# Quality Assessment of Panchromatic and Multispectral Image Fusion for the ZY-3 Satellite: From an Information Extraction Perspective

Xin Huang, *Member, IEEE*, Dawei Wen, Junfeng Xie, and Liangpei Zhang

***Abstract*—The first results of multispectral (MS) and panchromatic (PAN) image fusion for the ZiYuan-3 (ZY-3) satellite, which is China's first civilian high-resolution satellite, are announced in this study. To this end, the various commonly used image fusion (pan-sharpening) techniques are tested. However, traditionally, image fusion quality is assessed by measuring the spectral distortion between the original and the fused MS images. The traditional methods focus on the spectral information at the data level but fail to indicate the image content at the information level, which is more important for specific remote sensing applications. In this context, we propose an information-based approach for assessing the fused image quality by the use of a set of primitive indices which can be calculated automatically without a requirement for training samples or machine learning. Experiments are conducted using ZY-3 PAN and MS images from Wuhan, central China. One of the objectives of the experiments is to investigate the appropriate image fusion strategies for the ZY-3 satellite at both the data and information levels. On the other hand, the experiments also aim to reveal the inadequacy of the traditional image quality indices and the advantages of the proposed information indices for describing image content. It is suggested that an appropriate image quality index should take into account the global and local image features at both the data and information levels.**

***Index Terms*—High spatial resolution, image fusion, image quality, information extraction, pan-sharpening, ZiYuan-3 (ZY-3).**

## I. Introduction

THE ZiYuan-3 (ZY-3) satellite, launched in January 2012, is China's first civilian high-resolution optical satellite. It aims to provide a routine geospatial service for mapping, change monitoring, and resource surveying over urban areas, agricultural land, forest, etc. The important parameters of the ZY-3 satellite are shown in Table I. It provides 2.1-m resolution panchromatic (PAN) band and four 5.8-m resolution multispectral (MS) bands. Consequently, it is an important task to fuse the PAN and MS images, thereby combining the spatial features in the PAN images and the spectral information in the MS images. The first objective of this study is to assess the performance of PAN–MS image fusion (or pan-sharpening) for the ZY-3 satellite. The experiments are interesting since this study is the first announcement of image fusion results for ZY-3 and, more importantly, the ratio of the spatial resolution between the PAN and MS images is not the usual case (1 : 4).

TABLE I
PARAMETERS OF THE ZY-3 HIGH-RESOLUTION SATELLITE
(GSD = ground spatial distance)

| | |
|---|---|
| Sensors and GSD | PAN: 2.1 m; MS: 5.8 m |
| Wavelength | PAN: 450 ~ 800 nm<br>Blue: 450 ~ 520 nm<br>Green: 520 ~ 590 nm<br>Red: 630 ~ 690 nm<br>Infrared: 770 ~ 890 nm |
| Revisit cycle | 5 days |
| Orbital inclination | 97.421° |
| Orbital type | Sun synchronous |

The different pan-sharpening techniques have been extensively studied in the remote sensing community. In this letter, some state-of-the-art methods are chosen for the ZY-3 satellite and are briefly summarized as follows.

1) *Gram-Schmidt (GS)* [1]: a component substitution strategy based on the GS transform.
2) *Multiplicative (MP)* [2]: the MS bands are multiplied by the PAN image in order to enhance the spatial information.
3) *Brovey* [2]: the normalized version of MP.
4) *High-pass filter* [3]: this combines the high-frequency components in the PAN image and the low-frequency components in the MS bands.
5) *Principal component analysis (PCA)* [4]: a component substitution strategy based on the PCA transform, i.e., the first principal component of the MS bands is replaced by the PAN image.
6) *Wavelet (WT)* [5]: a component substitution strategy based on the WT transform. Two widely used WT fusion algorithms, WT-PCA and WT-MAX, are employed. PCA and MAX indicate the fusion operators.
7) *Contourlet (CT)* [6]: a component substitution strategy based on the CT transform. Similarly, CT-PCA and CT-MAX are carried out. Compared to WT, CT has the potential to consider precise selection of the details in the PAN image, due to its anisotropic character.

It should be noted that the fusion strategies are not the core aim of this study and readers can find further details in the references.

Manuscript received April 15, 2013; revised May 27, 2013 and June 26, 2013; accepted August 2, 2013. Date of publication September 6, 2013; date of current version December 2, 2013. This work was supported in part by the Natural Science Foundation of China (41101336), by the Program for New Century Excellent Talents in University of China (NCET-11-0396), and by the Program for Changjiang Scholars and Innovative Research Team in University (IRT1278).

X. Huang, D. Wen, and L. Zhang are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: huang_whu@163.com).

J. Xie is with the Satellite Surveying and Mapping Application Center, National Administration Surveying, Mapping and Geoinformation, Beijing 100862, China.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2013.2278551

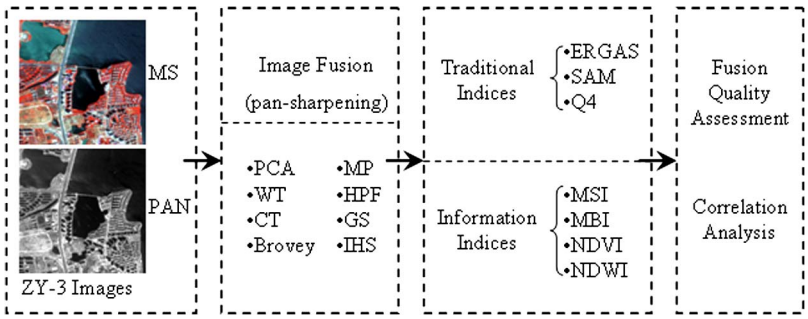


Fig. 1. Processing chain of this study for pan-sharpening of the ZY-3 satellite.

The different image fusion methods always alter the spectral and spatial components contained in the original PAN and MS images, which may significantly influence the accuracy of the information retrieval. Traditionally, the effectiveness of the fusion methods is assessed using image quality indices such as *Erreur Relative Globale Adimensionnelle de Synthese* (ERGAS) [7], the spectral angle [8], and Q4 [9]. These indices aim to measure the spectral distortion between the original and the pan-sharpened (or fused) MS bands, but they are not able to assess the quality of the fused images from the perspective of information content, which is more important for specific remote sensing applications. To this aim, in this study, we propose to use a set of information indices, which consist of the buildings, shadow, water, and vegetation indices, for evaluating the effectiveness of the image fusion methods. These primitive indices are calculated automatically without requiring machine learning or the collection of samples. Consequently, they are capable of directly reflecting the image content.

The contributions of this study lie in two aspects: 1) the first results of PAN–MS image fusion for the ZY-3 satellite are announced, and 2) a novel assessment strategy for fusion quality, based on a set of information indices, is proposed. The remaining parts of this letter are organized as follows. Section II introduces the traditional databased and the proposed information-based fusion quality indices. The results and analysis are presented in Section III. Finally, the conclusion is presented in Section IV.

## II. Methodology

This section describes the strategies for the quality assessment of pan-sharpened images, including the traditional indices at the data level and the proposed ones at the information level.

### A. Traditional Image Quality Indices

Three commonly used image quality indices are considered.

1) ERGAS [7]: This measures the root-mean-square error between the original and the fused image. Smaller ERGAS values correspond to a higher fusion quality.
2) Spectral angle mapper (SAM) [8]: This calculates the spectral similarity between the original and the fused MS images. Lower SAM values represent lower spectral distortion and better fusion quality.
3) Q4 index [9]: This is defined by measuring the correlation coefficient, spectral similarity, and contrast similarity between the original and pan-sharpened images. It should be noted that calculation of the Q4 index refers to a set of $N$ by $N$ blocks of an image and is finally obtained by averaging all the values over the whole image. Higher Q4 scores correspond to better spectral information preservation, since the Q4 index measures the similarity.

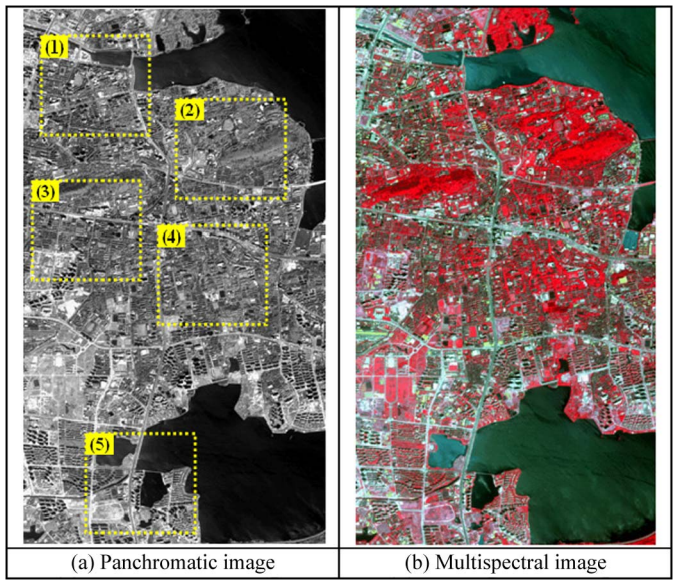


Fig. 2. (a) ZY-3 PAN and (b) MS images from Wuhan, central China. (a) PAN image. (b) MS image.

### B. Information-Level Image Quality Indices

The traditional image quality indices relate to the digital number of the spectral signals but do not evaluate the ability of information extraction, which is more meaningful for specific remote sensing applications. Consequently, in this study, we propose to assess the fused image quality by the use of a series of information indices, used to represent the image content. It should be noted that automatic information extraction from high-resolution imagery is a difficult problem. Based on our previous research, four primitive indices are employed in this letter: the morphological building/shadow indices and the normalized water/vegetation indices.

*MBI [10]:* The morphological building index (MBI) is able to automatically indicate the presence of buildings from high-resolution remotely sensed imagery [10]. It aims to represent the spectral and spatial characteristics of buildings using a series of morphological operators. The relatively high spectral reflectance of roofs with the spatially adjacent shadows leads to a high local contrast, which is described using the differential morphological profiles (DMPs) [10]. Moreover, buildings are

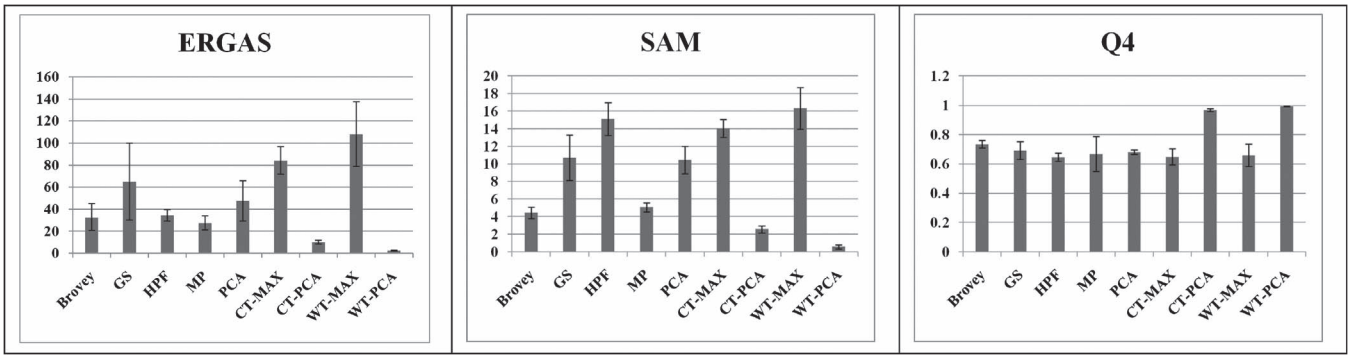


Fig. 3. Performance of the ZY-3 MS–PAN image fusion based on the traditional image quality indices.

relatively isotropic within a spatial range, which is depicted using a set of multidirectional and multiscale linear structural elements (SEs). Calculation of the MBI can be expressed by the following steps.

Step 1) Brightness image $b$. The maximum digital number of the visible bands for each pixel is used as its brightness value. The visible channels are of interest since they have the most significant contributions to the spectral property of buildings [10].

Step 2) Directional (white) top-hat by reconstruction. It is written as

$$\mathrm{WTH}(s,\mathrm{dir}) = b - \gamma_{\mathrm{RE}}(s,\mathrm{dir}) \tag{1}$$

where $\mathrm{WTH}(s,\mathrm{dir})$ represents the white top-hat transformation with a linear SE (length $= s$ and direction $=$ dir). $\gamma_{\mathrm{RE}}$ is the opening-by-reconstruction filter. The white top-hat transformation is used for the building index since it is capable of highlighting bright structures that have a size equal to or smaller than a given scale parameter (the length of a SE) and it suppresses other dark structures. Moreover, the linear SE is employed since it has the potential to discriminate between roads and buildings, considering that buildings are relatively compact and isotropic when compared to roads.

Step 3) DMPs. The DMPs of the white top-hat transformation ($\mathrm{DMP}_{\mathrm{WTH}}$) are used to model the building structures in a multiscale manner

$$\begin{cases} \mathrm{DMP}_{\mathrm{WTH}} = \{\mathrm{DMP}_{\mathrm{WTH}}(s,\mathrm{dir}) : s_{\min} \le s \le s_{\max}, \mathrm{dir} \in D\} \\ \mathrm{DMP}_{\mathrm{WTH}}(s,\mathrm{dir}) = |\mathrm{WTH}(s+\Delta s,\mathrm{dir}) - \mathrm{WTH}(s,\mathrm{dir})| \end{cases} \tag{2}$$

where $s_{\min}$ and $s_{\max}$ control the spatial size of the buildings and $D$ is the set of directions. In this way, the spectral–spatial characteristics of the buildings (e.g., brightness, local contrast, shape, and size) are embedded in the $\mathrm{DMP}_{\mathrm{WTH}}$ profiles.

Step 4) Calculation of the MBI. The MBI is defined as

$$\mathrm{MBI} = \frac{\sum_{s,\mathrm{dir}} \mathrm{DMP}_{\mathrm{WTH}}(s,\mathrm{dir})}{N_D \times N_S} \tag{3}$$

where $N_D$ and $N_S$ denote the directionality and the scale of the profiles, respectively. The definition of the MBI is based on the fact that building structures have high local contrast and, hence, have larger feature values in most of the directions of the profiles. Accordingly, buildings will have large MBI values.

TABLE II
DETECTION ACCURACIES (IN PERCENT) OF THE LAND-COVER CLASSES WITH THE DIFFERENT FUSION ALGORITHMS

| Methods | NDVI | NDWI | MSI | MBI |
|---|---|---|---|---|
| MS | 96.5 | 98.4 | 76.9 | 61.7 |
| Brovey | 96.5 | 98.4 | 70.2 | 80.1 |
| GS | 96.1 | 77.1 | 73.4 | 78.6 |
| HPF | 95.5 | 83.5 | 73.3 | 70.5 |
| MP | 96.5 | 98.4 | 82.2 | 69.9 |
| PCA | 95.7 | 86.0 | 74.5 | 74.9 |
| CT-MAX | 87.0 | 94.7 | 75.4 | 72.3 |
| CT-PCA | 96.3 | 98.1 | 77.4 | 66.3 |
| WT-MAX | 94.2 | 97.0 | 74.4 | 75.0 |
| WT-PCA | 96.3 | 98.2 | 76.7 | 62.8 |

*MSI [11]:* The morphological shadow index (MSI) is a related index to the MBI since shadows are spatially adjacent to buildings and they have similar structural features but different spectral properties. Accordingly, the shadow index is defined based on the DMP of the black top-hat transformation

$$\mathrm{MSI} = \frac{\sum_{s,\mathrm{dir}} \mathrm{DMP}_{\mathrm{BTH}}(s,\mathrm{dir})}{N_D \times N_S} \tag{4}$$

where $\mathrm{DMP}_{\mathrm{BTH}}$ is the black top-hat profiles. The MSI is able to highlight the dark structures that have high local contrast, which is in accordance with the spectral–spatial characteristics of shadow.

*NDVI:* Calculation of the normalized difference vegetation index (NDVI) is based on the principle that vegetation has a strong reflectance in the near-infrared (NIR) channel but a strong absorption in the red ($R$) channel. Consequently, it is defined as the normalized ratio of the NIR and R bands

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - R}{\mathrm{NIR} + R}. \tag{5}$$

*NDWI:* This is defined as

$$\mathrm{NDWI} = \frac{G - \mathrm{NIR}}{G + \mathrm{NIR}} \tag{6}$$

with $G$ being the green band. According to the spectral properties of a water body, the normalized difference water index (NDWI) aims to maximize the typical reflectance of the water

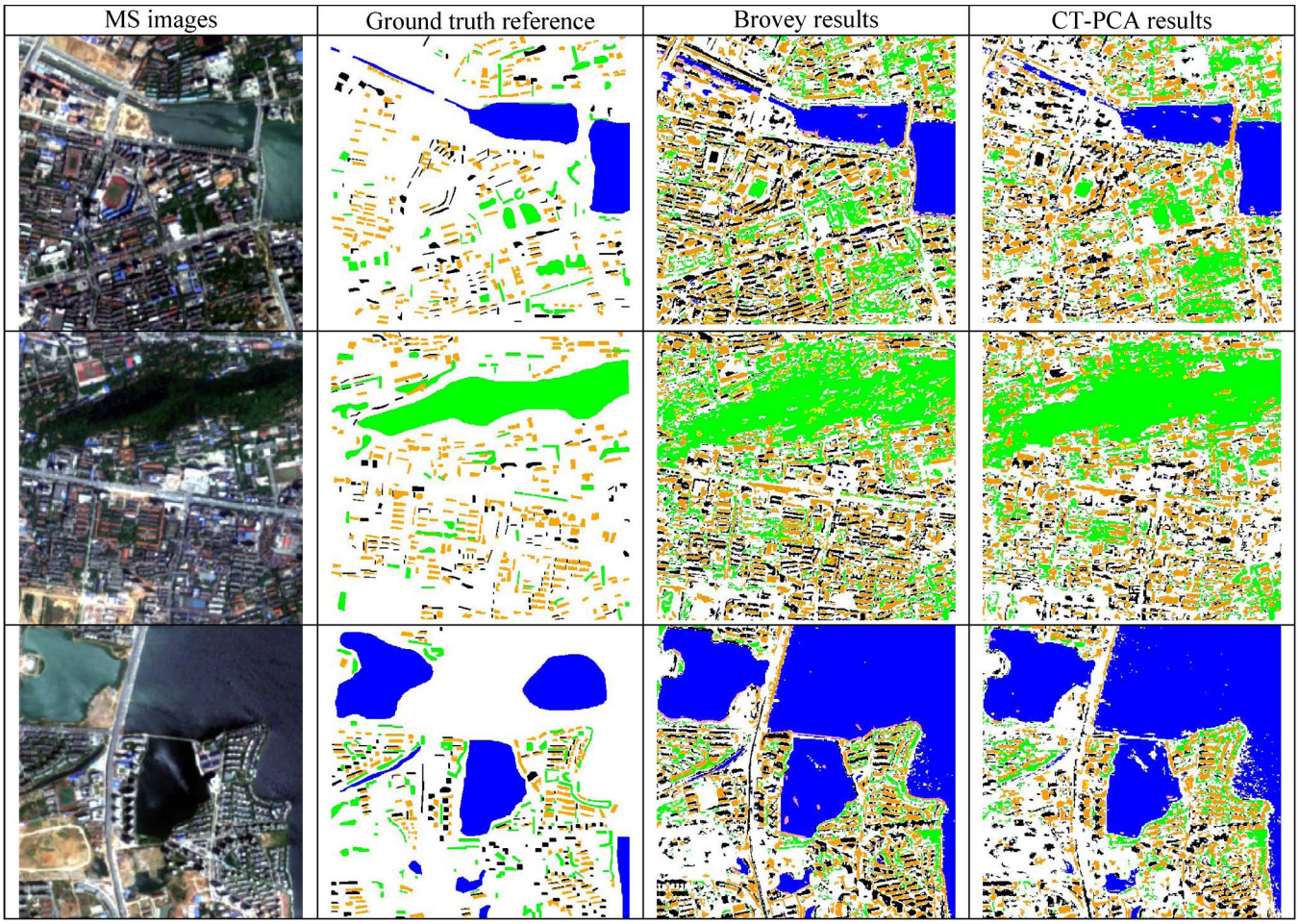


Fig. 4. Urban primitive information extraction based on the automatic information indices of MBI, MSI, NDVI, and NDWI. The first column shows the MS images of test regions #1, #2, and #3, respectively. The second column is their ground truth reference maps, which are manually delineated based on our prior knowledge and the field campaign in the study areas. Columns 3 and 4 are the automatic mapping results derived from the information indices (orange = buildings, black = shadow, green = vegetation, blue = water, and white = background).

features in the green band and, at the same time, minimizes their low reflectance in the NIR band.

## III. Results and Analysis

The pan-sharpening results of the ZY-3 satellite are presented and analyzed in this section. The processing chain is shown in Fig. 1. The PAN and MS images used in this study are shown in Fig. 2. The MS image is displayed in a color-infrared manner. The study area covers the city center of Wuhan, central China. The five rectangular regions in Fig. 2(a) are the areas where ground truth samples were collected for the accuracy assessment of the information indices.

### A. Data-Level Fusion Quality

In this experiment, the effectiveness of the PAN–MS image fusion methods is evaluated using a set of traditional image quality indices (i.e., ERGAS, SAM, and Q4). The experimental results are presented in Fig. 3, where all the indices are shown with the mean and standard deviation in the five validation regions. It can be seen that CT-PCA and WT-PCA are the most suitable for the ZY-3 MS-PAN fusion since they generate the smallest ERGAS and SAM values and the largest Q4 scores. Moreover, they give stable high-quality fusion results in terms of the small standard deviations.

### B. Information-Level Fusion Quality

In this experiment, the quality of the pan-sharpened ZY-3 images with different image fusion algorithms is assessed using a set of information indices. The accuracies of the four land-cover classes with the different image fusion methods are compared in Table II. It can be seen that, for the natural landscapes such as vegetation and water, the original MS bands achieve the highest accuracies. The Brovey and MP methods, which are able to preserve the original spectral values to the greatest degree, give the same high accuracies for the natural classes. Regarding the man-made structures, e.g., shadow and buildings, the accuracies of the original MS bands are not satisfactory. This is particularly true for buildings, where MS only gives a detection accuracy of 61.7%, which is significantly lower than the score of 80.1% achieved by the Brovey method. The highest accuracies for shadows and buildings are obtained by the MP and Brovey methods, respectively. It is therefore implied that, for man-made structures, the optimal fusion strategy varies according to the specific object of interest. The results of the information extraction are shown in Fig. 4. Although CT-PCA obtains better

TABLE III
$R^2$ PARAMETERS OF THE LINEAR REGRESSION FOR INDICATING THE LINEAR RELATIONSHIP BETWEEN THE TRADITIONAL IMAGE QUALITY INDICES AND THE INFORMATION EXTRACTION ACCURACIES

| Indices | ERGAS | SAM | Q4 |
|---|---|---|---|
| Vegetation (NDVI) | 0.006 | 0.027 | 0.035 |
| Water (NDWI) | 0.014 | 0.182 | 0.232 |
| Shadow (MSI) | 0.006 | 0.000 | 0.013 |
| Buildings (MBI) | 0.292 | 0.231 | 0.354 |

fusion quality than Brovey at the data level, it can be seen that, for the information level, the results generated by Brovey are more regular and homogeneous than CT-PCA, particularly for buildings.

### C. Correlation Between the Data and Information Levels

By comparing the results at the data and information levels, it can be found that a higher quality at the data level (i.e., a better preservation of spectral information between the original and the pan-sharpened images) does not signify higher accuracies for the information extraction. Consequently, a statistical analysis is performed between the traditional image quality indices and the proposed information quality indices. To this aim, linear regression, based on the accuracies of the four land-cover classes in the five test areas, is used to investigate their relationship. The $R^2$ parameters of the regression equations are used to describe the linear relationship (Table III). The $R^2$ values range from zero to one, and a small value means a weak linear relationship. The first comment with regard to Table III is that the values of $R^2$ are generally not high in all cases, showing that the traditional image quality indices are not effective for indicating the image information content. However, relatively speaking, Q4 is the optimal index to reflect the image content, since it obtains the highest $R^2$ values in all cases. This phenomenon can be attributed to the fact that Q4 is calculated based on a local window, where the neighboring information is considered as well as the spectral features.

## IV. CONCLUSION

The first results of MS and PAN image fusion for the ZY-3 satellite, which is China's first civilian high-resolution satellite, have been announced in this study. The commonly used pan-sharpening techniques are tested, and the image fusion quality is assessed using the traditional indices of ERGAS, SAM, and Q4. More importantly, we propose to use a set of information indices that are able to describe the image content for the fusion quality assessment. Four primitive information indices are employed in this study: the morphological building/shadow indices and the normalized vegetation/water indices. These indices can be automatically calculated without the need for machine learning or training samples and, hence, can directly reflect the image content.

The experimental results reveal several interesting and meaningful findings that have not been reported in the previous literature, to the best of our knowledge.

— *Observation 1:* The traditional image quality indices show that CT-PCA and WT-PCA are suitable for PAN and MS image fusion of the ZY-3 satellite at the data level, since they achieve the optimal spectral information preservation.
— *Observation 2:* The proposed information indices show that the Brovey and MP methods are appropriate for the ZY-3 satellite at the information level, since they give the highest detection accuracies for most of the land-cover classes.
— *Observation 3:* For the man-made urban structures, it is found that the optimal pan-sharpening method is related to the specific object of interest.
— *Observation 4:* The regression analysis reveals that the traditional image quality indices are not able to reflect the image information content effectively. Among them, the Q4 score gives the best response to the image content because it is calculated based on a moving window. This phenomenon provides guidance for the image quality assessment in that an appropriate image quality index should take into account the global and local image features at both the data and information levels.

## REFERENCES

[1] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," Eastman Kodak Company, Rochester, NY, USA, Tech. Rep. U.S. Patent 6 011 875, Jan. 2000.
[2] C. Pohl and J. L. Van Genderen, "Multisensor image fusion in remote sensing: Concepts, method and applications," *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, Nov. 1998.
[3] P. S. Chavez, 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, Mar. 1992.
[4] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 5, pp. 561–567, May 1992.
[5] G. Hong and Y. Zhang, "Comparison and improvement of wavelet-based image fusion," *Int. J. Remote Sens.*, vol. 29, no. 3, pp. 673–691, Feb. 2008.
[6] M. Lillo-Saavedra and C. Gonzalo, "Multispectral images fusion by a joint multidirectional and multiresolution representation," *Int. J. Remote Sens.*, vol. 28, no. 18, pp. 4065–4079, Aug. 2007.
[7] L. Wald, "Some terms of reference in data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1190–1193, May 1999.
[8] A. F. H. Goetz, R. H. Yuhas, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the Spectral Angel Mapper (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, pp. 147–149.
[9] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
[10] X. Huang and L. Zhang, "A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery," *Photogramm. Eng. Remote Sens.*, vol. 77, no. 7, pp. 721–732, Jul. 2011.
[11] X. Huang and L. Zhang, "Morphological building/shadow index for building extraction from high-resolution imagery over urban areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 161–172, Feb. 2012.