IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 14, 2019, accepted March 12, 2019, date of publication March 20, 2019, date of current version April 9, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2905615

No-Reference Quality Assessment for Pansharpened Images via Opinion-Unaware Learning

BINGZHONG ZHOU[1,2], FENG SHAO[1], (Member, IEEE), XIANGCHAO MENG[1], (Member, IEEE), RANDI FU[1], AND YO-SUNG HO[3], (Fellow, IEEE)

[1]Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China
[2]Department of Information Technology, Wenzhou Vocational & Technical College, Wenzhou 325035, China
[3]School of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500712, South Korea

Corresponding author: Feng Shao (shaofeng@nbu.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 61622109 and Grant 41801252, in part by the Zhejiang Natural Science Foundation of China under Grant R18F010008, in part by the Natural Science Foundation of Ningbo under Grant 2017A610112, and in part by the by K. C. Wong Magna Fund in Ningbo University.

ABSTRACT The high-quality pansharpened image with both high spatial resolution and high spectral fidelity is highly desirable in various applications. However, existing pansharpening methods may lead to spatial distortion and spectral distortion. To measure the degrees of distortion caused by the pansharpening methods, we conduct in-deep studies on the subjective and objective quality assessment of pansharpened images. We built a subjective database consisting of 360 images generated from 20 couples of panchromatic (PAN)/multispectral (MS) images using 18 pansharpening methods. Based on the database, we proposed a no-reference quality assessment method to blindly predict the quality of pansharpened images via opinion-unaware learning. The proposed method first extracted features from the MS images' spectral bands and typical information indexes which comprehensively reflect spatial distortion, spectral distortion, and the effects of pansharpening on applications. Based on the features extracted from the pristine MS image training dataset, a benchmark multivariate Gaussian (MVG) model is learned. The distance between the benchmark MVG and the MVG fitted on the test image is calculated to measure the quality. The experimental results show the superiority of our method on our database.

INDEX TERMS Pansharpening, no-reference quality assessment, spatial distortion, spectral distortion, remote sensing.

I. INTRODUCTION

Pansharpened images with high spatial and spectral resolution in remote sensing are highly desirable in many fields, such as visual interpretation and change detection. Typically, multispectral (MS) images have the high spectral resolution but low spatial resolution, while panchromatic (PAN) images have the high spatial resolution but low spectral resolution. The purpose of pansharpening is to achieve spatially enhanced MS images with the same spectral resolution of the MS and the same spatial resolution of the PAN by fusing the MS and PAN images [1]–[4], which is important for remote sensing applications.

In the past decades, a large number of pansharpening methods have been proposed, which can be broadly classified into three categories: component substitution (CS)-based methods [3], [5]–[10], multiresolution analysis (MRA)-based methods [11]–[20] and variational optimization (VO)-based methods [21]–[34]. Besides, deep learning (DL)-based pansharpening methods [35]–[37] have been proposed in recent years. Though there have proposed large numbers of pansharpening methods, how to effectively evaluate the quality of the pansharpened image has not been well addressed in the current research. Especially the no reference image quality evaluation is still a challenging problem, due to the lack of the ideal high spatial resolution reference MS image.

Generally speaking, the quality of the pansharpened image is evaluated from two aspects: the qualitative evaluation and

The associate editor coordinating the review of this manuscript and approving it for publication was Marco Anisetti.

the quantitative evaluation. The qualitative evaluation performs the subjective assessment of the pansharpened image visually. The quantitative evaluation objectively evaluates the pansharpened image based on the quality evaluation indexes, including the full-reference quality evaluation indexes and the no-reference quality evaluation indexes. Among them, the full-reference quality evaluation is based on the Wald protocol that the performances of the fusion methods are invariant among different scales [1]. For this, both the original MS and PAN images are degraded to a lower resolution before pansharpening process, so that the original MS image can be used as a reference for the evaluation. Currently, quality measurements for pansharpened images are developed from two aspects: distortion-based measurement and application-based measurement. For the first aspect, the widely used full-reference quantitative evaluation indexes include the correlation coefficient (CC) [3], the spectral angle mapper (SAM) metric [38], the Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS) index [39], the $Q2^n$ - index [40], etc. Compared to the full-reference image quality evaluation indexes, the no-reference quantitative evaluation of the pansharpened image is more challenging, due to the lack of the reference image for pansharpening. The typical no-reference methods include the Quality with No Reference (QNR) [41], and generalized QNR (GQNR) index [42], etc. For the second aspect, some widely used information indexes, such as normalized difference vegetation index (NDVI) [43], normalized difference water index (NDWI) [44] and RatioG [45], [46], are applied to evaluate the information presentation ability of fusion products. Conversely, quality assessment for the pansharpened image is taken as an indicator to guide the design of different pansharpening methods.

In this work, we conduct in-deep studies on quality prediction issue for pansharpened images from both subjective and objective perspectives. For this purpose, a subjective database of pansharpened images was built, which consists of 360 pansharpened images generated from 20 couples of PAN/MS images using 18 pansharpening methods. Based on the constructed dataset, we propose a no-reference image quality assessment (NR-IQA) method for pansharpened images. By analyzing the statistical properties of the pristine MS images and pansharpened images, features that comprehensively reflect spatial distortion, spectral distortion and remote sensing information indexes are extracted. The features extracted from the pristine MS training dataset are used to learn an MVG model as a benchmark. By computing the distance between the benchmark MVG and the MVG fitted on the testing image, the overall quality score of a testing pansharpened image is then obtained by pooling the distances from all patches.

Our work is inspired by the Natural Image Quality Evaluator (NIQE) model [47] and the Integrated Local NIQE (IL-NIQE) model [48], but it performs much better than NIQE and IL-NIQE for the following contributions:

1) A large-scale subjective performance evaluation database for pansharpening is firstly constructed.

2) This paper proposed a new no-reference image quality evaluation method by comprehensively considering the spatial and spectral distortions of the fused image. Furthermore, typical information indexes, including NDVI, NDWI and RatioG, are integrated to a form a quality-aware feature representation.

3) Different from the existing quality assessment methods for pansharpened images that need degrade the fused MS image to the same resolution of the reference image, our method does not require the reference images and ground-truth subjective scores or down-sample fused MS images, so that completely blind quality prediction can be achieved.

The rest of the paper is organized as follows: Section II introduces the related work. Section III describes the construction of the subjective database. Section IV presents the proposed quality assessment method. The experimental results are given and discussed in Section V, and finally, conclusions are drawn in Section VI.

II. RELATED WORK

A. PANSHARPENING METHODS

Pansharpening for remote sensing has been widely researched in the last thirty decades [3], [5]–[37]. Broadly, the pansharpening methods can be classified into three categories as suggested in [2] and [49]: CS-based pansharpening methods [3], [5]–[10], MRA-based pansharpening methods [11]–[20] and VO-based methods [21]–[34]. Besides, deep learning (DL)-based pansharpening methods [35]–[37] have been proposed in recent years, which can be regarded as a new type of pansharpening methods.

The CS-based methods are the simplest and most popular pansharpening methods, which have experienced the process from the traditional understanding to the general understanding. In the traditional understanding, the MS bands are first projected into a new space based on the spectral transformation and one component is substituted by the high-resolution PAN image. Then, an inverse projection is performed to obtain the fused image. The representative methods include Intensity-Hue-Saturation (IHS) fusion method [3], Principal Component Analysis (PCA) fusion method [5], and Gram Schmidt (GS) fusion method [6], etc. Subsequently, the CS-based methods have been generalized to a new formalization without explicit calculation of the forward and backward transformation. The general understanding is based on the simple substitution of a single component by the PAN image, and the component is generally obtained by a linear combination of the spectral bands of the MS images, such as Brovey method [7], Gram Schmidt Adaptive (GSA) method [8], Band-Dependent Spatial-Detail (BDSD) method [9], and Partial Replacement Adaptive Component Substitution (PRACS) method [10].

The MRA-based pansharpening methods are based on the spatial detail injection of the PAN image into the resampled MS image. In the early stage, these methods are based on high-pass filters, such as the High-Pass filter (HPF)

method [5] and smoothing filter-based intensity modulation (SFIM) method [11]. Subsequently, for better spectral preservation, methods based on different kinds of discrete wavelet transform were proposed. These methods include Additive À Trous Wavelet Transform with unitary injection model (ATWT) [12], À Trous Wavelet Transform using the Model 2 (ATWT-M2) [13], À Trous Wavelet Transform using the Model 3 (ATWT-M3) [13], Additive Wavelet Luminance Proportional (AWLP) [14], [15], etc. In addition, the Generalized Laplacian Pyramid (GLP) methods using filters to match the modulation transfer function (MTF) of the sensor are proposed, such as the MTF-GLP [16], MTF-GLP with context-based decision (MTF-GLP-CBD) [17], MTF-GLP with high pass modulation (MTF-GLP-HPM) [18], MTF-GLP-HPM with Post-Processing (MTF-GLP-HPM-PP) [19], etc.

The VO-based methods are generally based on the optimization of a variational model. The VO-based methods include two main parts: (1) the construction of the energy functional; and (2) the optimization solution. For the construction of the energy functional, the most widely used methods are based on the observation models [21]–[27] and the sparse representation models [28]–[31]. Among them, the observation model-based methods take the fusion process as an ill-posed inverse optimization problem. Based on the relationship between the desired fused image and MS/PAN images, an energy function is established, and a high-resolution fused image is obtained by optimizing the solution. The sparse-based methods are mainly based on sparse representation theory. It is assumed that the information of the remote sensing images is sparse in a basis set, and can be represented by a linear combination of relatively few base elements. The optimization solution of the energy functional is generally based on iterative optimization algorithms [32]–[34].

In addition to the above three classes of pansharpening methods, in recent years, several DL-based pansharpening methods are proposed. Huang *et al.* [35] proposed a pansharpening method based on deep neural network. Masi *et al.* [36] proposed convolutional neural networks-based pansharpening method. Yuan *et al.* [37] proposed the multiscale and multidepth convolutional neural network for pansharpening of remote sensing images.

Since CS-based methods and MRA-based methods are more widely used, our work mainly focuses on these two kinds of methods in this paper.

B. TRADITIONAL NR-IQA METHODS

Recent studies on NR-IQA focused on blind image quality assessment (BIQA) problem via ''opinion-aware'' or ''opinion-unaware'' learning. The representative ''opinion-aware'' BIQA methods include Blind Image Quality Indexes (BIQI) [50], Distortion Identification-based Image Verity and Integrity Evaluation (DIIVINE) [51], Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [52], Blind Image Integrity Notator using DCT Statistics-II (BLIINDS-II) [53], etc. These methods share a similar architecture that trains a regression model on a dataset consisting of distorted images and the corresponding subjective scores. However, the ''opinion-aware'' models need human subjective scores to learn the regression model, usually having weak generalization capability. In practice, the subjective scores for pansharpened images are hard to obtain, leading to poor applicability of ''opinion-aware'' models in quality assessment of pansharpened images. In contrast, ''opinion-aware'' BIQA model does not need training samples with human subjective scores. Mittal *et al.* [47] proposed a NIQE model, in which a set of local features are first extracted from images and fitted to a multivariate Gaussian (MVG) model. The quality is predicted by measuring the distance between its MVG model and the benchmark MVG model learned from pristine training images. Zhang *et al.* [48] proposed the IL-NIQE model, which integrates natural image statistics features derived from multiple cues to learn an MVG model. Considering the advantages of opinion-aware BIQA methods, it is of great interest to utilize ''opinion-aware'' model for quality assessment of pansharpened images.

C. IQA METHODS FOR PANSHARPENED IMAGES

The pansharpened images suffer from two main types of distortions: spectral distortion and spatial distortion. On one hand, some quality evaluation indexes mainly concern on the spectral distortion. Yuhas *et al.* [38] proposed a Spectral Angle Mapper (SAM), which can be used to measure the spectral distortion of a pansharpened image. Wald [39] proposed the Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS). Alparone *et al.* [54] proposed the *Q4* method which is an extension of the universal image quality index (UIQI) or *Q*-index [55] for pansharpened images. Garzelli and Nencini [40] proposed a generalization of the Q4-index which is called $Q2^n$-index. It is suitable to measure quality for multiband images having more than four spectral bands. On the other hand, some quality evaluation indexes mainly focus on spatial distortions. Zhou *et al.* [56] proposed to evaluate the spatial quality of the fused image based on the correlation coefficient between high frequency of the fused image and the PAN image. However, since the spatial distortion is usually co-existed with the spectral distortion, it is incomplete to assess the quality of pansharpened images only concerning the spectral distortion or the spatial distortion. To date, there are many works on both spectral and spatial distortions. Alparone *et al.* [41] proposed the QNR index to predict quality from both the spatial distortions and spectral distortions without reference. In [4], the existing pansharpening methods are thoroughly analyzed, and the quality of the fused image generated by various pansharpening methods is synthetically measured by SAM, root-mean-square error (RMSE), UIQI, Q4 and QNR methods. Palsson *et al.* [57] used full reference (FR) IQA methods (ERGAS, SAM, Q4/Q8 and Q) to measure the quality of pansharpened images with the degrading step consisting of synthesis property and consistency property, and adopted QNR

method [41] to evaluate the pansharpened images without any degrading steps. Based on QNR method, Khan *et al.* [58] proposed a no reference method optimizing the QNR spatial index. Kwan *et al.* [42] proposed the GQNR index for the fused WorldView-3 images. Hasanlou and Saradjian [59] propose to evaluate the quality of fused images based on the weighted average of the geometric index and radiation index. Rodriguez-Esparragon *et al.* [60] proposed an object-based quality assessment scheme for pansharpened remote sensing imagery, which extracted the segments of the images and assessed both the spectral and spatial properties of fused remote sensing images. Selva *et al.* [61] demonstrated that adopting the expanded image as reference is erroneous based on Wald's protocol.

There are some works in analyzing the effects of pansharpening on special applications. For change-detection (CD) applications, Bovolo *et al.* [62] found that change-detection maps computed from pansharpened data suffer from errors due to artifacts induced by the fusion processes. For land cover classification applications, Ibarrola-Ulzurrun *et al.* [63] assessed the influence of pansharpening techniques in obtaining precise vegetation maps and analyzed the accuracy assessment of the pixel-based classification algorithms for each fused image. Gilbertson *et al.* [64] found that pansharpening may lead to dramatic improvements in classification accuracy. For target detection, Garzelli *et al.* [65] compared different pansharpening methods by evaluating the performances of target detection on true MS and panchromatic data. In such applications, the effects on the corresponding information indexes play important roles. For example, Maglione *et al.* [66] used NDVI and NDWI derived from pansharpened images to extract coastline. In addition, there are some works focusing on the effects on information indexes. Johnson [67] evaluated the effects of pansharpened images on Vegetation Indexes (VIs) and found that the pansharpened MS image may lose some spatial information after VI calculations. Xu *et al.* [68] analyzed the relationship between pansharpened spectral indexes and soil total nitrogen (TN) and quantified the effect of pansharpened image based on the proposed soil TN models. Du *et al.* [69] compared Modified Normalized Difference Water Index (MDNWI) derived from four pansharpening algorithms.

However, the aforementioned methods only assessed one aspect of the pansharpened MS images. Recently, Li *et al.* [70] assessed the performances of eight state-of-the-art pansharpening methods using both quality indexes and information indexes, including NDVI, NDWI and morphological building index. However, the quality indexes and information indexes are evaluated separately.

III. DATABASE

To investigate quality evaluation of pansharpened images, we construct a new pansharpened image database, in which 360 pansharpened MS images are included. Subjective evaluation of these images is conducted to obtain the human opinion scores. In the next, we will describe the databases in details.

A. COLLECTION OF ORIGINAL AND PANSHARPENED MS IMAGES

We select the original MS images acquired from Quickbird sensor. The Quickbird sensor works in the visible and near-infrared spectrum range, obtaining 4-band multispectral components: Blue, Green, Red and near-infrared (NIR) bands, and the corresponding PAN band. The resolution cell is 2.44m×2.44m for MS bands and 0.61m×0.61m for PAN band. The sensor has the same spectral region for PAN and MS data, which guarantees good performance of various pansharpening methods. In the database, we select 20 couples of PAN/MS images as shown in Fig. 1, and use 18 different pansharpening methods to generate 360 pansharpened images. The scenes for the original MS images include typical ground objects, such as water, vegetation, buildings, and roads. The resolution of MS image is 512×512 pixels, and 2048×2048 pixels for the corresponding PAN image.

FIGURE 1. The reference images selected in the database (only RGB bands are displayed).

The pansharpening methods used in the database are listed in Table 1, from two typical CS and MRA based methods [4]. As shown by examples of fused images obtaining using different pansharpening methods in Fig.2, compared with the reference MS image in (b), there are two types of distortions in the pansharpened images: 1) Spectral distortion: The spectral distortions are quite obvious in the pansharpened images generated by BROVEY and IHS in (h) and (l), in which the color of the river and the riparian vegetation obviously deviate from the original MS images. The reason may be that the difference between the substituted component and the PAN is quite large. In addition, slight spectral distortions are also found in the pansharpened images generated by AWLP, HPF, INDUSION, MTF-GLP, MTF-GLP-HPM, MTF-GLP-HPM-PP and SFIM. 2) Spatial distortion: There are different degrees of spatial distortions in the pansharpened images generated by ATWT-M3, ATWT-M2, GS, GSA, MTF-GLP-HPM-CBD, PCA and PRACS, in which boundaries of the rivers and banks are blurred. Since spectral and spatial distortions are co-existed in a pansharpened image, quantifying their influence are very difficult. The existing methods evaluate the spectral and spatial distortions separately and integrate them to obtain a unified score [41], [42], [59]. From the perspective of subjective evaluation, it is meaningful to evaluate the two cues in a unified framework. In our subjective test, we evaluate the overall performance of a pansharpened image involving spectral and spatial distortions.

TABLE 1. Description of fusion methods used in the database.

Type	Method	Descriptions
CS	IHS[3]	Intensity-Hue-Saturation
	BROVEY[7]	Brovey transform
	PCA[5]	Principal Component Analysis
	GS[6]	Gram Schmidt
	GSA[8]	Gram Schmidt Adaptive
	BDSD[9]	Band-Dependent Spatial-Detail with local parameter estimation
	PRACS[10]	Partial Replacement Adaptive Component Substitution
MRA	HPF[5]	High-Pass Filtering with 5×5 box filter for 1:4 fusion
	SFIM[11]	High-Pass Modulation with 5×5 box filter
	Indusion[20]	Decimated Wavelet Transform using an additive injection model
	MTF-GLP[16]	Generalized Laplacian Pyramid (GLP) with MTF-matched filter with unitary injection model
	MTF-GLP-CBD[17]	GLP with MTF-matched filter and regression based injection model
	ATWT[12]	Additive À Trous Wavelet Transform with unitary injection model
	ATWT-M2[13]	À Trous Wavelet Transform using the Model 2
	ATWT-M3[13]	À Trous Wavelet Transform using the Model 3
	MTF-GLP-HPM[18]	GLP with MTF-matched filter [17] and multiplicative injection model [18]
	MTF-GLP-HPM-PP[19]	GLP with MTF-matched filter [17],multiplicative injection model and Post-Processing [19]
	AWLP[15][16]	Additive Wavelet Luminance Proportional

B. SUBJECTIVE TEST

We carried out the subjective experiment in the lab designed for subjective quality test at Ningbo University. The test environment and condition met ITU-R BT.500-11 [71] and ITU-R 1438 [72]. All the pansharpened images were displayed on a Samsung UA65F9000 65-inch Ultra HD 3D-LED TV. We adapted double stimulus continuous quality scale (DSCQS) test methodology in the experiment, in which the reference and test images are simultaneously presented (displayed side-by-side) on the screen. The interface for the subjective test is shown in Fig. 3. However, in the practical application, since the reference image with the same resolution of the test image cannot be obtained, according to Wald Protocol [1], we spatially degraded both the original MS and PAN images to a lower resolution so that the original MS image can be used as the reference. Since the scale ratio between PAN and MS images is 4 in our database, we degrade the original MS and PAN images by the factor 4.

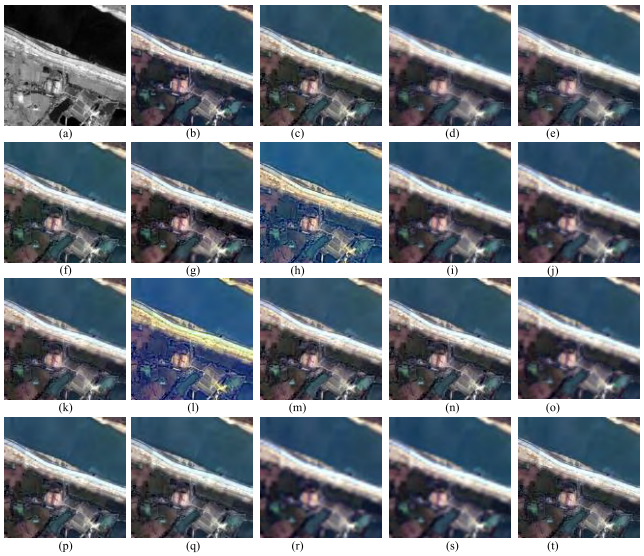


FIGURE 2. Examples of pansharpened images obtained by using different pansharpening methods.

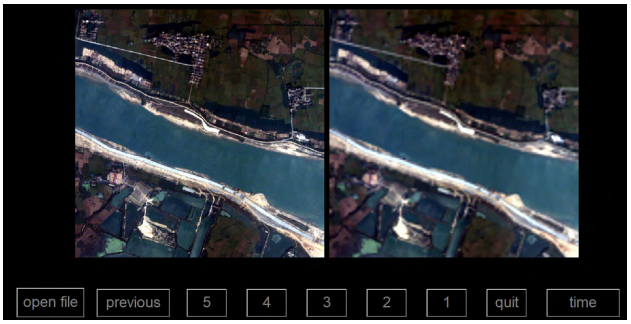


FIGURE 3. Subjective experiment interface.

Based on the properties of pansharpened images, grading criteria in our subjective study depend on two factors: spectral distortion and spatial distortion. Based on the grading criteria defined in Table 2, the participants were asked to rate the quality of images on a five-level scale: Level 1, Level 2, Level 3, Level 4 and Level5, corresponding to Excellent, Good, Fair, Poor and Bad. The descriptions of each level are shown in Table 2. We invite 40 graduate students participating in the subjective test.

After obtaining the raw subjective scores of all 360 pansharpened images from 40 subjects, the normalized z-scores is calculated by subtracting the mean and dividing by the standard deviation.

$$z_{i,j} = \frac{s_{i,j} - \mu_i}{\sigma_i} \tag{1}$$

where $s_{i,j}$ denotes the score assigned by subject i to the image j, and μ_i and σ_i denotes the mean and standard deviation calculated from all scores assigned by subject i. Then, after removing the outlier subjects (within 95% confidence interval), the normalized z-scores are scaled and mapped to

TABLE 2. Illustration of rating criteria.

Level	Descriptions
Level 1 (Excellent)	There is no blurring or color change compared with the original MS image. The edges of buildings, rivers and roads are clear.
Level 2 (Good)	There is small blurring or color change compared with the original MS image. The color changes have no influence on recognizing the object type.
Level 3 (Pair)	There is noticeable color change which has a little influence on recognizing the object type, or there is noticeable blurring on edges of buildings, rivers and roads.
Level 4 (Poor)	There is obvious color change partly influencing the recognition of object type, or the image is obviously blurred.
Level 5 (Bad)	There is serious blurring or color change over the entire image.

the range [0,100] to obtain MOS value for each sample:

$$z'_{i,j} = \frac{100(z_{i,j}+3)}{6} \tag{2}$$

$$\text{DMOS}_j = \frac{1}{U}\sum_i z'_{i,j} \tag{3}$$

where U is the number of final subjects. The DMOS score for each pansharpened image is used as the ground truth to represent the subjective quality. The distribution of the DMOSs in the database is shown in Fig. 4.

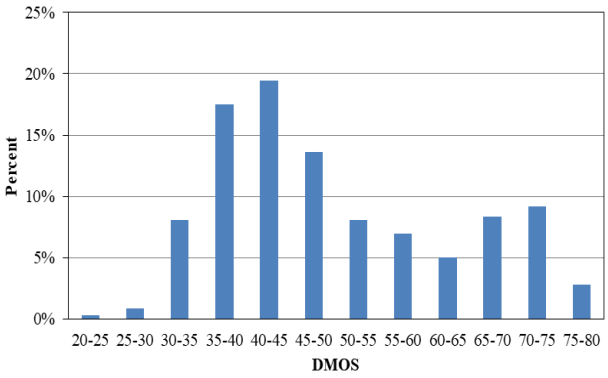


FIGURE 4. The DMOS distribution of the database.

IV. PROPOSED METHOD

As analyzed above, due to different properties of the MS and PAN in the process of fusing the pansharpened image, if we can directly extract the spatial and spectral features from the pansharpened images, the MVG models can be learned from the training and testing data. Followed the same architecture in [48], the processing flow of the proposed method is shown in Fig. 5. In the training stage, a benchmark MVG model is learned based on the features extracted from pristine MS training dataset. In the testing stage, the distance between the benchmark MVG and the MVG fitted on the testing

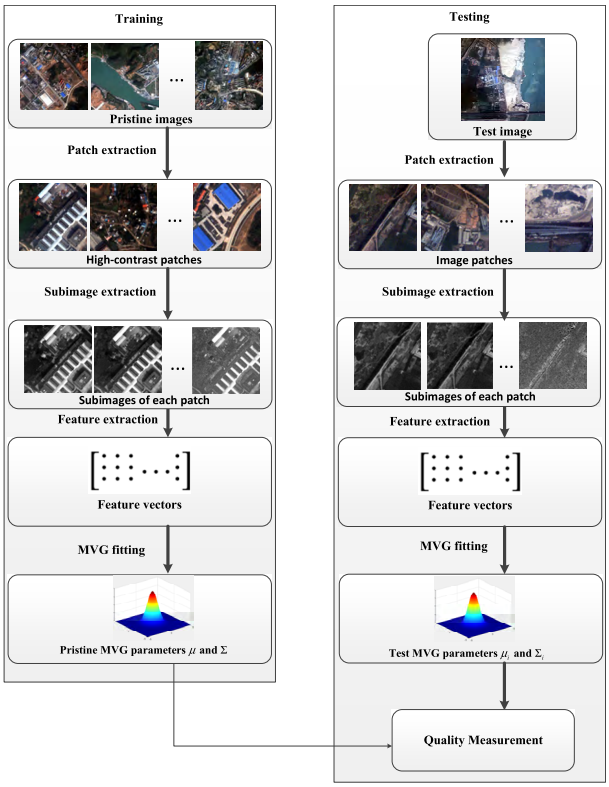


FIGURE 5. Flowchart of the proposed method.

pansharpened image is calculated as the final quality score. In the next, we will give the details for each step.

A. ANALYSIS OF PANSHARPENED IMAGES

A typical MS image is composed of 4 bands: red, green, blue and infra-red (NIR) bands. Each band data can be regarded as an individual image (subimages of MS image). The quality of pansharpened MS image can be directly determined from the quality of each band. However, each ground object has its own spectral properties, which can be measured by ground objects indexes. If there are obvious changes on the properties of the typical ground objects after pansharpening, the corresponding ground objects indexes or spectral features will reflect such changes. The typical ground objects include vegetation, water, roads and buildings.

To detect vegetation growth status, vegetation coverage and other information, the normalized difference vegetation index (NDVI) [43] is calculated as:

$$NDVI = \frac{NIR - R}{NIR + R} \tag{4}$$

To extract water information from images, the normalized difference water index (NDWI) [44] is calculated as:

$$NDWI = \frac{G - NIR}{G + NIR} \tag{5}$$

To distinguish roads and buildings, the RatioG [45], [46] is calculated as:

$$RatioG = \frac{G}{R+G+B+NIR} \tag{6}$$

We show an example to demonstrate the distributions of NDVI, NDWI and RatioG for the original and pansharpened image in Fig. 6.

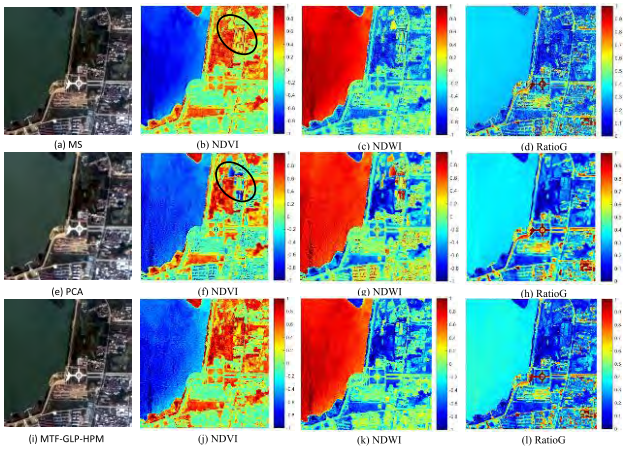


FIGURE 6. Comparison of the original MS image and pansharpened (PCA and MTF-GLP-HPM) MS image among the objects feature images. (a) MS. (b-d) the NDVI, NDWI and RatioG of (a). (e) PCA. (f-h) the NDVI, NDWI and RatioG of (e). (i) MTF-GLP-HPM. (j-l) the NDVI, NDWI and RatioG of (i).

It is clearly observed that the distributions of NDVI, NDWI and RatioG are changed caused by spectral distortions after pansharpening (obviously clear in the areas marked by black circles), which denotes NDVI, NDWI and RatioG can provide additional information to compensate the missing information in red, green, blue and NIR bands. As a result, we obtain seven subimages (e.g., red, green, blue, NIR, NDVI, NDWI and RatioG) to represent an MS image, and the following feature representation is performed for each subimage independently.

B. FEATURE REPRESENTATION

We observed that the distributions of locally normalized subimages of a pristine **I**, which is generated from a pristine MS image, can be well fitted by a sum of Gaussians. This normalization process for each subimage of **I** can be described as:

$$\hat{I}_s(i,j) = \frac{I_s(i,j) - \mu_s(i,j)}{\sigma_s(i,j) + C} \tag{7}$$

where s $\in$ {red, green, blue, NIR, NDVI, NDWI or RatioG} denotes the subimage index, and

$$\mu_s(i,j) = \sum_{k=-K}^{K}\sum_{l=-L}^{L} \omega_{k,l} I_s(i+k, j+k) \tag{8}$$

$$\sigma_s(i,j) = \sqrt{\sum_{k=-K}^{K}\sum_{l=-L}^{L} \omega_{k,l} \left(I_s(i+k,j+k) - \mu_s(i,j)\right)^2} \tag{9}$$

where $\omega = \{\omega_{k,l} | k = -K, ...K, l = -L, ...L\}$ is a 2D circularly-symmetric Gaussian filter ($K = L = 3$). Eq (7) represents mean subtracted and contrast normalized (MSCN) coefficients.

Refer to [73], we use a sum of Gaussian to accurately model the distributions of $\hat{I}_s(i,j)$. The sum of Gaussian model is given by:

$$g(x; a_1, a_2, \mu_1, \mu_2, \sigma_1, \sigma_2) = a_1 \exp\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + a_2 \exp\left(\frac{x-\mu_2}{\sigma_2}\right)^2 \tag{10}$$

where parameters a_1, a_2, μ_1, μ_2, σ_1 and σ_2 are used as quality-aware features, which can be estimated by least squares. In Fig. 7, we show the distributions of the original and pansharpened images as presented in Fig. 6. The details about the features are given in Table 3.

C. PRISTINE MVG MODEL LEARNING

To learn a pristine MVG model serves as a "reference" in evaluating the quality of an arbitrary image/block, we collected a set of high-quality MS images from Quickbird dataset (Qingshan Lake area of Nanchang) and Worldview-3 dataset (Washington DC area). The Worldview-3 sensor offers 8-band multispectral components and the PAN band. The resolution cell is 1.24m×1.24m for MS bands and 0.31m×0.31m for PAN band. In the experiment, we select 100 images from the objects of vegetation, buildings, roads and water as pristine images, shown in Fig. 8. None of the pristine images is included in our database.

Using 100 pristine MS images, we extracted the above described statistical features from the MSCN of NIR, Red, Green, Blue, NDVI, NDWI and RatioG, and learn a pristine MVG model from the statistical features. The key steps detailing the process are described:

Step 1 (Patch Extraction): The pristine MS images are separated into un-overlapping blocks with size of 96×96. To make the extracted features more efficient for quality prediction, we only adapted a subset of the blocks based on patch's contrast. Based on the contrast calculated in Eq. (9), only those patches having a high contrast greater than a given threshold are selected, in which the threshold is set as 75% of the peak patch contrast of each image.

Step 2 (Feature Extraction): Using the feature representation method in above, a m-dim feature vector is extracted for each block ($m = 84$). Let $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n] \in^{m\times n}$ be the matrix of feature vectors extracted from n selected image patches, the mean vector and covariance matrix are expressed as:

$$\boldsymbol{\mu} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{X}_i(j), \quad i = 1, \ldots, m \tag{11}$$

$$\boldsymbol{\Sigma} = (\sigma_{ij})_{m\times m} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_{mm} \end{bmatrix} \tag{12}$$

where σ_{ij} is the covariance of column vector $\mathbf{x}_i$ and $\mathbf{x}_j$.

TABLE 3. Description of features.

Feature number	Feature group	Feature description	Feature compute process
f_1 –f_{48}	Spectral band features	The parameters of a sum of Gaussians fitting the MCSN distributions the of blue, green, red and NIR in scale 1 and scale 2	A sum of Gaussians fitted by least square estimation
f_{49}-f_{84}	Ground objects features	The parameters of a sum of Gaussians fitting the MCSN distributions of NDVI, NDWI and RatioG in scale 1 and scale 2	A sum of Gaussians fitted by least square estimation

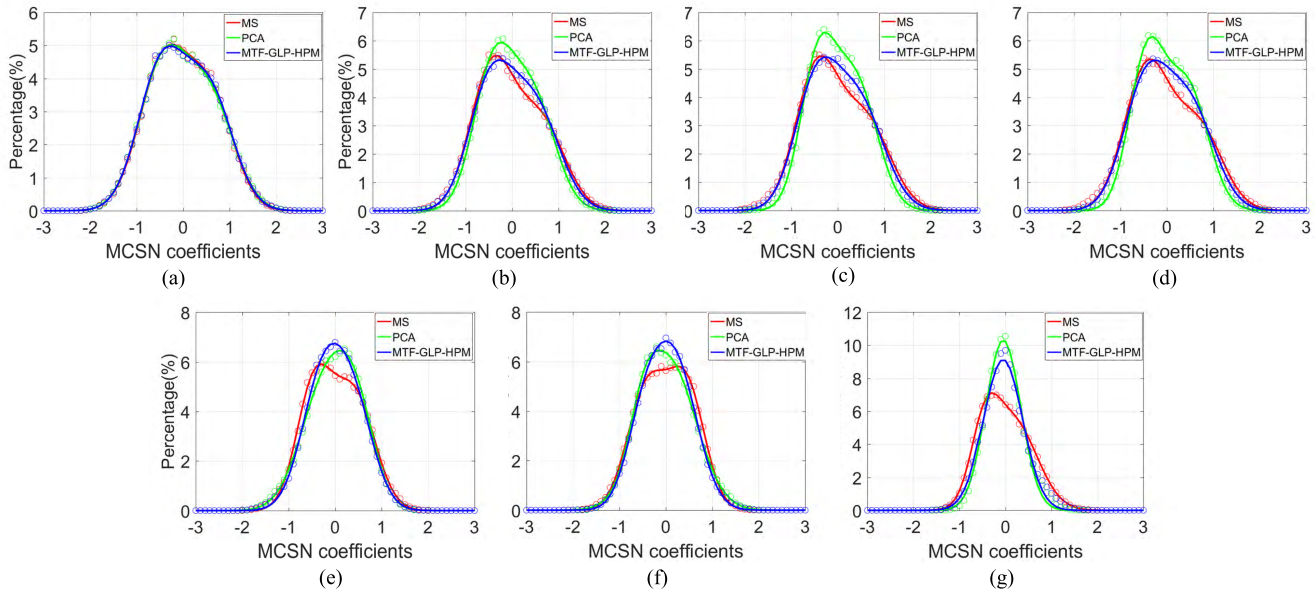


FIGURE 7. MSCN distributions and their corresponding fitting curves of the original and pansharpened images. (a) NIR. (b) Red. (c) Green. (d) Blue. (e) NDVI. (f) NDWI. (g) RatioG.

Step 3 (MVG Fitting): Once the mean vector ($\boldsymbol{\mu}$) and covariance matrix ($\boldsymbol{\Sigma}$) are obtained, the benchmark MVG model is fitted as:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) \tag{13}$$

D. TESTING

Once the benchmark MVG model ($\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$) is learned, it can be used to measure the quality of any patch for a given test image. For a test image, we first separate it into nonoverlapped patches with size of 96×96. The statistical feature is extracted for each patch, and the MVG model is learned using the same process described in the training stage. For a patch with MVG model ($\boldsymbol{\mu}_i$, $\boldsymbol{\Sigma}_i$) , the quality is obtained by calculating the distance between ($\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$) and ($\boldsymbol{\mu}_i$, $\boldsymbol{\Sigma}_i$) as:

$$\begin{aligned} q_i &= D(\boldsymbol{\mu}, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}, \boldsymbol{\Sigma}_i) \\ &= \sqrt{(\boldsymbol{\mu}-\boldsymbol{\mu}_i)^T\left(\frac{\boldsymbol{\Sigma}+\boldsymbol{\Sigma}_i}{2}\right)^{-1}(\boldsymbol{\mu}-\boldsymbol{\mu}_i)} \end{aligned} \tag{14}$$

It is evident that a small value q_i means high quality. The final quality score of the test image is calculated as the mean of all $\{q_i\}$.

V. EXPERIMENTAL RESULTS AND ANALYSES

A. EXPERIMENTAL DESCRIPTION

The evaluation experiments were conducted on the constructed database using three performance indicators: Pearson linear correlation coefficient (PLCC), Spearman rank order correlation coefficient (SRCC), and root mean square error (RMSE), between the objective scores after nonlinear regression and the subject scores. For the nonlinear regression, a five-parameter regression is used to map the objective scores to subjective scores [74]:

$$f(x) = \beta_1\left(\frac{1}{2} - \frac{1}{1+\exp(\beta_2(x-\beta_3))}\right) + \beta_4 x + \beta_5 \tag{15}$$

where $\beta_i (i = 1, 2, 3, 4, 5)$ are parameters determined by fitting.

The comparative methods selected in this paper include: (1) Five full reference IQA metrics: SSIM, PSNR, UIQI, SAM and ERGAS; (2) Three training-based BIQA metrics:

TABLE 4. Performance results with state-of-the-art FR and NR-IQA methods on our database (the top two results are in bold).

	PSNR	SSIM	UIQI	SAM	ERGAS	DIIVINE	BLIINDS-II	BIQI	NIQE	IL-NIQE	QNR	GQNR	Proposed
PLCC	0.4171	0.2651	0.3550	0.5158	0.2114	0.4296	0.6288	**0.7534**	0.322	0.5419	0.3454	0.6971	**0.7410**
SRCC	0.1748	-0.0657	0.0613	0.3564	-0.1037	0.5127	0.5727	**0.6744**	0.3102	0.3783	-0.1738	0.6109	**0.6913**
RMSE	12.7680	13.0054	13.0437	11.7141	13.0223	12.2233	10.2248	**8.8379**	12.6543	11.3191	12.6591	10.0299	**8.8399**

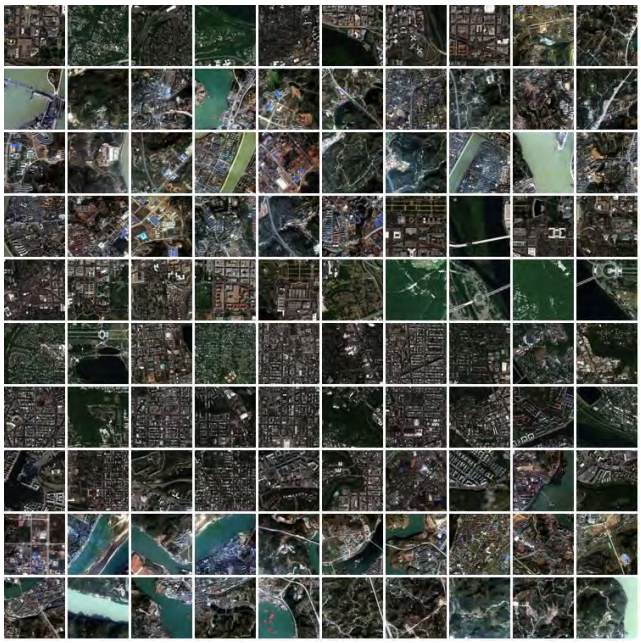


FIGURE 8. The selected 100 pristine MS images used to learn the benchmark MVG model.

BIQI, DIIVINE and BLIINDS-II; (3) Two opinion-unaware NR-IQA metrics: NIQE and IL-NIQE; (4) Two no-reference quality evaluation methods designed for pansharpened images: QNR and GQNR. Note that, for those full reference IQA metric, both the original MS and PAN images are degraded to a lower resolution so that the original MS image can be used as the reference, while for our method, we directly evaluate the quality of full resolution images without down-sampling.

B. EXPERIMENTAL RESULTS

To objectively evaluate the performance of our method, we compare our method with the selected eleven methods. Table 4 illustrates the comparisons of PLCC, SRCC and RMSE on our database. From the results, we can draw the following observations: 1) Compared with FR-IQA metrics, the evaluation performances of these metrics are extremely low, because spectral distortion is not considered in traditional PSNR, SSIM metrics and only spectral distortion is considered in ERGAS and SAM, leading to incomplete representation of the distortions. 2) Compared with three training-based BIQA metrics, our method is superior to DIIVINE and BLIINDS-II, but is superior to BIQA in some indictors. Since these metrics are highly dependent on the features and the training samples on which they trained, our method has stronger robustness than these training-based metrics. 3) Compared with two opinion-unaware NR-IQA metrics (NIQE and IL-NIQE) and two no-reference quality evaluation methods designed for pansharpened images (QNR and GQNR), our method has the best performance. The reason may be that those feature extraction methods used in these metrics are not suitable for our case. Overall, our method takes spectral distortion, spatial distortion and effects on information indexes into account, resulting in consistent distribution with the subjective observations.

In Fig. 9, we also provide the scatter plots of the predicted quality scores against the DMOS values for some representative objective metrics (such as PSNR, FSIM, ERGAS, SAM, QNR, GQNR, NIQE and IL-NIQE) on our database. From Fig. 9, it can be observed that our model has superior convergence and monotonicity, especially better than the PSNR, SSIM, ERGAS, NIQE and QNR metrics.

TABLE 5. Performance results of individual subimages.

Band	PLCC	SRCC	RMSE
Red	0.4510	0.3807	11.7489
Green	0.6625	0.6068	10.5167
Blue	0.7332	0.6823	8.9379
NIR	0.4524	0.4371	11.8358
NDVI	0.2052	0.1165	12.9960
NDWI	0.1693	0.0979	13.0110
RatioG	0.6661	0.5596	10.2306
Proposed	0.7410	0.6913	8.8399

C. PERFORMANCE EVALUATION OF EACH SUBIMAGE

Since we use seven bands (Blue, Green, Red, NIR, NDVI, NDWI and RatioG) to represent features, the influence of each band is further analyzed in this subsection. As shown in Table 5, for four red, green, blue and NIR bands in MS image, green and blue bands have the great influence on the performance, because spectral distortions are sensitive to in these bands. Although NDVI and NDWI have very low performance in the independent evaluation, they can capture special ground objects information and can compensate for the limitations of other bands. As a result, integrating these bands obviously promote the prediction performance compared with using independent ones. We further analyze the performance of different band combinations. As shown in Table 6, for simplicity, we design 6 schemes that only

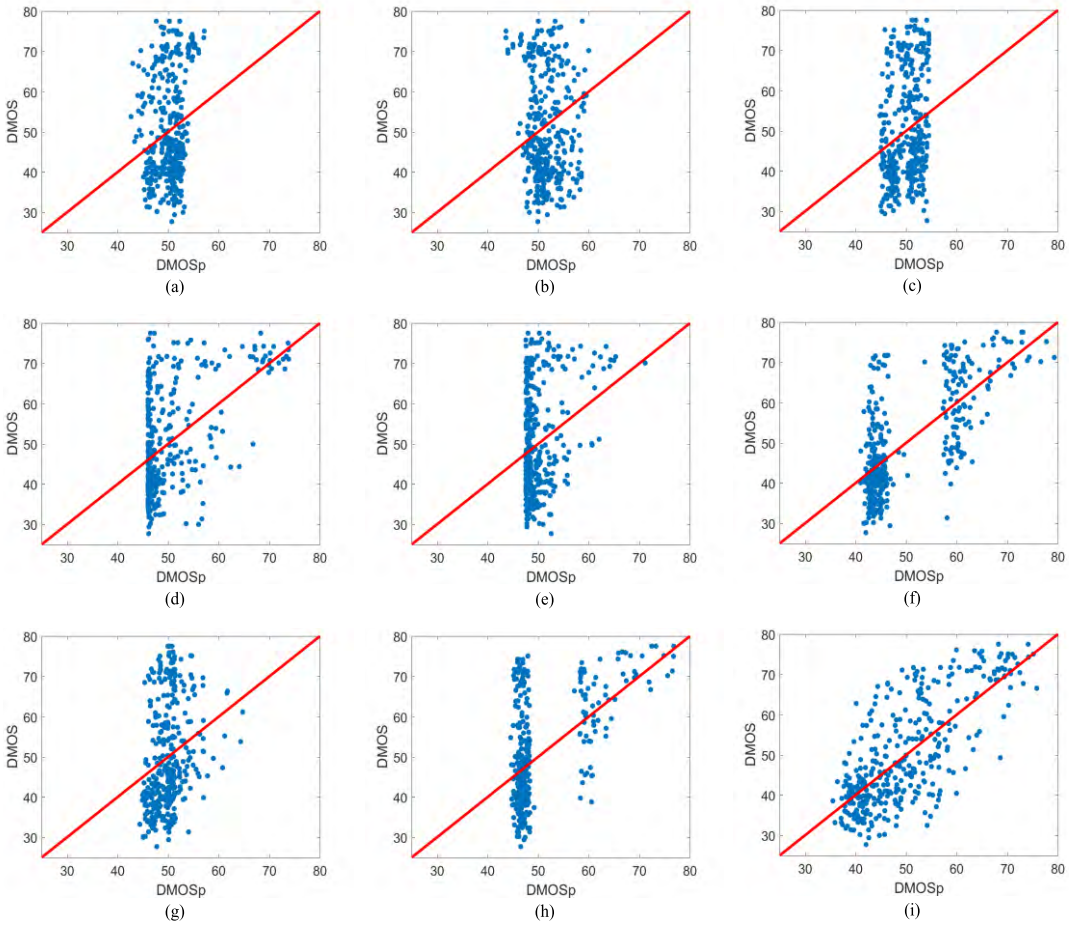


FIGURE 9. Scatter plots of DMOS versus DMOSp for various methods on our database. (a) PSNR. (b) SSIM. (c) ERGAS. (d) SAM. (e) QNR. (f) GQNR. (g) NIQE. (h) IL-NIQE. (i) PROPOSED.

TABLE 6. Performance results with different band combinations.

Band combination							PLCC	SRCC	RMSE
NIR	Red	Green	Blue	NDVI	NDWI	RatioG			
—	√	√	√	√	√	√	0.6525	0.6153	10.1973
√	—	√	√	√	√	√	0.6618	0.6183	10.1190
√	√	—	√	√	√	√	0.4990	0.4419	11.5507
√	√	√	—	√	√	√	0.4447	0.3879	11.9161
√	√	√	√	—	√	√	0.7184	0.6767	9.2292
√	√	√	√	√	—	√	0.6725	0.6066	9.9240
√	√	√	√	√	√	—	0.6703	0.5931	9.8226
√	√	√	√	√	√	√	0.7410	0.6913	8.8399

consider all possible combinations of 6 bands. Agreeing with the conclusion in the above analysis, each band has its unique contribution in characterizing the quality degradation.

D. PERFORMANCE VALIDATION

The most direct application of our metric is to guide the selection of optimal pansharpening methods. We provide an example to show the influence of different metrics on selecting the best pansharpening results. As shown in Fig. 10, for the pansharpened images obtained using six pansharpening methods: three CS methods (GSA, IHS and PCA) and three MRA methods (ATWT, ATWT-M2 and MTF-GLP), the image in (d) has significant spectral distortion, while the images in (c), (e) are spatially distorted. We use PSNR, SSIM, SAM, NIQE, IL-NIQE, QNR, GQNR and our method to measure the quality of the pansharpened images. Table 7 reports the objective evaluation results for the images shown in Fig. 10. It is obvious that our method can achieve the best consistent trend with the ground-truth DMOSs, while other metrics have more or fewer deviations in some pansharpened images. Overall, our method can well select the best pansharpened image in a completely blind way.

E. PERFORMANCE OF DIFFERENT PATCH SIZE

As shown in previous context, with the fixed patch size (96×96) our method has good performance. In this section, it is shown that the performance of our method is insensitive to changes in patch size. The patch size ranges from 48 to 132 with a step size of 12. The results are presented in Fig.11. SRCC and PLCC are used as the performance indicators.

TABLE 7. The predicted quality values of different metrics for the pansharpened images in fig. 10.

Methods	PSNR	SSIM	SAM	NIQE	IL-NIQE	QNR	GQNR	Proposed	DMOS
ATWT	16.2540 (4)	0.4980 (4)	8.6907 (2)	4.5944 (3)	33.8150 (3)	0.8867 (4)	4.8778 (1)	10.1594 (1)	33.1751 (1)
MTF-GLP	16.1629 (5)	0.4924 (5)	8.9340 (3)	4.6833 (4)	32.9608 (1)	0.8942 (6)	4.9414 (2)	10.7319 (2)	35.5699 (2)
ATWT-M2	17.3587 (1)	0.5452 (1)	8.2168 (1)	5.5395 (6)	37.0516 (4)	0.8745 (3)	6.3850 (5)	16.1772 (3)	48.8206 (3)
GSA	16.6533 (2)	0.5201 (3)	9.4321 (4)	4.3569 (1)	43.2178 (5)	0.8924 (5)	6.2851 (4)	17.8017 (4)	53.9231 (4)
IHS	14.4748 (6)	0.4376 (6)	12.5450 (6)	4.7408 (5)	33.3628 (2)	0.7838 (1)	5.9844 (3)	20.6070 (5)	71.7294 (5)
PCA	16.6498 (3)	0.5258 (2)	10.7513 (5)	4.4996 (2)	70.0378 (6)	0.8386 (2)	10.4482 (6)	23.3729 (6)	75.1548 (6)

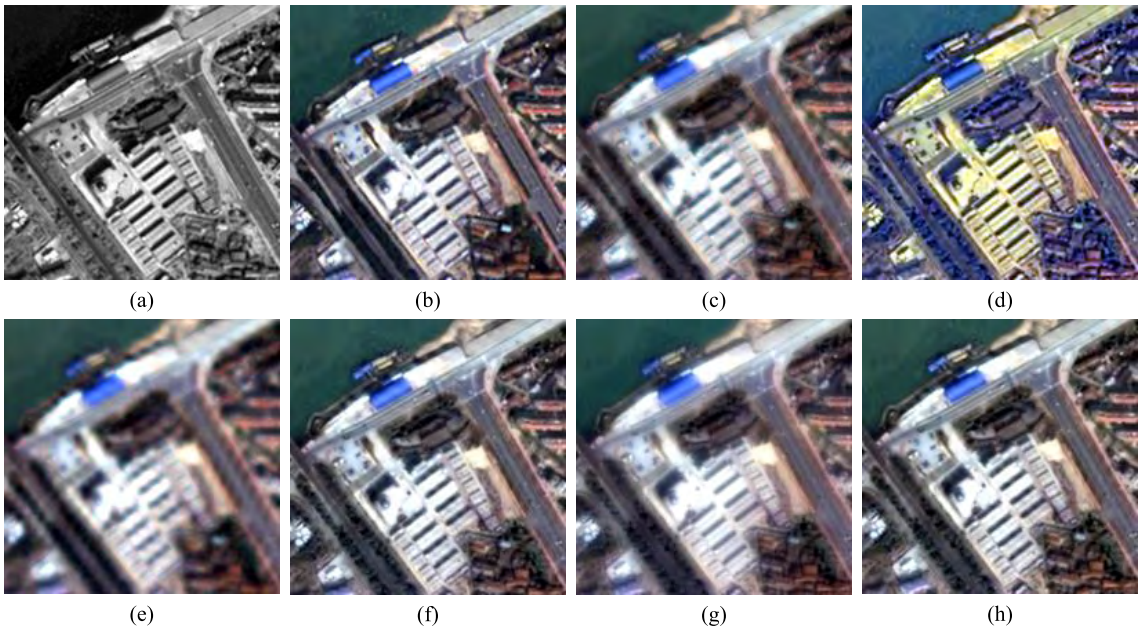


FIGURE 10. Example of original MS/PAN and its corresponding fused images generated by different pansharpening methods. (a) PAN. (b) MS. (c) GSA. (d) IHS. (e) PCA. (f) ATWT. (g) ATWT-M2. (h) MTF-GLP.

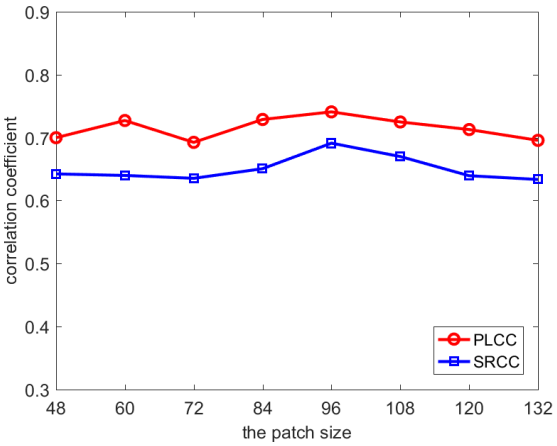


FIGURE 11. The performance (PLCC and SROCC) with different patch sizes.

From the results shown in Fig.11, it can be seen that the proposed method's performance is robust to the patch size variations in a considerable range.

F. DISCUSSION

Although the proposed method achieves good performance in evaluating the quality of pansharpened images, the following aspects deserve further attentions: 1) The current database only considers two types of pansharpening methods, which is inadequate for practical applications. Therefore, more types of pansharpening methods should be included in the further database construction. 2) We only select source MS images from Quickbird database, for other databases with more bands, such as WorldView-3, more effective feature representation method should be considered. 3) For the more critical subjective test, the absolute evaluation may be not a feasible way in estimating the quality of two images. Measuring relative quality (e.g., rank score) against other pansharpened images generated from the same source image may be more suitable for remote sensing applications. 4) Our work only concerns the quality assessment of multispectral pansharpened images. For other types of remote sensing image fusion, such as hyperspectral and multispectral image fusion, since hyperspectral images have more abundant bands information, it may also be applicable.

VI. CONCLUSIONS

In this paper, a subjective database of pansharpened images is constructed, and a no-reference quality evaluation method for pansharpened image is proposed. In order to measure the spectral and spatial distortions, our method combines the features for NIR, Red, Green, Blue, NDVI, NDWI and RatioG.

We built a benchmark MVG model from pristine MS images. Quality score for a test image is computed by measuring the distance between the benchmark MVG model and the fitted MVG model of the test image. Compared with other metrics on the database, our metric yields more consistent with subjective observation. In future work, we will focus on constructing a more large-scale database.

REFERENCES

[1] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.

[2] X. Meng, H. Shen, H. Li, L. Zhang, and R. Fu, "Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges," *Inf. Fusion*, vol. 46, pp. 102–113, Mar. 2019.

[3] P. S. Chavez, Jr., S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, 1991.

[4] G. Vivone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.

[5] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001.

[6] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6 011 875, Jan. 4, 2000.

[7] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images. II. Channel ratio and 'chromaticity' transformation techniques," *Remote Sens. Environ.*, vol. 22, no. 3, pp. 343–365, Aug. 1987.

[8] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS + pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.

[9] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pan sharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 228–236, Jan. 2008.

[10] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, Jan. 2011.

[11] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Nov. 2000.

[12] G. Vivone, R. Restaino, M. Dalla Mura, G. Licciardi, and J. Chanussot, "Contrast and error-based fusion schemes for multispectral image pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 930–934, May 2014.

[13] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 49–61, Jan. 2000.

[14] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, May 1999.

[15] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005.

[16] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.

[17] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and PAN imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[18] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "An MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas," in *Proc. 2nd GRSS/ISPRS Joint Workshop Remote Sens. Data Fusion URBAN Areas*, 2003, pp. 90–94.

[19] J. Lee and C. Lee, "Fast and efficient panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 155–163, Jan. 2010.

[20] M. M. Khan, J. Chanussot, L. Condat, and A. Montanvert, "Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 98–102, Jan. 2008.

[21] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rouge, "A variational model for P+XS image fusion," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 43–58, Aug. 2006.

[22] F. Fang, F. Li, C. Shen, and G. Zhang, "A variational approach for pan-sharpening," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2822–2834, Jul. 2013.

[23] D. Fasbender, J. Radoux, and P. Bogaert, "Bayesian data fusion for adaptable image pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1847–1857, Jun. 2008.

[24] P. Liu, L. Xiao, J. Zhang, and B. Naz, "Spatial-hessian-feature-guided variational model for pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2235–2253, Apr. 2016.

[25] N. D. A. Mascarenhas, G. J. F. Banon, and A. L. B. Candeias, "Multispectral image data fusion under a Bayesian approach," *Int. J. Remote Sens.*, vol. 17, no. 8, pp. 1457–1471, May 1996.

[26] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, "A new pansharpening algorithm based on total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 318–322, Jan. 2014.

[27] L. Zhang, H. Shen, W. Gong, and H. Zhang, "Adjustable model-based fusion method for multispectral and panchromatic images," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 42, no. 6, pp. 1693–1704, Dec. 2012.

[28] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.

[29] M. Guo, H. Zhang, J. Li, L. Zhang, and H. Shen, "An online coupled dictionary learning approach for remote sensing image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1284–1294, Apr. 2014.

[30] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "A practical compressed sensing-based pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 629–633, Jul. 2012.

[31] X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.

[32] H. Shen, X. Meng, and L. Zhang, "An integrated framework for the spatio–temporal–spectral fusion of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7135–7148, Dec. 2016.

[33] H. Shen, L. Zhang, B. Huang, and P. Li, "A MAP approach for joint motion estimation, segmentation, and super resolution," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 479–490, Feb. 2007.

[34] Q. Wei, N. Dobigeon, and J. Y. Tourneret, "Bayesian fusion of multi-band images," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1117–1127, Sep. 2015.

[35] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, May 2015.

[36] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, p. 594, Jul. 2016.

[37] Q. Yuan, Y. Wei, X. Meng, H. Shen, and L. Zhang, "A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 978–989, Mar. 2018.

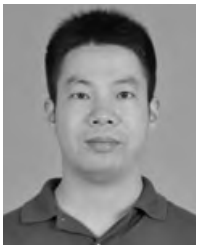
[38] R. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the Spectral Angle Mapper (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, vol. 1, 1992, pp. 147–149.

[39] L. Wald, *Data Fusion. Definitions and Architectures—Fusion of Images of Different Spatial Resolutions*. Paris, France: Les Presses de l' École des Mines, 2002.

[40] A. Garzelli and F. Nencini, "Hypercomplex quality assessment of multi/hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 662–665, Oct. 2009.

[41] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.
[42] C. Kwan, B. Budavari, A. C. Bovik, and G. Marchisio, "Blind quality assessment of fused WorldView-3 images by using the combinations of pansharpening and hypersharpening paradigms," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1835–1839, Oct. 2017.
[43] J. W. Rouse, Jr., R. H. Haas, J. A. Schell, D. W. Deering, and J. C. Harlan, "Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation," NASA/GSFC, Greenbelt, MD, USA, Final Rep. E74-10676, 1974, p. 371.
[44] B. C. Gao, "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, Dec. 1996.
[45] *eCognition Developer (8.64.0) User Guide*, eCognition, Trimble Germany GmbH, Munich, Germany, 2010.
[46] B. Salehi, Y. Zhang, M. Zhong, and V. Dey, "Object-based classification of urban areas using VHR imagery and height points ancillary data," *Remote Sens.*, vol. 4, no. 8, pp. 2256–2276, Aug. 2012.
[47] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
[48] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.
[49] X. Meng, H. Shen, Q. Yuan, H. Li, L. Zhang, and W. Sun, "Pansharpening for cloud-contaminated very high-resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, to be published.
[50] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, May 2010.
[51] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.
[52] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
[53] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
[54] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.
[55] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
[56] J. Zhou, D. L. Civco, and J. A. Silander, "A wavelet transform method to merge Landsat TM and SPOT panchromatic data," *Int. J. Remote Sens.*, vol. 19, no. 4, pp. 743–757, 1998.
[57] F. Palsson, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Quantitative quality evaluation of pansharpened imagery: Consistency versus synthesis," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1247–1259, Mar. 2016.
[58] M. M. Khan, L. Alparone, and J. Chanussot, "Pansharpening quality assessment using the modulation transfer functions of instruments," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3880–3891, Nov. 2009.
[59] M. Hasanlou and M. R. Saradjian, "Quality assessment of pan-sharpening methods in high-resolution satellite images using radiometric and geometric index," *Arabian J. Geosci.*, vol. 9, no. 1, p. 45, 2016.
[60] D. Rodríguez-Esparragón, J. Marcello, F. Eugenio, A. García-Pedrero, and C. Gonzalo-Martín, "Object-based quality evaluation procedure for fused remote sensing imagery," *Neurocomputing*, vol. 255, pp. 40–51, Sep. 2017.
[61] M. Selva, L. Santurri, and S. Baronti, "On the use of the expanded image in quality assessment of pansharpened images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 320–324, Mar. 2018.
[62] F. Bovolo, L. Bruzzone, L. Capobianco, A. Garzelli, S. Marchesi, and F. Nencini, "Analysis of the effects of pansharpening in change detection on VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 53–57, Jan. 2010.
[63] E. Ibarrola-Ulzurrun, C. Gonzalo-Martín, and J. Marcello, "Influence of pansharpening in obtaining accurate vegetation maps," *Can. J. Remote Sens.*, vol. 43, no. 6, pp. 528–544, 2017.
[64] J. K. Gilbertson, J. Kemp, and A. Van Niekerk, "Effect of pan-sharpening multi-temporal Landsat 8 imagery for crop type differentiation using different classification techniques," *Comput. Electron. Agricult.*, vol. 134, pp. 151–159, Mar. 2017.
[65] A. Garzelli, L. Capobianco, and F. Nencini, "On the effects of pansharpening to target detection," in *Proc. Geosci. Remote Sens. Symp.*, 2010, pp. II-136–II-139.
[66] P. Maglione, C. Parente, and A. Vallario, "Coastline extraction using high resolution WorldView-2 satellite imagery," *Eur. J. Remote Sens.*, vol. 47, no. 1, pp. 685–699, 2014.
[67] B. Johnson, "Effects of pansharpening on vegetation indices," *ISPRS Int. J. Geo-Inf.*, vol. 3, no. 2, pp. 507–522, 2014.
[68] Y. Xu, S. E. Smith, S. Grunwald, A. Abd-Elrahman, and S. P. Wani, "Effects of image pansharpening on soil total nitrogen prediction models in South India," *Geoderma*, vol. 320, pp. 52–66, Jun. 2018.
[69] Y. Du, Y. Zhang, F. Ling, Q. Wang, W. Li, and X. Li, "Water bodies' mapping from sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band," *Remote Sens.*, vol. 8, no. 4, p. 354, Apr. 2016.
[70] H. Li, L. Jing, and Y. Tang, "Assessment of pansharpening methods applied to WorldView-2 imagery fusion," *Sensors*, vol. 17, no. 1, p. 89, 2017.
[71] *Subjective Video Quality Assessment Methods for Multimedia Applications*, document Rec. ITU-T P.910, ITU Telecommunication Sector of ITU, 1999.
[72] *Subjective Video Quality Assessment Methods for Multimedia Applications*, document Rec. ITU-T P.911, ITU Telecommunication Sector of ITU, 1999.
[73] A. Goshtasby and W. D. Oneill, "Curve fitting by a sum of Gaussians," *Graph. Models Image Process.*, vol. 56, pp. 281–288, Jul. 1994.
[74] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.

BINGZHONG ZHOU received the B.S. and M.S. degrees from Hohai University, Changzhou, China, in 2002 and 2006, respectively. He is currently pursuing the Ph.D. degree with Ningbo University, Ningbo, China. He is currently an Engineer with the Department of Information Technology, Wenzhou Vocational and Technical College, China. His current research interests include image/video processing and quality assessment.



FENG SHAO (M'16) received the B.S. and Ph.D. degrees in electronic science and technology from Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively. He was a Visiting Fellow with the School of Computer Engineering, Nanyang Technological University, Singapore, from 2012 to 2012. He is currently a Professor with the Faculty of Information Science and Engineering, Ningbo University, China. He has published over 100 technical articles in refereed journals and proceedings in the areas of 3-D video coding, 3-D quality assessment, and image perception. He received the Excellent Young Scholar Award from the NSF of China (NSFC), in 2016.

XIANGCHAO MENG (M'18) received the B.S. degree in geographic information system from the Shandong University of Science and Technology, Qingdao, China, in 2012, and the Ph.D. degree in cartography and geography information system from Wuhan University, Wuhan, China, in 2017. He is currently a Lecturer with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China. His research interests include variational methods and remote sensing image fusion.

RANDI FU received the M.S. degree from PLA Information Engineering University, China, in 2001. He is currently an Associate Professor with the Faculty of Information Science and Engineering, Ningbo University, China. His research interests mainly include pattern recognition and remote sensing image processing.

YO-SUNG HO (SM'06–F'16) received the B.S. and M.S. degrees in electronic engineering from Seoul National University, Seoul, South Korea, in 1981 and 1983, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, in 1990. In 1983, he joined the Electronics and Telecommunications Research Institute (ETRI), Daejon, South Korea. From 1990 to 1993, he was with Philips Laboratories, Briarcliff Manor, NY, where he was involved in the development of the advanced digital high-definition television (AD-HDTV) system. In 1993, he re-joined the Technical Staff of ETRI, where he was involved in development of the Korean DBS digital television and high-definition television systems. Since 1995, he has been with the Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea, where he is currently a Professor with the Information and Communications Department. His research interests include digital image and video coding, image analysis and image restoration, advanced video coding techniques, digital video and audio broadcasting, 3-D video processing, and content-based signal representation and processing.