieve end-to-end dehazing. The n CNN individuals and the fusion unit are connected to constitute the whole dehazing network. Putting the hazy image into it, the restored image can be obtained at the output end.

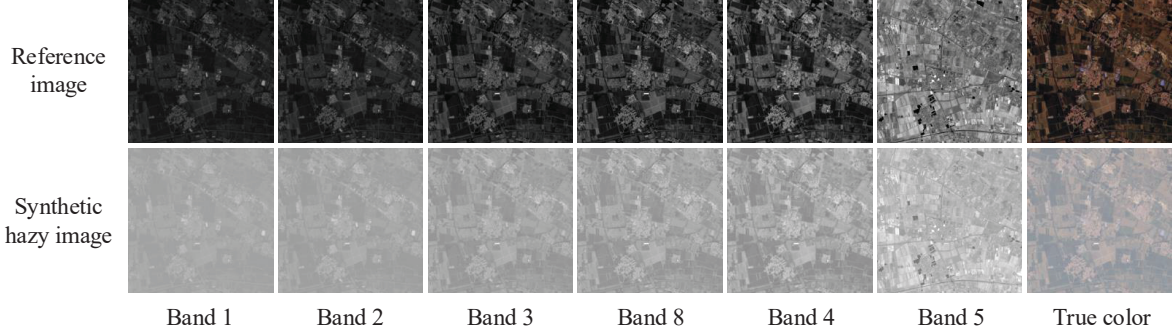


Fig. 5. Synthetic image with uniform haze.

IV. HAZE SIMULATION

Learning-based dehazing methods always adopt the haze synthesis manner to produce abundant labeled data for training. Synthetic hazy images can also be used for quantitative performance assessment of dehazing methods [19] [20]. Therefore, haze simulation plays an important role in dehazing problem.

Synthesizing hazy images is based on the atmospheric scattering model (1). Given a haze-free image $\mathbf{J}(\mathbf{x})$, the atmospheric light $\mathbf{A}$ and transmission $t \in (0, 1)$, the hazy image can be synthesized using (1). Usually, $\mathbf{A}$ is set to 1 for simplification [19] [20]. For the transmission t , Tang [19] and Cai [20] set it to be global-constant, which can generate the image with uniform haze. Pan [27] extracted the transmission mask from a real hazy image and added it to a clear image to synthesize the image with nonuniform haze. They assumed that haze between different bands is wavelength-independent, and generated the same transmission for all channels. However, according to the atmospheric scattering theory, haze is highly wavelength-dependent and decreases progressively with the wavelength increasing. In this paper, we propose a new haze synthesis method based on Rayleigh's law, which takes the wavelength correlation into account and can generate hazy images closer to real conditions.

A. Relation of transmissions between two bands

Taking band 1 as standard band, we derive the relational expression of transmissions between it and other bands. According to (2), the medium transmission is exponential function of the scattering coefficient β and the light path d . Taking a natural logarithm on both sides of (2), we can get:

$$\ln t = -d\beta(\lambda) \quad (7)$$

Then, the ratio of $\ln t$ between band 1 and band i can be written as:

$$\ln t_1 : \ln t_i = -d\beta_1 : -d\beta_i = \beta_1 : \beta_i, \quad i = 2, 3, 4, 5, 8 \quad (8)$$

where t_i and β_i are the transmission and scattering coefficient of the i th band. Thus, transmission t_i can be represented as:

$$t_i = e^{\frac{\beta_i}{\beta_1} \ln t_1} \quad (9)$$

According to Rayleigh's law (3), the scattering coefficient β is inversely proportional to the γ th power of the wavelength λ . Putting (3) into (9), we have:

$$t_i = e^{(\frac{\lambda_1}{\lambda_i})^\gamma \ln t_1} \quad (10)$$

where γ depends on the size of suspended particles in the atmosphere, and λ_i denotes the wavelength of the i th band. Usually, γ is suggested to 0.5, 0.7 and 1, corresponding to dense haze, haze and moderate haze respectively [5], and λ_i is the center wavelength of each band [5] [11], for example, the center wavelength of band 1 is $0.443 \mu\text{m}$ (see Table I).

Equation (10) is the final derived relational expression of transmissions between band 1 and band i . Through this function, given the transmission t_1 , t_i can be calculated directly.

B. Haze synthesis method

When synthesizing hazy images, the transmission of band 1 (standard band) can be set to a global-constant value or extracted from a real hazy image using Pan's method [27], and other bands' transmissions are calculated by taking the transmission of band 1 into (10). Adding these transmissions on the corresponding bands of a clear image through (1), the synthetic hazy image is obtained.

Setting the transmission of band 1 to constant 0.5, Fig. 5 shows a synthetic image with uniform haze, where the first six columns are band images (their center wavelengths are increasing from left to right), and the last column is the RGB composite images. As can be seen, synthetic haze gets weak gradually from left to right with the wavelength increasing, which is consistent with the atmospheric scattering theory.

Fig. 6 shows a synthetic image with nonuniform haze. The first row exhibits a real hazy image, which provides the transmission template for band 1. The second row shows the generated transmission masks, in which the transmission mask of band 1 is extracted from the real hazy image in the first row using Pan's method [27], and the remaining five transmission masks are calculated by taking the generated transmission mask of band 1 into (10). Adding these transmission masks to the clear image in the third row, the synthetic hazy image can be obtained and exhibited in the last row. It can be seen that, the distribution of synthetic haze is highly close to that

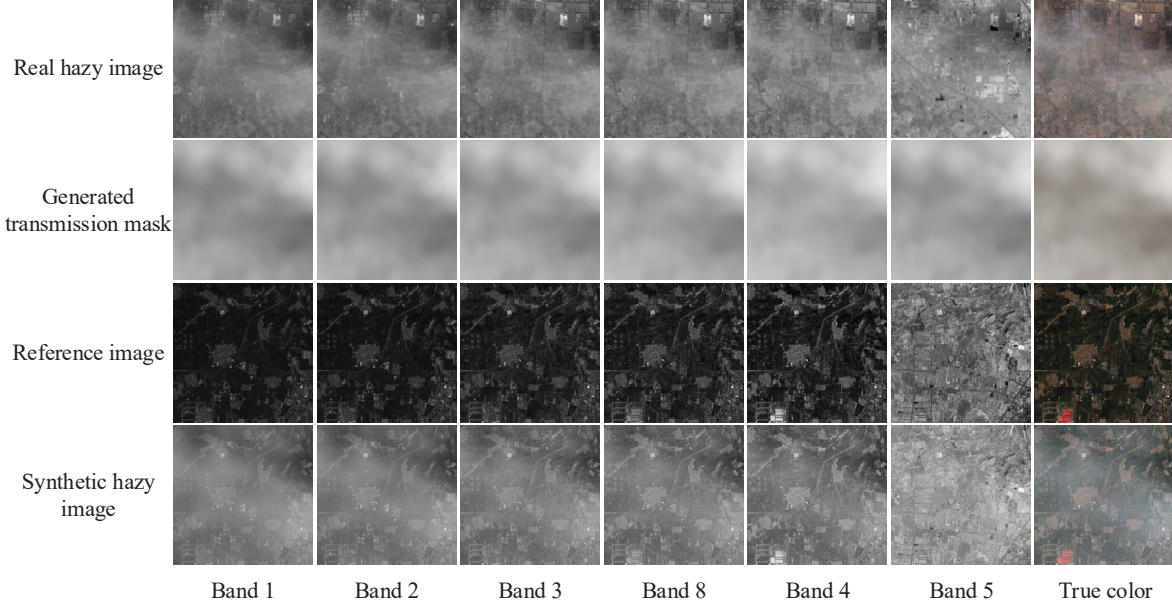


Fig. 6. Synthetic image with nonuniform haze.

of the real image in vision, and more importantly, the haze correlation between different bands is preserved.

V. EXPERIMENT

In this paper, the number of CNN individuals n in the designed dehazing network (see Fig. 1) is set to 5. The designed network is trained on a PC with Nvidia GeForce GTX 1080 GPU and tested in Python code. The whole work is implemented with Caffe package [28].

To verify the effectiveness of the proposed dehazing framework, the experiments are conducted in three aspects: the effectiveness of the designed network architecture, the dehazing performance of the proposed framework, and the influence of the haze simulation method on dehazing results.

A. Datasets for training and testing

Since the collection of hazy images and their corresponding haze-free images from multispectral remote sensing images is difficult, we use the haze synthesis method to produce labeled data for training and testing.

Synthetic images with uniform haze are used to train the network to ensure the diversity and balance of haze samples [19] [20]. We collected 400 haze-free Landsat 8 OLI images with size 200×200 and 6 spectral bands from the website (<http://ids.ceode.ac.cn/query.html>). Following the description of haze synthesis method in Section IV-B, the standard transmission mask t for band 1 (standard band) is set to global-constant and changed from 0 to 1 with interval 0.1. For each transmission mask of band 1, the corresponding transmission masks of other 5 bands are calculated using (10). γ is set to 0.5, 0.7 and 1 according to [5]. Thus, $10 \times 3 = 30$ sets of transmission masks are generated, and each set includes 6 transmission masks corresponding to the 6

bands. Adding these masks to 400 clear images according to (1) respectively, totally $30 \times 400 = 12000$ hazy images with different haze densities are obtained. Of the 12000 images, 9000 images are used as training set to train the network, and the remained 3000 images are used as validation set to determine the optimal parameters. In the first training stage, the training set and validation set are divided into five groups according to the value of t , where the values of t for the five groups are $\{0.1, 0.2\}$, $\{0.3, 0.4\}$, $\{0.5, 0.6\}$, $\{0.7, 0.8\}$, $\{0.9, 1\}$, respectively. Each group is used to train a CNN individual of the designed network. In the second training stage, the five groups are merged.

The test set is used for qualitative and quantitative assessment. Nonuniform haze is more common and more challenging to be removed. We synthesize images with nonuniform haze to verify the effectiveness of the proposed method. We extract 8 transmission maps from real hazy images with different haze distributions using the method in [27]. Using the 8 extracted transmission maps, we generate 80 standard transmission masks for band 1 through $t_1 = m_i \times k$, where m_i is the i th extracted transmission map, and k is a coefficient varying from 0 to 1 with interval 0.1. Fixing γ to 1 and calculating the corresponding transmission masks of other 5 bands using (10), totally 80 sets of transmission masks are obtained (each set includes 6 channels). Adding them on 100 haze-free images respectively (do not overlap with the 400 haze-free images used in the training set), the test set including $80 \times 100 = 8000$ images with nonuniform haze is established.

B. Verification of network architecture

In our proposed dehazing framework, five CNN individuals are used to learn the regression from the hazy image to the clear image, which have different dehazing abilities, and then

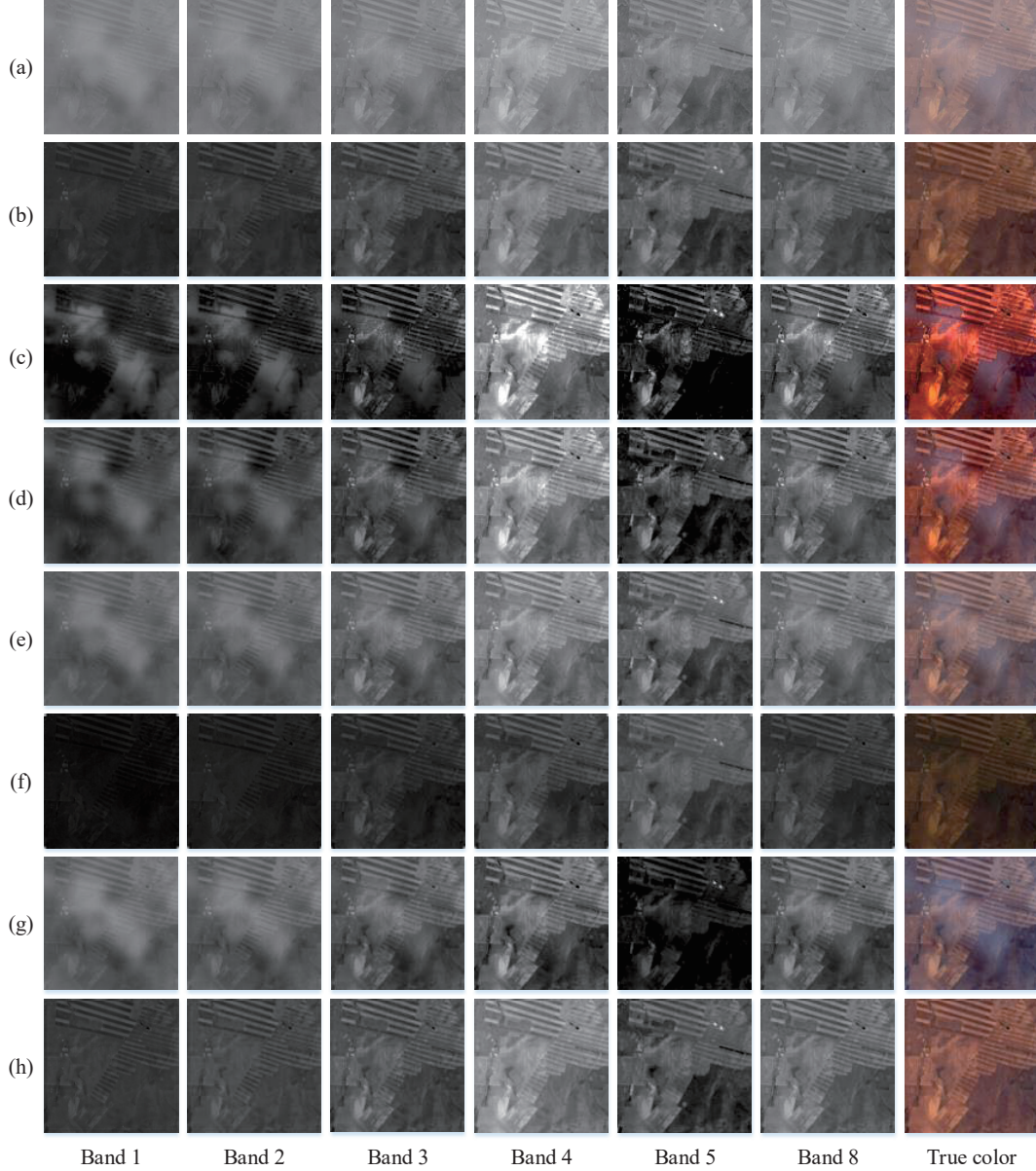


Fig. 7. Dehazing instance for a synthetic hazy image. (a) Synthetic image. (b) Ground truth. From (c) to (h) are Long's, Pan's, Feng's, Makarau's, Cai's and our results, respectively.

TABLE II
EUCLIDEAN DISTANCE FOR DIFFERENT NETWORKS ON VALIDATION SET

Network type	Euclidean Distance
Non-residual	1326.76
Simple residual	1303.06
Multiscale residual	1266.19
Cascade connection of 5 CNNs	1228.26
Parallel connection of 5 CNNs	1085.16
Proposed	983.78

their outputs are combined with weight maps to generate the final restored image. Each of the CNN individuals is with residual structure and uses multiscale convolutions to extract the haze feature. To demonstrate the effectiveness

and performance of our network architecture, we compare it with five counterparts: non-residual network, simple residual network, multiscale residual network (a CNN individual of the designed network in Fig. 1), the cascade connection of 5 CNN individuals, and the parallel connection of 5 CNN individuals (the designed network without the guidance of weight maps, see Fig. 1), where the first three networks are single networks and they have the same number of layers. For the cascade connection network of 5 CNN individuals, the training is using the whole training set, while for the parallel connection network, the training set is divided to 5 groups according to the haze thickness, and each group is used to train a CNN individual.

We use the Euclidean Distance between the dehazed image

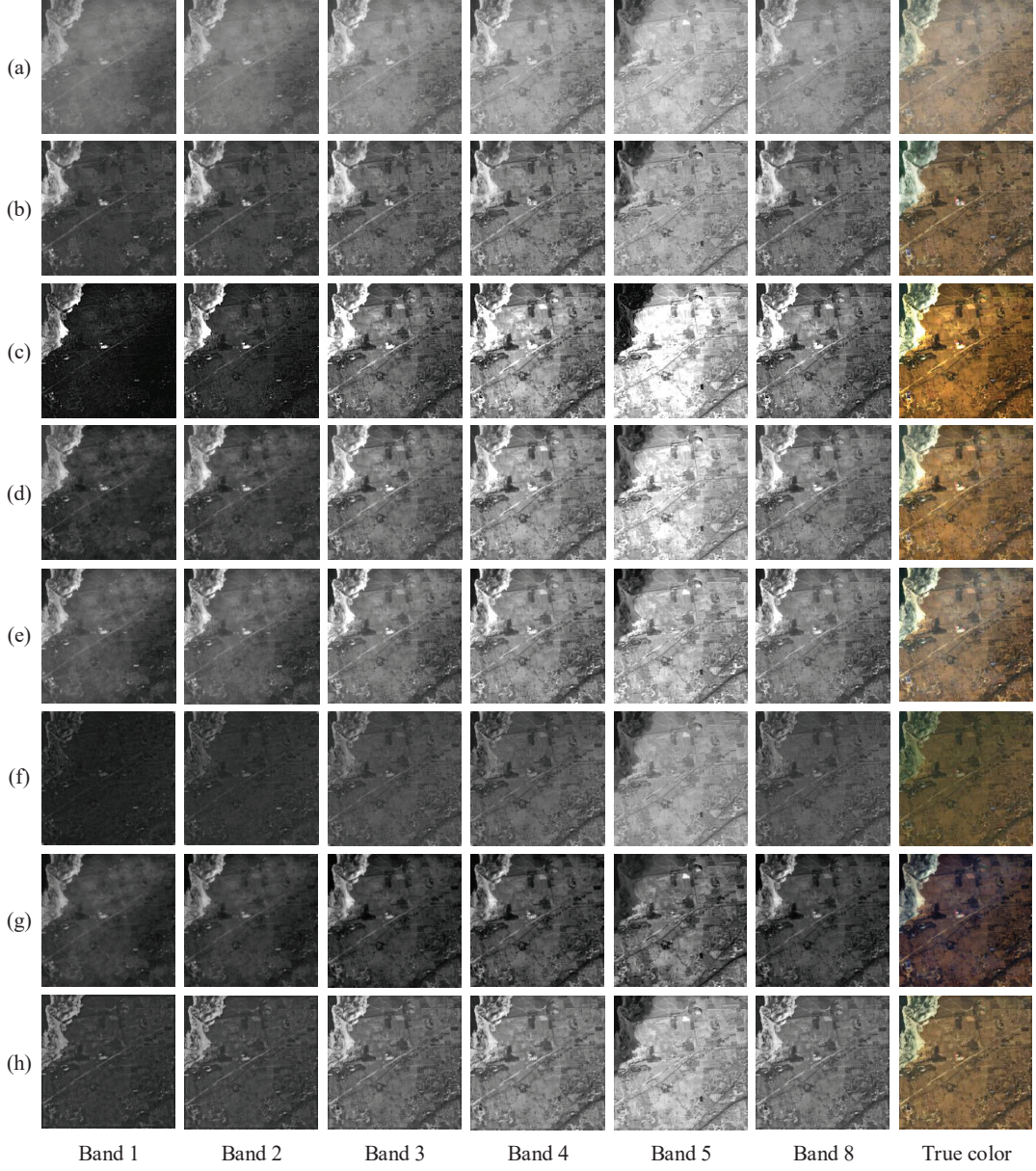


Fig. 8. Dehazing instance for a real hazy image. (a) Real hazy image (ID: LC81230322016365LGN00). (b) Reference image collected in the same scene at a different time (ID: LC81230322016333LGN00). From (c) to (h) are Long's, Pan's, Feng's, Makarau's, Cai's and our results, respectively.

and the ground truth to evaluate the performances of these networks. A lower value of Euclidean Distance indicates a better performance. Table II gives each network's average result on the validation set. As can be seen, when using the single network to dehaze, the multiscale residual network (a CNN individual in Fig.1) outperforms the non-residual and simple residual network, which demonstrates the advantage of the multiscale features for dehazing. In addition, when multiple network individuals are connected in cascade or in parallel, the performance is improved, see the fifth and the sixth rows in Table II. Compared with the simple cascade connection network and parallel connection network, our final designed network fuses the outputs of 5 CNN individuals

under the guidance of weight maps, and achieves the best dehazing result.

C. Comparison with state-of-the-art methods

To verify the effectiveness of our proposed dehazing framework, we compare it with five state-of-the-art dehazing methods, including Long's [15], Pan's [17], Feng's [2], Makarau's [6] and Cai's [20] methods. In [20], Cai used a CNN to regress the transmission to dehaze for outdoor images. In this paper, we retrain Cai's network on multispectral remote sensing images and compare it with our designed network to verify the performance of our dehazing framework.

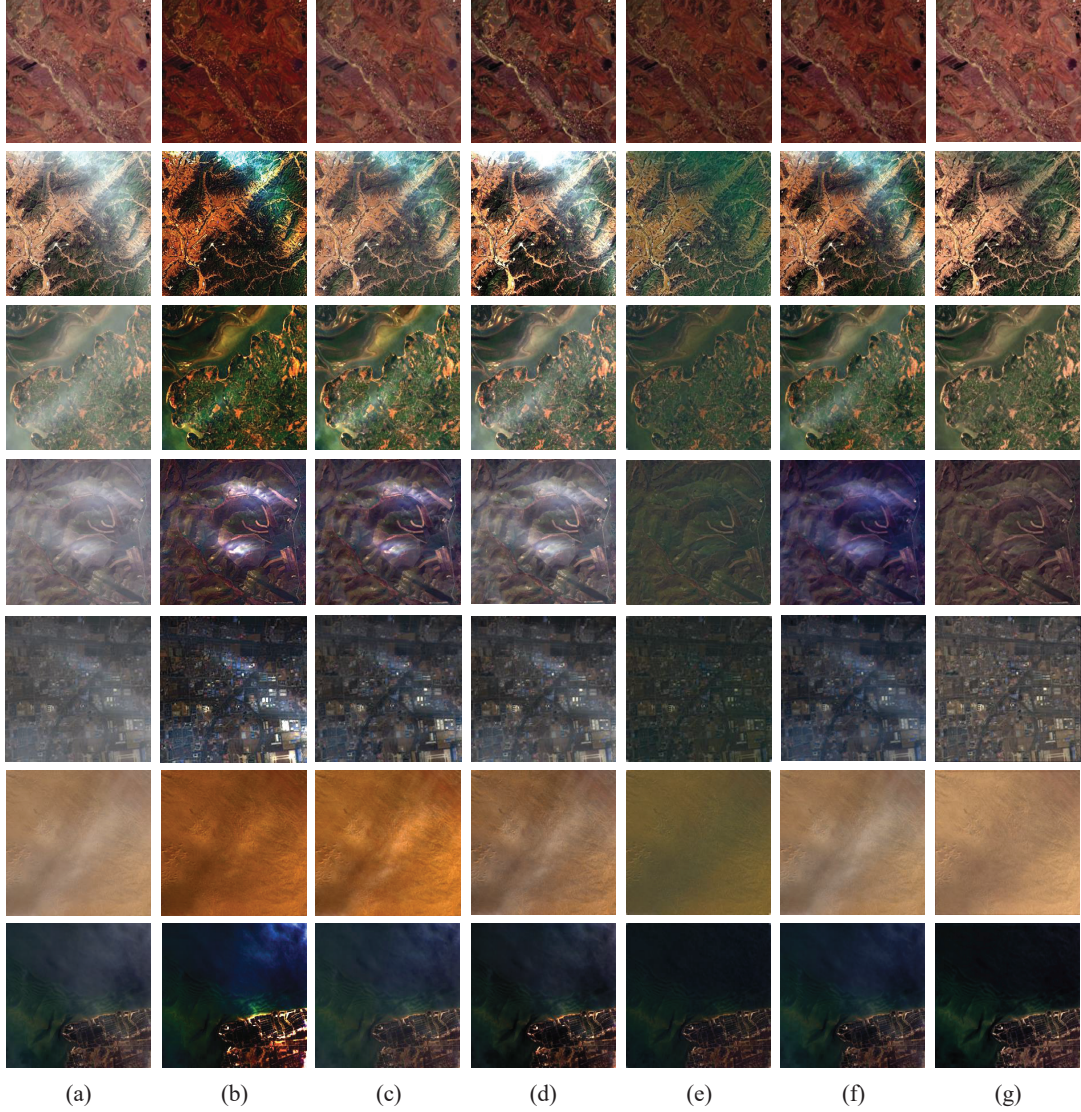


Fig. 9. Dehazing instances for real hazy images under different scenes. (a) Original image. From (b) to (g) are Long's, Pan's, Feng's, Makarau's, Cai's and our results, respectively.

TABLE III
VALUES OF EVALUATION INDICATORS ON TEST SET

Metric	MSE	SSIM	PSNR	WPSNR	UQI
Long's	0.0400	0.7220	15.5313	21.8908	0.7847
Pan's	0.0334	0.8211	17.5418	23.6010	0.7674
Feng's	0.0552	0.7923	14.4788	20.5485	0.7042
Makarau's	0.0172	0.8053	20.3359	26.4522	0.8235
Cai's	0.0197	0.8421	19.5578	25.8724	0.8192
Our	0.0074	0.8964	23.8243	30.0059	0.8986

1) *Comparison on synthetic images:* Synthetic images can provide quantitative and qualitative evaluations for dehazing methods. Fig. 7 shows a dehazing instance on a multispectral image using our method and five compared methods, where the first row is the synthetic hazy image, the second row is the reference image (ground truth), and the remaining six rows

are the dehazing results for the first row. Long's and Pan's methods were developed for the remote sensing images with three channels from Google Earth. When using them to remove haze from multispectral images, Long's method is failed with obvious spectral distortions and Pan's method cannot remove haze completely. Feng's and Makarau's methods were originally developed for multispectral images. Obviously, Feng's result is under-dehazed and Makarau's result is over-dehazed. As for the deep learning based methods, Cai's method used a CNN to estimate the transmission, which fails for the dense haze condition and leads to an under-dehazing result, while our method directly regresses the clear scene and can correctly dehaze for each band. Therefore, among the six methods, our method can accurately restore the clear scene in each band, and the result is the closest to the reference image.

We calculate five metrics between dehazed images and

their reference images (ground truth) to quantitatively evaluate the performance of these dehazing methods, including Mean Squared Error (MSE), structural similarity (SSIM) [29], peak signal-to-noise ratio (PSNR), weighted peak signal-to-noise ratio (WPSNR) [30], and universal image quality index (UQI) [31]. The dehazing performance is negatively correlated with the MSE metric, and positively associated with the other four metrics. Table III gives the statistical results of these metrics on test set. Obviously, with the best values in all metrics, our method greatly outperforms the compared methods and achieves the best dehazing result.

2) *Comparison on real hazy images:* We verify the performance of our proposed framework on real hazy images. Fig. 8 shows a dehazing instance for a real image with nonuniform haze using the six dehazing methods, where the first row is the real hazy image (ID: LC81230322016365LGN00), and the second row is the corresponding haze-free image (ID: LC81230322016333LGN00) collected under the same scene at a different time, as reference. It can be seen that, Long's result is over-dehazed and its RGB composite image is over-saturated compared with the reference image. Some texture information is seriously lost in Band 1 and Band 5. Pan's and Feng's results are under-dehazed, and some haze is still remained in the first two bands. Makarau's and Cai's methods lead to color distortions to different degrees in the RGB composite images. What's more, Cai's method again under-estimates the haze, leading to under-dehazing result for Band 1 and Band 2. Different from the compared methods, our method properly removes haze in each band along with good color fidelity.

Fig. 9 shows true color composite images of more dehazing instances with different land cover types including red ground, village, forest, mountain, urban, desert and sea. The first row is an instance for a clear image. It can be seen that, Pan's and our results are close to the original image and achieve good color fidelity, while others are over-enhanced to different degrees. From the second row to the last row, the thickness and spatial distribution of the haze is varied, and as can be seen, our method is superior to the compared methods not only in dehazing ability but also in color fidelity.

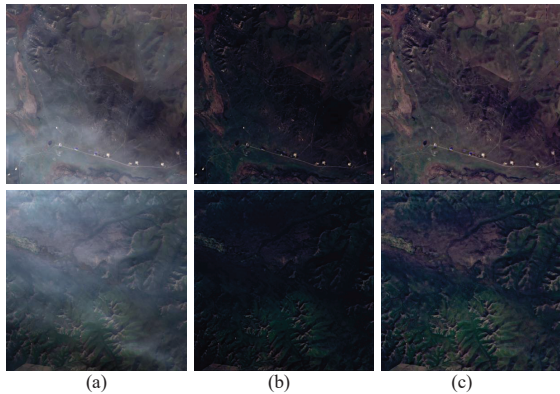


Fig. 10. The influence of haze synthesis methods on dehazing results. (a) Real hazy image. (b) and (c) are dehazing results using wavelength-independent and our wavelength-dependent haze synthesis methods, respectively.

D. Influence of haze simulation method on dehazing results

The performance of learning-based methods is highly dependent on the training data. Therefore, a good haze simulation method is of great importance to the dehazing results. In [19], [20] and [27], the wavelength-independent haze simulation methods were used to synthesize hazy images, in which all bands shared the same transmission. However, haze in multispectral images is wavelength-dependent according to the atmospheric scattering theory. In this paper, taking the wavelength correlation into account, we propose a new haze synthesis method.

We analyze the influence of different haze synthesis methods on dehazing results. When using a network trained by synthetic hazy images to remove haze from real images, the better the dehazing results, the closer the synthetic haze is to the real haze and the better the performance of the corresponding haze simulation method. Two dehazing networks are trained using the synthetic images generated by the wavelength-independent method and our wavelength-dependent method, respectively. Using the two trained dehazing networks to dehaze for real hazy images respectively, Fig. 10 shows the true color composite results of two dehazing instances. It can be seen that, the result images generated by our method have more details and higher color fidelity, which means that our synthesis method can generate hazy images closer to real conditions than the compared method. Therefore, using our synthetic hazy images to train the network, better dehazing results can be achieved.

VI. CONCLUSION

Haze often appears in multispectral remote sensing images, which decreases the image visibility and causes wrong interpretation. In this paper, a novel haze removal method based on deep CNN is proposed for multispectral remote sensing images. The designed network includes two parts. One is the parallel connection of multiple CNN individuals with residual structure. Each individual is used to learn a regression from the hazy image to the clear image, which mines multiscale features through multiscale convolutions to obtain stronger representation for haze. Adopting the residual structure in the designed network, only the haze component is learned and represented by convolutional layers, and the complex texture and color information can be directly provided by the input image. Thus, the learning difficulty is decreased. These individuals are trained using different levels of haze samples, and they have different dehazing abilities. The second part of the designed network is the fusion of multiple CNN individuals to generate the final dehazed image. We calculate the weight maps to guide the fusion of these individuals' outputs. Because haze is spatially-varying, the fusion model is changed with the haze distribution, and the haze in the image is adaptively removed. Learning-based dehazing methods always adopt haze synthesis manner to produce abundant labeled data for training, and a good haze simulation method is of great importance to dehazing results. In this paper, according to the atmospheric scattering model and Rayleigh's law, the relational expression of transmissions between different bands is derived, and a wavelength-dependent haze simulation method is proposed.

The proposed haze synthesis method can generate haze highly close to real conditions, using which to train the dehazing network, more accurate dehazing results can be obtained. Qualitative and quantitative experiments are conducted on multispectral remote sensing images from Landsat 8 OLI. Results indicate that compared with the state-of-the-art methods, our proposed dehazing method can effectively remove haze in each band of multispectral images under different scenes.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant 61471016) and SAST innovation foundation (Grant SAST201417).

REFERENCES

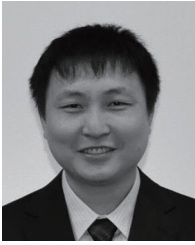
- [1] R. Richter, "Atmospheric correction of satellite data with haze removal including a haze/clear transition region," *Computers & Geosciences*, vol. 22, no. 6, pp. 675–681, 1996.
- [2] F. Chun, M. Jian-wen, D. Qin, and C. Xue, "An improved method for cloud removal in aster data change detection," in *Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International*, vol. 5. IEEE, 2004, pp. 3387–3389.
- [3] Y. Du, B. Guindon, and J. Cihlar, "Haze detection and removal in high resolution satellite image with wavelet analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 1, pp. 210–217, 2002.
- [4] C. Liu, J. Hu, Y. Lin, S. Wu, and W. Huang, "Haze detection, perfection and removal for high spatial resolution satellite imagery," *International journal of remote sensing*, vol. 32, no. 23, pp. 8685–8697, 2011.
- [5] P. S. Chavez, "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data," *Remote sensing of environment*, vol. 24, no. 3, pp. 459–479, 1988.
- [6] A. Makarau, R. Richter, R. Muller, and P. Reinartz, "Haze detection and removal in remotely sensed multispectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5895–5905, 2014.
- [7] Y. Zhang, B. Guindon, and J. Cihlar, "An image transform to characterize and compensate for spatial variations in thin cloud contamination of landsat images," *Remote Sensing of Environment*, vol. 82, no. 2, pp. 173–187, 2002.
- [8] G. D. Moro and L. Halounova, "Haze removal for high-resolution satellite data: a case study," *International Journal of Remote Sensing*, vol. 28, no. 10, pp. 2187–2205, 2007.
- [9] X. Y. He, J. B. Hu, W. Chen, and X. Y. Li, "Haze removal based on advanced haze-optimized transformation (ahot) for multispectral imagery," *International Journal of Remote Sensing*, vol. 31, no. 20, pp. 5331–5348, 2010.
- [10] H. Shen, H. Li, Y. Qian, L. Zhang, and Q. Yuan, "An effective thin cloud removal procedure for visible remote sensing images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 96, pp. 224–235, 2014.
- [11] H. Li, L. Zhang, H. Shen, and P. Li, "A variational gradient-based fusion method for visible and swir imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 78, no. 9, pp. 947–958, 2012.
- [12] A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmio, "Fusion-based variational image dehazing," *IEEE Signal Processing Letters*, vol. 24, no. 2, pp. 151–155, 2017.
- [13] X. Lan, L. Zhang, H. Shen, Q. Yuan, and H. Li, "Single image haze removal considering sensor blur and noise," *EURASIP journal on advances in signal processing*, vol. 2013, no. 1, p. 86, 2013.
- [14] W. Ni, X. Gao, and Y. Wang, "Single satellite image dehazing via linear intensity transformation and local property analysis," *Neurocomputing*, vol. 175, pp. 25–39, 2016.
- [15] J. Long, Z. Shi, W. Tang, and C. Zhang, "Single remote sensing image dehazing," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 1, pp. 59–63, 2014.
- [16] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.
- [17] X. Pan, F. Xie, Z. Jiang, and J. Yin, "Haze removal for a single remote sensing image based on deformed haze imaging model," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1806–1810, 2015.
- [18] Q. Liu, X. Gao, L. He, and W. Lu, "Haze removal for a single visible remote sensing image," *Signal Processing*, vol. 137, pp. 33–43, 2017.
- [19] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2995–3000.
- [20] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.
- [21] H. Koschmeider, "Theorie der horizontalen sichtweite. beifs," *Phys. frei*, pp. 33–53, 1924.
- [22] E. J. McCartney, "Optics of the atmosphere: scattering by molecules and particles," *New York, John Wiley and Sons, Inc.*, 1976. 421 p., 1976.
- [23] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," *IEEE transactions on pattern analysis and machine intelligence*, vol. 25, no. 6, pp. 713–724, 2003.
- [24] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [26] J. Kim, J. Kwon Lee, and K. Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1646–1654.
- [27] X. Pan, F. Xie, Z. Jiang, Z. Shi, and X. Luo, "No-reference assessment on haze for remote-sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1855–1859, 2016.
- [28] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 2014, pp. 675–678.
- [29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [30] J. Mannos and D. Sakrison, "The effects of a visual fidelity criterion of the encoding of images," *IEEE transactions on Information Theory*, vol. 20, no. 4, pp. 525–536, 1974.
- [31] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002.



Manjun Qin received the B.Eng. degree from University of Science and Technology Beijing in 2015. She is a postgraduate majored in Pattern Recognition and Intelligent System at School of astronautics in Beihang University. Her research interests include remote sensing image processing, image quality assessment, and deep learning.



Fengying Xie received the Ph.D. Degree in Pattern Recognition and Intelligent System from Beihang University, Beijing, China in 2009. She was a visiting scholar in the Laboratory for Image and Video Engineering (LIVE) at the University of Texas at Austin from 2010 to 2011. She is now a professor at the School of Astronautics in Beihang University. Her research interests include biomedical image processing, remote sensing image understanding and application, image quality assessment, and pattern recognition. Her homepage is: <http://xfy.buaa.edu.cn/>.



Wei Li received the Ph.D. degree in Signal and Information Processing from Beihang University, China, in 2011. He is currently a senior engineer with Shanghai Institute of Satellite Engineering. His major research interests include satellite system design and system performance analysis.



Zhenwei Shi (M'13) received the Ph.D. degree in mathematics from the Dalian University of Technology, Dalian, China, in 2005. He was a Post-doctoral Researcher with the Department of Automation, Tsinghua University, Beijing, China, from 2005 to 2007. He was a Visiting Scholar with the Department of Electrical Engineering and Computer Science, Northwestern University, USA, from 2013 to 2014. He is currently a Professor and the Dean of the Image Processing Center, School of Astronautics, Beihang University, Beijing, China. He has

authored or coauthored more than 100 scientific papers in refereed journals and proceedings, including the IEEE Transactions on Pattern Analysis and Machine Intelligence, the IEEE Transactions on Neural Networks, the IEEE Transactions on Geoscience and Remote Sensing, the IEEE Geoscience and Remote Sensing Letters, and the IEEE Conference on Computer Vision and Pattern Recognition. His current research interests include remote sensing image processing and analysis, computer vision, pattern recognition, and machine learning.



Haopeng Zhang received his B.S. and Ph.D. degrees from Beihang University, Beijing, China, in 2008 and 2014, respectively. He is currently an assistant professor in the Image Processing Center, School of Astronautics, Beihang University. He is a member of IEEE. His main research interests include remote sensing image processing, multi-view object recognition, 3D object recognition and pose estimation, and other related areas in pattern recognition, computer vision, and machine learning.