


## Article

# Rectangular-Normalized Superpixel Entropy Index for Image Quality Assessment

矩形归一化超像素熵指数用于图像质量评估

Tao Lu <sup>1</sup> , Jiaming Wang <sup>1,\*</sup>, Huabing Zhou <sup>1</sup>, Junjun Jiang <sup>2</sup>, Jiayi Ma <sup>3,4</sup> and Zhongyuan Wang <sup>5</sup>

<sup>1</sup> Hubei Key Laboratory of Intelligent Robot, School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan 430073, China; lut@wit.edu.cn (T.L.); zhou@wit.edu.cn (H.Z.)

<sup>2</sup> School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China; jiangjunjun@hit.edu.cn

<sup>3</sup> Electronic Information School, Wuhan University, Wuhan 430072, China; jiaiyima@whu.edu.cn

<sup>4</sup> Beijing Advanced Innovation Center for Intelligent Robots and Systems, Beijing Institute of Technology, Beijing 10081, China

<sup>5</sup> School of Computer Science, Wuhan University, Wuhan 430072, China; wzy\_hope@whu.edu.cn

\* Correspondence: wjmecho@163.com or wjmecho@wit.edu.cn; Tel.: +86-027-8799207

Received: 21 November 2018; Accepted: 8 December 2018; Published: 10 December 2018



**Abstract:** Image quality assessment (IQA) is a fundamental problem in image processing that aims to measure the objective quality of a distorted image. Traditional full-reference (FR) IQA methods use fixed-size sliding windows to obtain structure information but ignore the variable spatial configuration information. In order to better measure the multi-scale objects, we propose a novel IQA method, named RSEI, based on the perspective of the variable receptive field and information entropy. First, we find that consistence relationship exists between the information fidelity and human visual of individuals. Thus, we reproduce the human visual system (HVS) to semantically divide the image into multiple patches via rectangular-normalized superpixel segmentation. Then the weights of each image patches are adaptively calculated via their information volume. We verify the effectiveness of RSEI by applying it to data from the TID2008 database and denoise algorithms. Experiments show that RSEI outperforms some state-of-the-art IQA algorithms, including visual information fidelity (VIF) and weighted average deep image quality measure (WaDIQaM).

**Keywords:** image quality assessment; mutual information; superpixel segmentation

## 1. Introduction

With the rapid development of digital communication, images are playing increasingly important role in modern society. However, the quality of images is naturally degraded due to image acquisition, compression, storage, and transmission. Image quality assessment (IQA) is a basic problem in the field of image processing research. Image quality determination using only human-in-the-loop-based qualitative measures is time consuming, labor intensive, and cannot be applied to real-time or autonomous systems. Generally, objective IQA metrics can be divided into full-reference (FR), no-reference (NR), and reduced-reference (RR) methods [1–3].

A human visual system (HVS) is sensitive to visual quantifiable features such as brightness [4], contrast [5], inter-patch and intra-patch similarities [6], visual saliency [7], fuzzy gradient similarity deviation [8], and frequency content of an image [9]. FR-based IQA metrics are divided into two classes, namely, error statistic-based and HVS-based classes. Error statistic-based methods measure the distance between a distorted image and a reference image at the pixel level, and are less consistent with the HVS. Peak signal-to-noise ratio (PSNR) [10] and mean-squared error (MSE) are the widely used error statistic-based methods. HVS-based methods use visual quantifiable features to construct a visual

model. These factors are important in simulating the human perception of image distortion [11–13]. A noise quality measure (NQM) [11] uses these factors to find an image quality measure. Wavelet-based visual SNR (VSNR) [12] compares the low-level HVS property of the perceived contrast and the mid-level HVS property of global precedence.

The structural similarity (SSIM) index [14] suggests that the human eye is more sensitive to structural information based on field-of-view, and quantifies the degree of distortion by comparing brightness, contrast, and mechanism similarities. Wang et al. [15] combined the multi-scale SSIM (MS-SSIM) of wavelet domain and obtained improved performance. From the perspective of information theory, Sheikh et al. [16] proposed an information fidelity criterion (IFC) to quantify the mutual information (MI) of reference and distorted images. In [13], IFC was expanded to contain visual information fidelity (VIF). The feature similarity (FSIM) index [17] determines the visual difference of images in the feature domain by comparing the gradient and phase consistency. Li et al. [18] demonstrated the effectiveness of regional MI for IQA. Existing studies [17,19] have shown that VIF and FSIM are more consistent with the subjective results, compared to other traditional algorithms. Recently, Bosse et al. [20] proposed a deep neural networks (the network is based on HVS model) for image quality assessment (WaDIQaM), and achieved state-of-the-art performance.

Distortions, such as noise and blur, are inevitable in non-ideal image degradation and transmission [21,22]. In fact, the scenarios of different IQA applications are also different. Although the abovementioned methods are general, their performances are degraded when the images undergo specific degradations [23–26]. Our experimental results also prove this point (more details in Section 3.2).

Generally, IQA metrics frequently use fixed-size sliding windows to simulate HVS, such as receptive field [17]. However, they ignore the irregular and inhomogeneous the content and distribution of images, especially the variable spatial configuration information in satellite image. Image segmentation can divide the image into image patches with similar semantics. Existing traditional image segmentation algorithms are mainly divided into three categories, namely, turbopixel/superpixel [27,28] segmentation, watershed segmentation [29,30], active contour [31,32] algorithms. Recent studies conducted in [33,34] show that superpixel [28] provides an state-of-the-art representation of image data.

This study proposes a novel RSEI for IQA. Image patches provided by sliding window ignore the spatial structure information in the image and the correlation between the pixels, and can only measure the quality of the image from the low semantic information level. RSEI utilizes the superpixel segment [28] of the reference image and then clusters the content of the image. The superpixel is used to fully exploit the spatial information to obtain high-level semantic image patches. The distorted image is segmented based on clustering information. Therefore, the weights are automatically generated based on the IE of the reference image patch. RSEI uses MI to describe the changes between image patches.

Overall, the contributions of this paper are highlighted as follows:

1. The proposed IQA metric semantically divides the image into multiple flexible patches based on superpixel to accurately measure multi-scale objects in images. Here, the superpixel of images provides the variable spatial configuration information.
2. A weighting scheme that determines the importance of an image patch based on its information volume is introduced. This weighting scheme reveals the attention-seeking mechanism of HVS.
3. The proposed IQA metric focuses on the inevitable problems of image degradation and compression.

The remainder of this paper is structured as follows. In Section 2, we describe the framework of the proposed method. Experiment signals is analyzed in Sections 3. In addition, comparison is performed among the proposed metric and some representative IQA methods to show the superiority of the proposed method. Discussion and conclusion are summarized in Sections 4 and 5, respectively.

## 2. Rectangular-Normalized Superpixel Entropy Index

MI measures the degree of image distortion by quantifying the information dependence between reference image  $Y$  and distorted image  $\hat{Y}$  [35,36]. The joint entropy  $H(\hat{Y}, Y)$  of images  $\hat{Y}$  and  $Y$  is defined as follow:

$$H(\hat{Y}, Y) = \sum_{s \in \hat{Y}} \sum_{t \in Y} p(s, t) \log p(s, t). \quad (1)$$

where  $s$  and  $t$  represent the gray-scale value of the image, and  $p(s, t)$  is the joint probability of  $s$  and  $t$ . Thus, MI is defined as follow:

$$\begin{aligned} I(\hat{Y}, Y) &= H(\hat{Y}) + H(Y) - H(\hat{Y}, Y) \\ &= \sum_{s \in \hat{Y}} \sum_{t \in Y} p(s, t) \log \frac{p(s, t)}{p(s)p(t)}, \end{aligned} \quad (2)$$

where  $H(\hat{Y})$  and  $H(Y)$  are the entropies,  $p(s)$  represents the probability distribution.

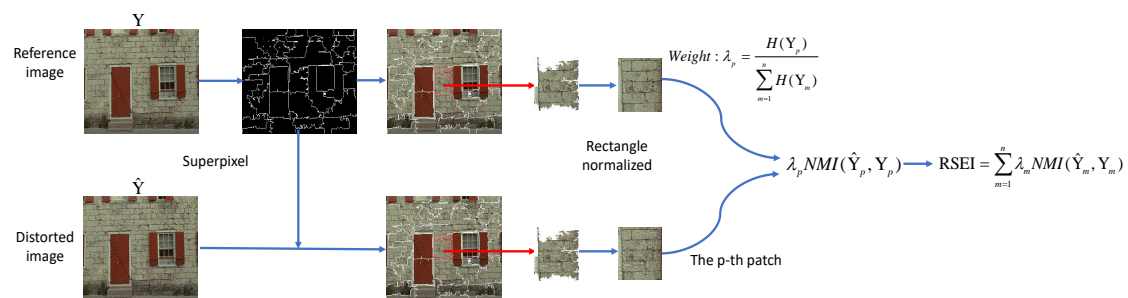
The greater the MI between the images, the greater the similarity information between images will be. However, MI ignores the visual perception characteristics of HVS, such as image patch weighting and contrast sensitivity. For an intuitive comparison, MI is normalized as follows [37]:

$$NMI(\hat{Y}, Y) = \frac{2I(\hat{Y}, Y)}{H(\hat{Y}) + H(Y)}. \quad (3)$$

IE reflects the amount of information in an image. However, the interference caused by distorted images increases the amount of information in that image, and this additional information causes a negative impact. Therefore, MI is not robust to the interference in the image, resulting in an inaccurate evaluation of image quality of the distorted image.

Various semantic image patches are important to the overall image. Different semantic objectives in the image have different levels of importance to the image. For objects with small variation, the amount of information, such as the sky and sea, and distortion have a small effect on the subjective quality of the image. For larger patches of information, such as airplanes and ships, each image section contains considerable gradients and structural information [38].

RSEI is illustrated in Figure 1. Reference image  $Y$  is divided into  $n$  image patches by semantic segmentation [28]. The corresponding segmentation label is recorded as  $l_Y$ . The resulting distorted image  $\hat{Y}$  is divided by label  $l_Y$ .



**Figure 1.** Illustration of the RSEI. Superpixel segmentation provides clustering information of spatial pixels which contains flexible image semantic information.

The shape of the image patch after semantic segmentation is irregular. In order to be able to do calculations, it is usually filled as a rectangle. Padding areas of 0 or 255 [39] are treated as exactly the same area, whether pixel-based or HVS-based, which will add additional error terms. Therefore, we use the minimum area of bounding rectangle-normalized to normalize the image patch by self-padding

to avoid excessive filling that affects segmentation. The convex hull of the image patch is recorded as  $\{hx_i, hy_i\}_{i=1}^t$ , where  $t$  is the number of convex hull points. Minimum area  $S$  is the rectangle defined as:

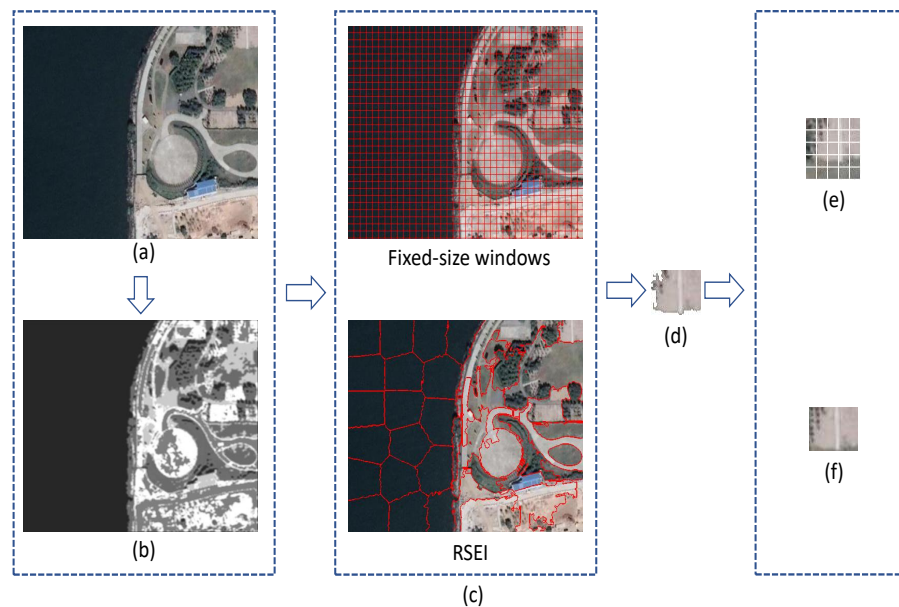
本文的精华部分

$$\theta = \text{unique} \left( \text{mod} \left( \arctan \left( \frac{hy_{i+1} - hy_i}{hx_{i+1} - hx_i} \right), \frac{\pi}{2} \right) \right), \quad (4)$$

$$S = \arg \min_{\theta} \left\{ \left( \max \left( [hx, hy] * \begin{bmatrix} \sin(-\theta) \\ \cos(-\theta) \end{bmatrix} \right) - \min \left( [hx, hy] * \begin{bmatrix} \sin(-\theta) \\ \cos(-\theta) \end{bmatrix} \right) \right) * \left( \max \left( [hx, hy] * \begin{bmatrix} \cos(-\theta) \\ -\sin(-\theta) \end{bmatrix} \right) - \min \left( [hx, hy] * \begin{bmatrix} \cos(-\theta) \\ -\sin(-\theta) \end{bmatrix} \right) \right) \right\}, \quad (5)$$

where *unique* is used to obtain the non repetitive element, *mod* is the modulo function.  $\theta$  is the angle set of bounding rectangle that is moved into the first quadrant.

Figure 2 shows the differences between the fixed-size sliding window and the proposed method. Image saliency [40] is an important visual feature in an image, that emphasizes the degree of importance of a region for human eye perception. The segmentation results of RSEI conform to the saliency map of the image. RSEI covers an irregular semantic patch 3(d) with the smallest rectangle, and the sliding windows ignore the semantics of the image patch and cover many unrelated areas.



**Figure 2.** Differences between RSEI and the fixed-size sliding window. (a) Reference image. (b) Saliency map of the reference image. (c) Segmentation results. (d) Image patch in the upper-right corner of semantic segmentation map. (e) Completely overlay the image patch (d) by fixed-size windows. (f) Completely overlay the image patch (d) by RSEI.

Then, the weight of the image patch is determined based on its IE. The larger the amount of information in the image is, the more important each object information will be to the overall assessment of the image. Patches with a small amount of information should occupy a small proportion. The weight of the  $p$ -th patch is defined as follows:

$$\lambda_p = \frac{H(Y_p)}{\sum_{m=1}^n H(Y_m)}, \quad (6)$$

where  $n$  represents the total number of segmented image patches and  $Y_p$  denotes the  $p$ -th image patch. RSEI is defined as follows:

$$RSEI = \sum_{m=1}^n \lambda_m NMI(\hat{Y}_m, Y_m) \quad (7)$$

Our proposed RSEI is described in Algorithm 1. The code has been made publicly available at <https://github.com/jiaming-wang/RSEI>.

---

**Algorithm 1** RSEI index for IQA
 

---

**Input:** Initialize the following parameters.

- (1) Reference images  $Y$ ;
- (2) distorted images  $\hat{Y}$ ;
- (3) number of image patches  $n$ ;

**Superpixel:**

- (4) Superpixel segmentation is performed on the reference image  $Y$ . The corresponding segmentation label

is recorded as  $l_Y$ ;

- (5) Normalizing reference images  $Y$  by label  $l_Y$ ;

- (6) Minimum area bounding rectangular-normalized by using Formula (5). The divided image patches are

normalized to  $\{Y_i\}_{i=1}^n$ , the distorted images  $\hat{Y}$  are normalized to  $\{\hat{Y}_i\}_{i=1}^n$ .

**RSEI:**

**for each patch**  $i = 1$  to  $n$

- (7) Compute weighted  $\lambda_i$  by Formula (6);

- (8) Compute MI  $NMI(\hat{Y}_i, Y_i)$  by Formula (3);

**end**

- (9) Compute RSEI by Formula (7).
- 

### 3. Experiments

#### 3.1. Databases

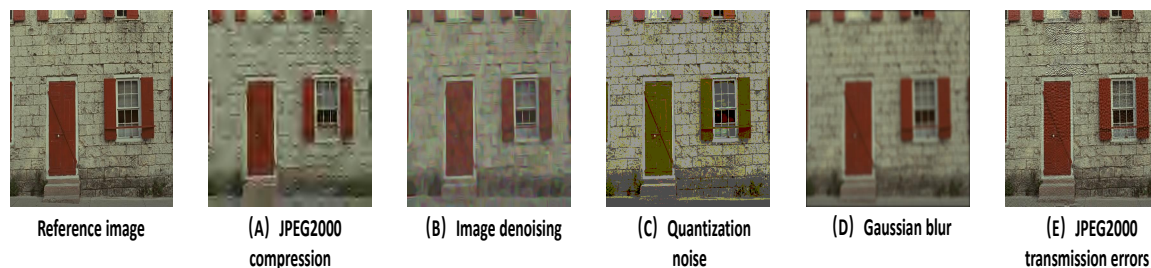
The TID2008 dataset [41] is a commonly used public database in the IQA community. The dataset contains 25 reference images and 1,700 distorted images. Each reference image corresponds to 68 different distorted images and includes 17 types of distortion. The MOS of the images is scored by 838 observers. The image size is  $512 \times 384$  pixels. All images are RGB images. However, all IQA algorithms are used in single-channel images. We convert the image pixels to YCbCr color space by using only the Y channel for testing.

Four common performance metrics are used to evaluate the performance of assessment methods. The Kendall rank-order correlation coefficient (KROCC) and Spearman rank-order correlation coefficient (SROCC) can be effectively used to measure the prediction monotonicity of an IQA metric. The two other metrics are the Pearson linear correlation coefficient (PLCC) and root MSE (RMSE) between MOS and objective scores after nonlinear regression. An excellent method indicates high KROCC, SROCC, and PLCC while low RMSE score [42].



### 3.2. Limitations of Existing IQA Algorithms

Figure 3, shows five types of distortion, namely, JPEG2000 compression, image denoising, quantization noise, Gaussian blur, and JPEG2000 transmission errors, all of which are selected from TID2008 [41]. As shown in Table 1, the last column is the mean opinion scores (MOS) of the images, and the first five columns are the results of PSNR, SSIM, VIF, FSIM, and rectangle-normalized superpixel entropy index (RSEI), respectively. Neither the state-of-the-art deep learning-based algorithm WaDIQaM nor the best performing VIF and FSIM can accurately describe these distortion changes.



**Figure 3.** (A–E) are the distorted versions of a reference image in TID2008 database. The distortion types of (A–E) are JPEG2000 compression, image denoising, quantization noise, Gaussian blur, and JPEG2000 transmission errors, respectively.

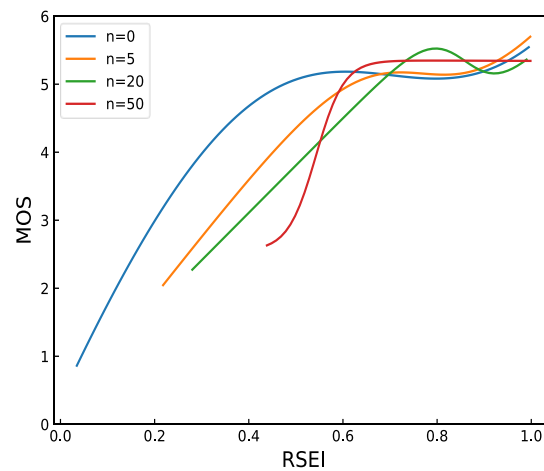
**Table 1.** Subjective scores with different type of distortion. From this table, traditional image quality measures such as PSNR, SSIM, VIF and FSIM are not always complied with distribution of MOS scores.

Type of Distortion	PSNR	SSIM	VIF	FSIM	WaDIQaM	RSEI	MOS
JPEG2000 compression	21.0628	0.3939	0.0858	0.7472	1.9727	0.2524	1.0000
Image denoising	21.0453	0.4397	0.1305	0.7688	2.2150	0.2605	1.2000
Quantization noise	21.1297	0.7806	0.3330	0.8907	2.4559	0.2592	2.1667
Gaussian blur	21.0833	0.3859	0.1125	0.7422	2.4214	0.2640	2.1765
JPEG2000 transmission errors	20.7817	0.7106	0.3617	0.9112	4.1493	0.3100	3.1765

### 3.3. Parameter Settings

The number of image patches is an important parameter used in RSEI. The first six reference images are selected from the TID2008 [41] reference image and the corresponding 408 distorted images for parameter tuning. We set  $n$  at 0, 5, 20, and 50 and then perform RSEI evaluation. The curve fitting results are shown in Figure 4. The patch is inaccurate when  $n$  is 0 and 5. RSEI cannot accurately evaluate all images with high MOS, and performance is not constantly improved with the increase of  $n$ . The image patch increases when  $n$  is 50, which leads to the low recognition degree of RSEI for MOS greater than 5 and high time complexity. Thus, we set  $n$  to 20.

However, there is no MOS score in the actual scenario, so that the optimal solution of  $n$  cannot be directly obtained. The larger the value of  $n$ , the finer the classification, and the longer it will take. Therefore, the value of  $n$  depends on the accuracy requirement of the task in the actual scene.



**Figure 4.** Fitting curve of number  $n$  of different image patches. RSEI has better performance when the values of  $n$  are 20 and 50.

### 3.4. Performance Comparison with State-of-the-Art IQAs

The evaluation results are compared with some representative FR IQA metrics, including some state-of-art algorithms:

- SSIM [14]—a widely used methods for IQA
- MS-SSIM [15]—the multi-scale SSIM (MS-SSIM) of wavelet domain
- FSIM [17]—the best-performance IQA method based on structural information
- VSNR [12]—wavelet-based visual SNR
- IFC [16]—an information fidelity criterion using natural scene statistics
- NQM [11]—a noise quality measure
- VIF [13]—the excellent image quality assessment method based on information theory
- WaDIQaM [20]—the state-of-the-art image quality assessment method based on deep learning

Here, the wavelet domain version of VIF is used. Other comparison algorithm results are provided by TID2008 [41] datasets, except for FSIM and VIF. For FSIM and WaDIQaM, we directly use the open-source codes provided by the author and the parameters in this study. The experiment uses 25 reference images and the corresponding six types of distorted images, which are JPEG2000 compression, JPEG compression, image denoising, quantization noise, Gaussian blur, and JPEG2000 transmission errors.

The curve fitting to MOS and image objective scoring is shown in Figure 5. All the scores of the IQA metrics are listed in Table 2. RSEI obtains the best objective score for SROCC, PLCC, and RMSE, and the next best score for KROCC. For the abovementioned distortion types, the performance of RSEI correlates more consistently with the subjective evaluations than do the other methods.

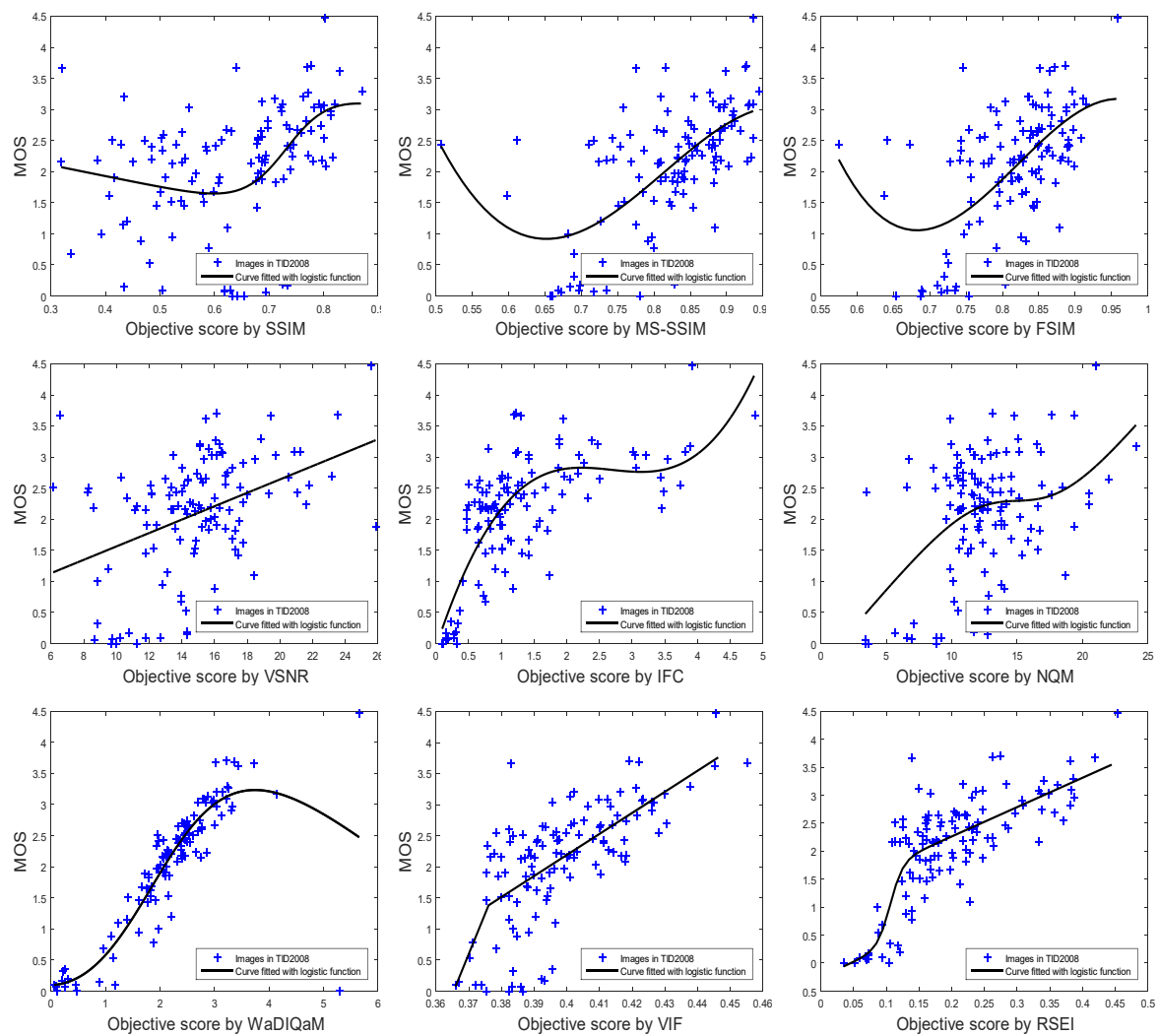


Figure 5. Scatter plots of subjective MOS versus scores obtained by model prediction.

Table 2. Performance comparison of IQA metrics. Red color indicates the best performance and blue color indicates the second best performance.

Category	Assessment Methods	KROCC	SROCC	PLCC	RMSE
Error statistic-based	PSNR	0.1328	0.2201	0.2876	0.9183
	SSIM	0.4095	0.5983	0.7521	0.6669
	MS-SSIM	0.4476	0.5643	0.8200	0.5737
	NQM	0.3333	0.4893	0.7106	0.7053
HVS-based	VSNR	0.2881	0.3357	0.8561	0.5131
	IFC	0.4286	0.6000	0.7279	0.6873
	VIF	0.6000	0.7057	0.8578	0.5152
	FSIM	0.2571	0.3436	0.6532	0.7590
	WaDIQaM	0.6067	0.7067	0.8592	0.4911
	RSEI	0.5619	0.7107	0.8593	0.5128



### 3.5. Running Time

The running time comparisons of all algorithms are shown in Figure 6. WaDIQaM has a shorter test time, however it takes a lot of time for data training, which could not be ignored. RSEI remembers the semantic segmentation labels, so it does not increase dramatically when testing large amounts of data. Combining the data in Table 2, it is clear that RSEI has the best performance by sacrificing a little time. We implement all algorithms in the experiments under the same hardware configuration, which are as follows: Intel Core i5-6300HQ CPU @2.30 GHz, 8 GB RAM, and NVIDIA GeForce GTX 950M.

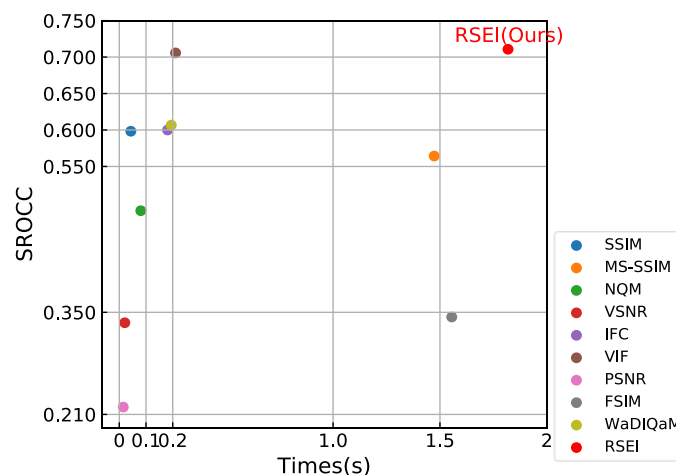


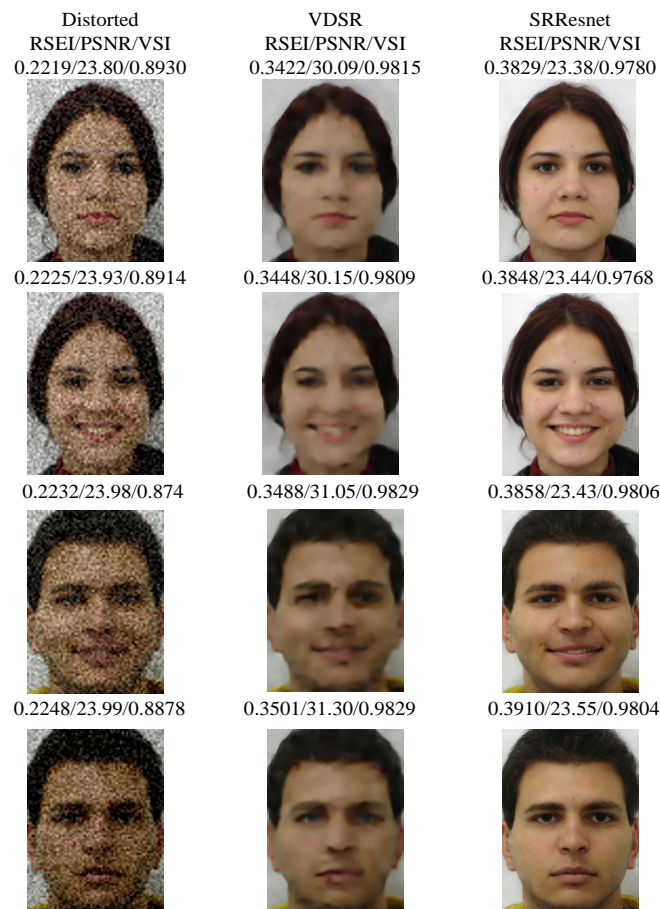
Figure 6. Mean running time (seconds) of all 150 samples for different algorithms.

### 3.6. Application of Denoise Algorithmic Scenario

Although existing IQA metrics are designed to measure standard datasets, they do not evaluate images that undergo a complex nonlinear transformation in a real algorithmic scenario. Since there is no MOS score in this algorithmic scenario, the deep learning algorithm is difficult to apply [43,44].

The FEI faces database [45] includes 400 images of 200 people (100 men and 100 women). Each person has two frontally aligned face images, where the first one is a frontal facial image and second one is a smiling image. The image size is  $260 \times 360$  pixels. Figure 7 shows several denoising results on the FEI face dataset [45] in which the images are added with different levels of Gaussian noise.  $\sigma$  is the standard deviation of the added noise to the images. The first column is the distorted image used as the input noise image with  $\sigma = 20$ . The last two columns are the denoising results of VDSR [46] and SRResnet [47], respectively. Clearly, SRResnet has a stronger denoising effect, the results also show more fluent details and color deviations.

The result of SRResnet has the highest perceptual quality, where its PSNR/VSI [7] values are low and its RSEI value is high. The performance of RSEI consistently correlates with subjective evaluations. PSNR focuses on pixel-level differences and cannot measure the image quality in HVS. VSI effectively shows that VDSR is similar in visual saliency to the ground truth but is insensitive to the detailed information of the image. However, RSEI measures the information fidelity in semantic structure patches, combines the advantages of two methods, and obtains accurate evaluation results.



**Figure 7.** Inconsistency between PSNR/VSI values and perceptual quality. The images are distorted image with noise level  $\sigma = 20$ , VDSR, and SRResnet. The result of SRResnet has the highest perceptual quality, where its PSNR/VSI values are low and RSEI value is high.

## 4. Discussion

### 4.1. Traditional Methods

Traditional algorithms use fixed-size sliding windows to simulate receptive fields, which does not incur additional running time. They ignore the spatial structure information in the image and the correlation between the pixels, and can only measure the quality of the image from the low semantic information level. RSEI semantically segments images for flexible processing of image patches, that benefited from the contributions of superpixel. However, it also leads to an increase in the running time of the algorithm.

### 4.2. Deep Learning-Based Methods

Deep learning-based methods rely on the priori information provided by a large number of training datasets to obtain the excellent performance. Since there is less public dataset for IQA, complex preprocessing is required to obtain data augmentation [20]. Meantime, the network training process depends on the acceleration of hardware devices (GPU) which is the limitation of WaDIQaM. The distortion in the real application scenario is more complex, and it is difficult to ensure the deep learning-based algorithm with good generalization ability.

## 5. Conclusions

This study introduced a novel RSEI for QIA. Semantic segmentation was applied to the variable spatial configuration information rather than using a fixed-size sliding window. RSEI is based on the perspective of a variable receptive field and IE to better measure multi-scale objects in satellite images. RSEI assigns weights to the image patch based on its degree of information richness. We verified the effectiveness of RSEI by applying it to the data from the TID2008 database, denoising algorithms, and inaccurate supervision application scenarios. We believed that the proposed approach is appropriate to this satellite application scenario, in which the ground truth of satellite images is frequently unavailable.

**Author Contributions:** Conceptualization and Writing—Original Draft preparation, T.L. and J.W.; supervision, H.Z.; software, J.J. and J.M.; editing, Z.W.

**Funding:** This work is supported by the NSFC grants (61502354, 61671332, 41501505, 61771353), the Central Guide to Local Science and Technology Development (2018ZYYD059), the Natural Science Foundation of Hubei Province of China (2018ZYYD059, 2012FFA099, 2012FFA134, 2013CF125, 2014CFA130, 2015CFB451), Scientific Research Foundation of Wuhan Institute of Technology (K201713), the Beijing Advanced innovation Center for intelligent Robots and Systems under Grant (2016IRS15).

**Acknowledgments:** We would like to thank the anonymous reviewers for their very helpful and professional comments to improve the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Wang, Z.; Sheikh, H.R.; Bovik, A.C. Objective video quality assessment. In *The Handbook of Video Databases: Design and Applications*; CRC Press: Boca Raton, FL, USA, 2003; Volume 41, pp. 1041–1078.
2. Anbarjafari, G.; Jafari, A.; Jahromi, M.N.S.; Ozcinar, C.; Demirel, H. Image illumination enhancement with an objective no-reference measure of illumination assessment based on Gaussian distribution mapping. *Eng. Sci. Technol. Int. J.* **2015**, *18*, 696–703. [[CrossRef](#)]
3. Magni, S.; Magni, S.; Tagliasacchi, M.; Tubaro, S. No-Reference Pixel Video Quality Monitoring of Channel-Induced Distortion. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 605–618.
4. Wang, Z.; Bovik, A.C. A universal image quality index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84. [[CrossRef](#)]
5. Liu, A.; Lin, W.; Narwaria, M. Image quality assessment based on gradient similarity. *IEEE Trans. Signal Process. Soc.* **2012**, *21*, 1500.
6. Zhou, F.; Lu, Z.; Wang, C.; Sun, W.; Xia, S.T.; Liao, Q. Image quality assessment based on inter-patch and intra-patch similarity. *PLoS ONE* **2015**, *10*, e0116312. [[CrossRef](#)] [[PubMed](#)]
7. Zhang, L.; Shen, Y.; Li, H. VSI: A Visual Saliency-Induced Index for Perceptual Image Quality Assessment. *IEEE Trans. Image Process.* **2014**, *23*, 4270–4281. [[CrossRef](#)]
8. Guo, S.; Xiang, T.; Li, X. Image quality assessment based on multiscale fuzzy gradient similarity deviation. *Soft Comput.* **2017**, *21*, 1145–1155. [[CrossRef](#)]
9. Wang, Z.; Li, Q. Information content weighting for perceptual image quality assessment. *IEEE Trans. Image Process.* **2011**, *20*, 1185. [[CrossRef](#)]
10. Ponomarenko, N.; Ieremeiev, O.; Lukin, V.; Egiazarian, K.; Carli, M. Modified image visual quality metrics for contrast change and mean shift accounting. In Proceedings of the 2011 11th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), Polyana-Svalyava, Ukraine, 23–25 February 2011; pp. 305–311.
11. Damara-Venkata, N.; Kite, T.D.; Geisler, W.S.; Evans, B.L.; Bovik, A.C. Image quality assessment based on a degradation model. *IEEE Trans. Image Process.* **2000**, *9*, 636–650. [[CrossRef](#)]
12. Chandler, D.M.; Hemami, S.S. VSNR: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Trans. Image Process.* **2007**, *16*, 2284–2298. [[CrossRef](#)]
13. Sheikh, H.R.; Bovik, A.C. Image information and visual quality. *IEEE Trans. Image Process.* **2006**, *15*, 430–444. [[CrossRef](#)] [[PubMed](#)]
14. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]

15. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale structural similarity for image quality assessment. In Proceedings of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 9–12 November 2004; Volume 2, pp. 1398–1402.
16. Sheikh, H.R.; Bovik, A.C.; Veciana, G.D. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Process.* **2005**, *14*, 2117–2128. [[CrossRef](#)] [[PubMed](#)]
17. Zhang, L.; Zhang, L.; Mou, X.; Zhang, D. FSIM: A Feature Similarity Index for Image Quality Assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [[CrossRef](#)] [[PubMed](#)]
18. Li, J.; Zhang, X.; Ding, M. Image Quality Assessment Based on Regional Mutual Information. *AEUE Int. J. Electron. Commun.* **2012**, *66*, 784–787. [[CrossRef](#)]
19. Sheikh, H.R.; Sabir, M.F.; Bovik, A.C. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Trans. Image Process.* **2006**, *15*, 3440–3451. [[CrossRef](#)] [[PubMed](#)]
20. Bosse, S.; Maniry, D.; Müller, K.R.; Wiegand, T.; Samek, W. Deep neural networks for no-reference and full-reference image quality assessment. *IEEE Trans. Image Process.* **2018**, *27*, 206–219. [[CrossRef](#)]
21. Ma, J.; Jiang, J.; Liu, C.; Li, Y. Feature guided Gaussian mixture model with semi-supervised EM and local geometric constraint for retinal image registration. *Inf. Sci.* **2017**, *417*, 128–142. [[CrossRef](#)]
22. Ma, J.; Jiang, J.; Zhou, H.; Zhao, J.; Guo, X. Guided locality preserving feature matching for remote sensing image registration. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4435–4447. [[CrossRef](#)]
23. Lu, T.; Zeng, K.; Xiong, Z.; Jiang, J.; Zhang, Y. Group Embedding for Face Hallucination. *IEEE Access* **2018**, *6*, 63402–63415, doi:10.1109/ACCESS.2018.2877422. [[CrossRef](#)]
24. Zhou, H.; Kuang, Y.; Yu, Z.; Ren, S.; Zhang, Y.; Lu, T.; Ma, J. Image Deformation with Vector-field Interpolation based on MRLS-TPS. *IEEE Access* **2018**. [[CrossRef](#)]
25. Hore, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In Proceedings of the 20th International Conference on Pattern recognition (icpr), Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369.
26. Nasiri-Avanaki, M.R.; Ebrahimpour, R. In-service video quality measurements in optical fiber links based on neural network. *Neural Netw. World* **2007**, *17*, 457.
27. Levinshtein, A.; Stere, A.; Kutulakos, K.N.; Fleet, D.J.; Dickinson, S.J.; Siddiqi, K. Turbopixels: Fast superpixels using geometric flows. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 2290–2297. [[CrossRef](#)] [[PubMed](#)]
28. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [[CrossRef](#)] [[PubMed](#)]
29. Ciecholewski, M. Automated coronal hole segmentation from Solar EUV Images using the watershed transform. *J. Vis. Commun. Image Represent.* **2015**, *33*, 203–218. [[CrossRef](#)]
30. Cousty, J.; Bertrand, G.; Najman, L.; Couprie, M. Watershed cuts: Thinnings, shortest path forests, and topological watersheds. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 925–939. [[CrossRef](#)] [[PubMed](#)]
31. Balla-Arabé, S.; Gao, X. Geometric active curve for selective entropy optimization. *Neurocomputing* **2014**, *139*, 65–76. [[CrossRef](#)]
32. Ding, K.; Xiao, L.; Weng, G. Active contours driven by region-scalable fitting and optimized Laplacian of Gaussian energy for image segmentation. *Signal Process.* **2017**, *134*, 224–233. [[CrossRef](#)]
33. Jampani, V.; Sun, D.; Liu, M.Y.; Yang, M.H.; Kautz, J. Superpixel Sampling Networks. *arXiv* **2018**, arXiv:1807.10174.
34. Fan, F.; Ma, Y.; Li, C.; Mei, X.; Huang, J.; Ma, J. Hyperspectral image denoising with superpixel segmentation and low-rank representation. *Inf. Sci.* **2017**, *397*, 48–68. [[CrossRef](#)]
35. Maes, F.; Collignon, A.; Vandermeulen, D.; Marchal, G.; Suetens, P. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging* **1997**, *16*, 187–198. [[CrossRef](#)] [[PubMed](#)]
36. Rueckert, D.; Clarkson, M.; Hill, D.; Hawkes, D.J. Non-rigid registration using higher-order mutual information. In Proceedings of the Medical Imaging 2000: Image Processing, International Society for Optics and Photonics, San Diego, CA, USA, 12 February 2000; Volume 3979, pp. 438–448.
37. Flannery, B.P.; Flannery, B.P.; Teukolsky, S.A.; Vetterling, W.T. *Numerical Recipes: the Art of Scientific Computing*; Cambridge University Press: Cambridge, UK, 1986; pp. 120–122.
38. Peng, L.; Zhang, Y.; Zhou, H.; Lu, T. A robust method for estimating image geometry with local structure constraint. *IEEE Access* **2018**, *6*, 20734–20747. [[CrossRef](#)]

39. Lu, T.; Chen, X.; Zhang, Y.; Chen, C.; Xiong, Z. SLR: Semi-coupled locality constrained representation for very low resolution face recognition and super resolution. *IEEE Access* **2018**. [[CrossRef](#)]
40. Fu, H.; Cao, X.; Tu, Z. *Cluster-Based Co-Saliency Detection*; IEEE Press: New York, NY, USA, 2013; pp. 3766–78.
41. Ponomarenko, N.; Lukin, V.; Zelensky, A.; Egiazarian, K.; Carli, M.; Battisti, F. TID2008—A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics. *Adv. Mod. Radioelectron.* **2004**, *10*, 30–45.
42. Greg, C.; Philip, C.; Mylène, C.Q.F.; Taali, M.; Margaret, P.; Filippo, S.; Arthur, W. Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II (FR\_TV2). 2003. Available online: [https://www.its.bldrdoc.gov/media/4150/vqegii\\_final\\_report.doc](https://www.its.bldrdoc.gov/media/4150/vqegii_final_report.doc) (accessed on 7 December 2018).
43. Wang, G.; Liu, Y.; Zhao, T. A quaternion-based switching filter for colour image denoising. *Signal Process.* **2014**, *102*, 216–225. [[CrossRef](#)]
44. Lu, T.; Xiong, Z.; Zhang, Y.; Wang, B.; Lu, T. Robust face super-resolution via locality-constrained low-rank representation. *IEEE Access* **2017**, *5*, 13103–13117. [[CrossRef](#)]
45. Thomaz, C.E.; Giraldi, G.A. A new ranking method for principal components analysis and its application to face image analysis. *Image Vis. Comput.* **2010**, *28*, 902–913. [[CrossRef](#)]
46. Kim, J.; Lee, J.K.; Lee, K.M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR Oral), Las Vegas, NV, USA, 27–30 June 2016.
47. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.P.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *CVPR* **2017**, *2*, 4.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).