Texture Smoothing Quality Assessment via Information Entropy

CHONG LIU¹,², CUI YANG², MINGQIANG WEI¹, AND JUN WANG¹

¹College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
²College of Mathematics and Computing Science, Anqing Normal University, Anqing 246003, China

Corresponding authors: Mingqiang Wei (mqwei@nuaa.edu.cn) and Jun Wang (wjun@nuaa.edu.cn)

This work was supported by the Key Laboratory of Modeling, Simulation and Control of Complex Ecosystem, Dabie Mountains of Anhui Higher Education Institutes, Anqing Normal University.

ABSTRACT Image texture smoothing (ITS) aims at completely removing textures while preserving as much as possible different-scale structures of an image. However, few quality assessment metrics have been formulated to objectively evaluate the ITS results, due to the lack of ground-truth texture-free images. Considering both the smoothness of textures and the preservation of structures, we make the debut of objective evaluation of ITS results, and design an intuitive Texture Smoothness and Structural Similarity Index (T3SI) based on human perception. Specifically, we first intuitively select some patches which contain relatively more texture/structure information. We then employ the Edge Preservation Index and the Structural Similarity Index Measure to evaluate the texture smoothness and the structure similarity on the selected patches, respectively. We finally formulate the novel exponential entropy function to balance the texture smoothness and the structure similarity for the quality assessment of ITS. T3SI is objective, easy to use, simple to implement, and stable. An independent User Study is performed to verify our proposed T3SI, and large experiments show that T3SI competes successfully to the state-of-the-art metrics. Our code is publicly available.

INDEX TERMS Texture filtering, IQA, smoothness, structure similarity, entropy.

I. INTRODUCTION Texture, as a set of texture elements or texels occurring in some regular or repetitive patterns, can be artificially created in an image or found in the captured natural scene images, such as mosaics, sand, and rocks. Image texture smoothing is commonly used to remove those redundant background textures that may be global (or local) details existing in an image. Unlike denoised images, the detextured image quality assessment needs to measure the loss of texture details while also counting the retention of structure features in relative areas. Since 1) the texture-free images as the ground truths are missing, and 2) a comprehensive consideration of texture smoothness and structure retention is required, it is challenging to effectively assess texture filtered image quality. So far, existing image quality assessment (IQA) metrics that can be usable for measuring detextured images are very rare. Designing an objective and quantitative evaluation index is not only conducive to the comparisons of texture filtering works, but also beneficial to many practical low- and high-level vision applications.

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

For decades, User Study is widely used to evaluate de-textured image quality, which is required to design an independent user evaluation system and normalizes the viewer’s score to perform a qualitative assessment of a filtered image. The intuitive and simple assessment of User Study without complex mathematical formulas is suitable to practical applications. However, when some filtered images are observed very similar by human perceptions, the intuitive results from User Study are difficult to achieve a reasonable evaluation. In order to fully comply with human visions, the human visual system (HSV) [1], [2] based on salient physiological and psychovisual features have been presented, which capture the statistics of important natural signals to match the result of the subjective evaluation. Similarly, Information Fidelity Criterion (IFC) [3] and Blind Image Quality Assessment (BIQA) [4] are also devised by using the statistically significant feature information. Due to the lack of real human perceptions, these metrics could not fully consider the unity of subjective textures and objective features of texture filtered images, and lead to a large deviations from subjective results.

Other metrics focus on structure details determined by human perceptions, such as subjective methods [5]–[7]. Edge Preservation Index (EPI) [5] based on regions of
interest (ROI) is obtained according to the correlation of signal intensity in ROI, which has a good performance for measuring local structure sharpness. Besides, the improvement of PSNR (IPSNR) [8] adopts local blocks to calculate PSNR, which concentrates on the lower bounds of filtering efficiency for the visual quality of images. Differently, texture filtering should focus on the texture smoothness rather than the sharpness of the texture. These quality evaluation methods based on local details only analyze the edge features of the image while not synthesizing the texture smoothness and the structural retention, which could not achieve a comprehensive quality measurement. To fully compare the overall performance on the measurement of texture filtered images, the early common methods: Mean Square Error (MSE) [9], Peak-Signal to Noise Ratio (PSNR) [10] and Structure Similarity Index Measure (SSIM) [11] have been utilized for comprehensively evaluating filtered images by some texture filters [12]–[14]. These common methods use a global measurement scheme, yet they could not provide an effectively comprehensive evaluation result through the flatness of texture and the sharpness of structure from a de-textured image.

The aforementioned objective evaluation metrics perform well on de-noised or de-blurred images, while they are limited to comprehensively measure the texture smoothness and the structure retention. Especially, when texture details in some texture images need to be subjectively recognized by human perception, some texture details could be easily misjudged as structures by these assessment methods. In this work, aiming at the subjective definition of textures, the locality of texture details, and the comprehensiveness of IQA, we integrate texture smoothness and structure similarity to construct a comprehensive objective evaluation index T3SI, based on the intuitive selection of textures and features. We first manually select some patches from the original image and the filtered image, and these patches should respectively contain relatively more textures and features defined by human perception. Thus, these selected patches can represent texture and structure regions. Then, EPI is employed to measure the texture smoothness through some selected patches representing texture regions, and SSIM is used to measure the edge structure similarity based on these patches including significant features. Combining EPI and SSIM by using the exponential cross entropy, we can construct a final comprehensive exponential cross-entropy assessment function for texture filtered images.

The contributions of our approach are as follows.

- We intuitively identify textures and structural features by human perception, due to the lack of ground-truth texture-free images. The difficult quantifiable task of texture smoothing evaluation can be easily transformed into a quantified problem.
- We use local patches to distinguish image texture details from structural details, which could ensure the accurate measurements for the texture smoothness and the structural similarity.
- Our T3SI synthesizes the local evaluated results of texture smoothness and structure retention via the exponential entropy, which can provide a comprehensive and quantitative measurement for texture filtering.

The remainder of this paper is organized as follows. In Section II, we briefly reviewed the related work, including no-reference and full-reference IQA methods. In Section III & IV, we specifically described the idea of our approach, including the recognition and definition of texture regions, the measurement of texture smoothness and structure similarity, and the construction of comprehensive index T3SI. In Section V, we analyzed the implementation of our method and compared with several similar methods, followed by the conclusion in Section VI.

II. RELATED WORK

Among IQA metrics, some full-reference approaches could be used for measuring low-scale texture images. Generally, since most texture images obtained from nature usually have no real background, no-reference IQA models have a certain applicability in texture images. In this section, we mainly introduce the related no-reference and full-reference methods of image quality assessment.

The current no-reference image quality assessment is a very popular research, and various excellent algorithms [15]–[19] have been designed for evaluating different distorted images. NIQE [15] mainly considers the construction of a ‘quality aware’ collection of statistical features based on a natural scene statistic (NSS) model. Similarly, Appina et al. [16] proposed a non-reference image quality evaluation model based on natural scene statistics, which can accurately extract the luminance coefficient and the parallax subband coefficient and achieve excellent performance. PIQE [17] is an opinion unaware methodology that attempts to quantify distortion by extracting local features for predicting quality, and BRISQUE [18] uses scene statistics of locally normalized luminance coefficients to quantify possible losses of “naturalness” in the image. Differently, Akhter et al. [19] first extracts the segmentation and disparity map features from the left and right views, and uses the logistic regression to predict the image quality score. Besides, some no-reference IQA models [19]–[22] have been developed for perceptual quality evaluation of stereoscopic images. These no-reference methods can analyze image quality based on the ‘sharpness’ of statistically feature details of distorted images. Whereas for complex texture filtered images, the result of adequate smoothing texture image will result in the opposite evaluation.

Unlike no-reference IQA approaches, full-reference models are more widely applied for performing the quality assessment of images with a truth-background, such as images added virtual noise, enhanced images and segmented images. The evaluation metrics mainly have PSNR [23], MSE [9], [24], [25] and SSIM [11]. PSNR is commonly used to evaluate noise images. MSE and its extension PAMSE [26] and CW-MSE [27] were proposed based on average error.
They measure the quality of an image by calculating the global size of the pixel error between the distorted image and the reference image, which are relatively simple and easy to implement. Considering the statistical correlation of structure features, SSIM and its improvements: MS-SSIM [28], MS-SSTEM [29], RFSIM [30] are also widely used for various image quality assessment. The similar methods of SSIM define the structural information as the structure of the object according to the brightness and contrast, which could perform well on the image with obvious edges feature. To highlight the local region of interest, QILV [31] utilizes the comparison of the local variance distribution of two images to assess the non-stationarity of images. In addition, SFF [32] adopts the sparse features by a feature detector, and IFS [33] uses only a part of the reference image information. Since the loss of texture details is regarded as a significant index, these full-reference metrics are also difficult to provide a comprehensive measurement of textures and structures.

Our evaluation approach is devised to comprehensively measure the smoothness of texture and the retention of structure on a relative full-reference.

**III. OVERVIEW**

To provide an overall view, we illustrate the pipeline in Fig. 1. and succinctly introduce it as follows.

First, we intuitively identify some texture/structure regions on the original image by human perception and give some clicks on the centers of these regions by the mouse. Setting an appropriate radius \( r \), we can obtain some texture/structure patches with the size \((r + 1) \times (r + 1)\) centered at these clicked points. Similarly, according to the positions clicked on the original image, we can also crop some texture/structure patches from the filtered image.

Second, based on these texture/structure patches, we employ EPI and SSIM to calculate the negative texture smoothness and the structure similarity, respectively.

Finally, combining the results of 1-EPI and SSIM, we formulate an exponential cross-entropy function (T3SI) to evaluate texture filtered images comprehensively.

**IV. METHOD**

Due to the locality and unrecognizability of textures/structures, we construct a comprehensive assessment metric from the following three aspects: 1) the determination of texture and structure, 2) the measurement of texture smoothness and structural retention, 3) the formulation of comprehensive evaluation index.

**A. DETERMINATION OF TEXTURE AND STRUCTURE**

The accurate identification of textures and structures is a prerequisite for IQA. Generally, the textures and structures in an image are not distinguished by a strict criteria, and even the different texture regions may be defined from the same image by various filtering methods. For instance, the flecks on giant salamander in Fig. 2 (a) are structural features but are regarded as texture details by some filtering methods in specific applications. In order to suit the actual applications, we adopted the intuitive method that is also the most direct and effective approach to identify textures/structures by human perception.

Since the texture region has a structural consistency between the local region and the whole region, we could utilize some local texture patches to substitute the whole texture region and achieve the measurement of the texture smoothness. Recognized texture areas by human perception and given \( N \) clicks inside texture areas (black background in Fig. 2 (a) ) by the mouse, the positions \( p_i (i = 1, 2, \ldots, N) \) (yellow ‘⋆’ in Fig. 2 (a)) representing texture could be marked on the original texture image. We set an appropriate patch radius \( r \) to ensure that each selected patch \( \Omega_i \) with the size \((2r + 1) \times (2r + 1)\) centered at \( p_i \) is located inside texture areas, and abstract some box patches according to the selected points \( p_i (i = 1, 2, \ldots, N) \). Once all box patches are cropped from one texture region, we obtain an accurate texture patches’ set \( S = \bigcup_{i=1}^{N} \Omega_i \) which can represent all texture regions (see yellow boxes in Fig. 2 (b)).

The structure positions \( q_i (i = 1, 2, \ldots, N) \) (red ‘⋆’ in Fig. 2 (a)) should be located on some areas including lots of structure positions.

**FIGURE 1.** The pipeline of our evaluation approach for texture filtered images.
of structure features. Similar to texture patches, we can also obtain $N$ structural patches' set $\tilde{S} = \bigcup_{i=1}^{N} \Phi_i$ (see red boxes in Fig. 2 (b)). Here $\Phi_i$ represents the i-th structural patch.

According to the same positions as the corresponding clicks' positions $p_i$, $q_i (i = 1, \ldots, N)$, we could also obtain patches representing texture regions $S_1$ and structure regions $\tilde{S}_1$ from the filtered image (see Fig. 2 (c)). By using these selected patch sets $S, \tilde{S}, S_1, \tilde{S}_1$ from the original image and the filtered image, we can calculate the texture smoothness and the structure retention.

### B. TEXTURE SMOOTHNESS

We consider the residual edge structures in texture regions as the negative measurement of the texture smoothness, and adopt the edge preservation index $EPI$ [5] to count the retention of edges, since $EPI$ based on the region of interest (ROI) can provide a good supplementary evaluation for edge structures in the local texture region. Not difficult to find, when a texture region is fully smoothed, there is a large difference between the clean original region and the filtered region. It means that a smoother texture region corresponds to a smaller value of $EPI$. Hence, the negative value of $1-EPI$ could be directly used to measure the smoothness of texture regions.

$EPI$ [5] can be expressed as

$$EPI = \frac{\Gamma(x - \bar{x}, \hat{x} - \bar{\hat{x}})}{\sqrt{\Gamma(x - \bar{x}, x - \bar{x})\Gamma(\hat{x} - \bar{x}, \hat{x} - \bar{\hat{x}})}}$$

where $\bar{x}$ and $\bar{\hat{x}}$ are mean pixel values in the sets $S$ and $\tilde{S}_1$ from the original image and the filtered image, and

$$\Gamma(x, \hat{x}) = \sum_{i=1}^{T} x_i \hat{x}_i$$

where $T$ represents the number of pixels in $S$, $x_i$ and $\hat{x}_i$ are the pixel value in $S$ and $S_1$. To get significant edge information, the original and filtered images need to be operated by a Laplacian high-pass filter, respectively.

### C. STRUCTURE SIMILARITY

The structure similarity of $\tilde{S}$ and $\tilde{S}_1$ could be regarded as the retention of structure features on the filtered image. Unlike the texture smoothness, the structure similarity may require more information to measure the distortion of structural details, such as brightness, contrast and structural composition, etc. Since Structure Similarity Index Measure (SSIM) [11] could reflect these properties of the object structure in the scene, we use SSIM as the structure similarity index to measure the retention of features in structure regions.

SSIM [11] can be expressed as

$$SSIM = \frac{(2\tilde{S}v_{\tilde{S}_1} + C_1)(2\tilde{S}\tilde{S}_1 + C_2)}{(\tilde{S}^2 + v_{\tilde{S}_1}^2 + C_1)(\tilde{S}_1^2 + v_{\tilde{S}}^2 + C_2)}$$

where $v_{\tilde{S}}$ and $v_{\tilde{S}_1}$ are the average intensity of patches $\tilde{S}$ and $\tilde{S}_1$, respectively. The parameters $\sigma_{\tilde{S}}^2$ and $\sigma_{\tilde{S}_1}^2$ represent corresponding standard deviations of the sets $\tilde{S}$ and $\tilde{S}_1$. $\sigma_{\tilde{S}\tilde{S}_1}$ is the covariance between $\tilde{S}$ and $\tilde{S}_1$. $C_1$ and $C_2$ are constant, which can be usually set: $C_1 = (K1 * L)^2$, $C_2 = (K2 * L)^2$, $K1 = 0.01$, $K2 = 0.03$, $L = 255$.

### D. EVALUATION INDEX T3SI

According to the analysis of $1 - EPI$ in section IV-B, $1 - EPI$ is positive related to the texture smoothness, which could be used as a measure of the loss information of texture details. Similarly, the structure similarity index $SSIM$ can also positively measure the retention information of feature details. $1 - EPI$ and $SSIM$ belong to $[0,1]$, which could be regarded as information energy. Higher values of them indicate higher information energy. As is well known, the entropy function $-xlog_2(x)$ is very suitable for the formulation of the information energy loss. Consequently, we could synthesize the two information indexes $1 - EPI$ and $SSIM$ to construct a cross entropy function, named Texture Smoothness and Structure Similarity Index (T3SI). T3SI is used to evaluate texture smoothed images, and can be defined as

$$T3SI = -(1 - \xi) \log_2 (1 - \xi) - \theta \log_2 \theta$$

where $\xi = EPI$ is calculated by sets $S$ and $\tilde{S}$, and $\theta = SSIM$ is computed through the sets $S_1$ and $\tilde{S}_1$.

Since $\xi$ and $\theta$ belong to the interval $[0,1]$, T3SI is not a monotonic increasing function on the variable of $\xi$ and $\theta$ (see Fig. 3 (a)). The non-uniform change of the function T3SI produced by the value of the variable $\xi$ and $\theta$ will lead to an unsuccessful comprehensive quality assessment. Analyzing the entropy formulation $-xlog_2(x)$, we can easily obtain the
Then T3SI is written as follows.

\[
T_3SI = \exp\left(-\bar{1} - \bar{\theta} \log_2 \left(1 - \bar{\xi} - \theta \log_2 \bar{\theta}\right)\right)
\]

Due to the loss of calculation error, the value of T3SI for the similar filtered images will be same. The same value of \(\bar{\xi}\) and \(\bar{\theta}\) that belong to the interval (0, 0.4).

The normalized formulation of \(\xi\) and \(\theta\) can be expressed as

\[
\bar{\xi} = \xi(0.36 - \varepsilon) + \varepsilon \\
\bar{\theta} = \theta(0.36 - \varepsilon) + \varepsilon
\]

\(\varepsilon\) is set as an infinitesimal constant, such that \(\bar{\xi}\) and \(\bar{\theta}\) are not equal to 0.

Algorithm 1 Texture IQA Index T3SI

| Input: | Original image \(I_0\) and filtered image \(I_1\) |
|---|---|
| Output: | T3SI |
| 1: | \(p_i, q_i \leftarrow\) Manual positions selected in \(I_0\) |
| 2: | \(r \leftarrow\) Set a patch radius |
| 3: | \(S, \tilde{S} \leftarrow\) Select patches centered at \(p_i, q_i\) from \(I_0\) |
| 4: | \(S_1, \tilde{S}_1 \leftarrow\) Select patches centered at \(p_i, q_i\) from \(I_1\) |
| 5: | \(1 - EPI \leftarrow\) Computed by Eq. (1) according to \(S, \tilde{S}\) |
| 6: | \(SSIM \leftarrow\) Computed by Eq. (2) according to \(S_1, \tilde{S}_1\) |
| 7: | \(\bar{\xi}, \bar{\theta} \leftarrow\) Normalize \(1 - EPI\) and \(SSIM\) by Eq. (4) - (5) |
| 8: | T3SI \(\leftarrow\) Obtain the evaluation result through Eq. (6) |

V. EXPERIMENT ANALYSIS

We collect some classic and meaningful texture images (see Fig. 5) as a reference set, and use excellent texture methods: RGF [34], BF [35], RTV [36], mRTV [37] and ROG [38] to perform on the set. After the set is processed by each filtering method with different parameters, we can obtain 5 groups of test sets, and each group test set contains 48 filtered images that could be used to detect the evaluation effect of different metrics in subsequent analysis and experiments.

A. IMPACT OF PATCH

The result of T3SI ultimately depends on the texture and structure patches, and a different selection of patches will result in a different value of T3SI. In order to analyzing the impact of the patches on T3SI, we chose a typical mosaic filtered picture 'harp' from one group of test set, and performed experiments on different positions and number of patches, respectively.

1) IMPACT OF PATCHES’ LOCATIONS

Obviously, if we mistake the structural area as a texture region to calculate EPI, we will inevitably obtain a wrong evaluation result. Therefore, texture patches should contain as few feature details as possible for subjective applications, and structure patches should also include significant edge features. Putting the same number of clicks inside different texture regions and structure regions four times, we can obtain four texture and structure patches’ sets. Fig. 6 (a) - (d) only shows four different locations of texture and structure patches on the filtered image, respectively. Based on these patches of four random locations, EPI, SSIM and T3SI of the filtered image were computed in Tab. 1.

The selection of texture patches in Fig. 6 (a) - (c) are fixed in some positions, and the structural patches are located in different positions each time. Conversely, the selection of structure patches in Fig. 6 (e) - (g) does not change. We could easily find from the results in Tab. 1 that the deviation of...
FIGURE 5. Texture image set.

FIGURE 6. Selection of patches’ location. Patches representing texture regions and structure regions are selected six times, and each patch is set to the same radius $r = 15$. The corresponding value of T3SI is labeled below each sub-figure.

TABLE 1. The results of EPI, SSIM and T3SI calculated by the patches selected on the different positions as Fig. 2 (a) - (d).

| Position of patches | EPI   | SSIM  | T3SI  |
|---------------------|-------|-------|-------|
| Position#1          | 0.0886| 0.8022| 2.8438|
| Position#2          | 0.0886| 0.7968| 2.8439|
| Position#3          | 0.0886| 0.8132| 2.8494|
| Position#4          | 0.0758| 0.7881| 2.8428|
| Position#5          | 0.0897| 0.7881| 2.8407|
| Position#6          | 0.0705| 0.7881| 2.8435|

SSIM and EPI are very small. It indicates that patches’ positions have a less effect on T3SI. T3SI has global stability in the selection of patches’ locations.

2) IMPACT OF PATCHES’ NUMBER
We randomly select nine filtered images with typical textures and features from a test set, and set a moderate patch radius $r = 12$ to ensure all patches inside texture regions for each test image. The values of T3SI are calculated through the different number of patches, and the relationship between T3SI and the number of patches is shown in Fig. 7. We could find that T3SI is basically stable when choosing 5 patches. It shows that only 4-6 patches containing general textures and significant structures could accurately calculate the index T3SI.

B. COMPARATIVE ANALYSIS
We design an independent User Study for verifying the comparison of our T3SI and other IQA metrics. For the image sharpness, the texture smoothness and the structure retention, we invite 100 volunteers to comprehensively evaluate the visual quality according to ten scores from 1 to 10. The final comprehensive evaluation result of each filtered image is given by the average score of 100 volunteers. Compared User Study with our proposed metric, the results of test images demonstrate our proposed comprehensive evaluation for each test image set is almost consistent with the subjective User Study. More experimental results are exhibited in the following comparison.

1) OVERALL PERFORMANCE
Choosing recent excellent IQA metrics: NIQE [15], PIQE [17], BRISQUE [18], SFF [32], IFS [33], QILV [31] and our T3SI to perform one group of test images, we compared the results of this metrics and User Study. According to the results of the User Study ranging from the smallest
FIGURE 8. Comparisons between various assessment methods and user study.

FIGURE 9. A simple reference image and its filtered test images. (a) Reference image. (b)-(f) Five filtered test images #01 - #05. The area in each blue box represents the texture region, and the area in each red box contains typical structure details. The two corresponding local areas are zoomed on the right of each image.

to the largest, the each group of test filtered images are sequentially labeled \(i(i = 1, 2, \cdots)\) accordingly, and the results of other evaluation methods are also sorted in the order of image labels. Normalizing the results of User Study and each metric into the interval [0,1], we could plot the comparative curves between User Study and each quality assessment method. Fig. 8 shows the comparative results with User Study, from which we can easily find that, the assessment result of T3SI is almost similar to User Study (see Fig. 8 (g)). There is a large difference between the results of other methods and User Study, since User Study may adopt through the comprehensive measurement of texture smoothness and the structure sharpness that are two essential indexes for the texture image assessment. It verifies that our proposed objective evaluation T3SI is applicable to subjective computer vision.

2) INDIVIDUAL PERFORMANCE

We respectively select a simple, a medium-scale and a complex texture image from Fig. 5, and the three corresponding texture filtered images are also taken from the 5 test sets. Using the three references and corresponding test images, the specific results of all aforementioned evaluation methods are listed in Tab. 2. For the subjective observation, we also
TABLE 2. The different measurement of simple texture filtered images (see Fig. 9 (b) - (e)) provided by five filtering methods for the same input. The data in the table is the assessment values of various metrics: NIQE [15], PIQE [17], BRISQUE [18], SFF [32], IFS [33], QILV [31], T3SI and user study.

| Texture pattern | Image   | NIQE | PIQE | BRISQUE | SFF | IFS | QILV | T3SI | User Study |
|-----------------|---------|------|------|---------|-----|-----|------|------|------------|
| simple          | Image#01| 3.8888 | 34.6190 | 19.3638 | 0.9364 | 0.8222 | 0.9975 | 2.2517 | 7.1156 |
| Image#02       | 4.0605 | 26.5581 | 20.2005 | 0.9242 | 0.8059 | 0.9956 | 2.3255 | 7.3489 |
| Image#03       | 7.0500 | 83.6318 | 53.8428 | 0.8935 | 0.7053 | 0.9943 | 2.6352 | 8.3274 |
| Image#04       | 7.2167 | 88.2207 | 56.8858 | 0.8967 | 0.7135 | 0.9944 | 2.6409 | 8.3454 |
| Image#05       | 6.1335 | 79.8131 | 58.8933 | 0.9038 | 0.7394 | 0.9956 | 2.7216 | 8.2846 |
| medium-scale    | Image#11| 4.8803 | 23.7772 | 16.7665 | 0.9687 | 0.8935 | 0.9953 | 2.2314 | 7.0528 |
| Image#12       | 4.9413 | 36.8941 | 42.2886 | 0.9372 | 0.8482 | 0.9835 | 2.3596 | 7.4579 |
| Image#13       | 5.3839 | 85.9535 | 49.6996 | 0.9225 | 0.7626 | 0.9857 | 2.7714 | 8.7596 |
| Image#14       | 5.5066 | 85.0237 | 46.4916 | 0.9223 | 0.7738 | 0.9842 | 2.7793 | 8.8161 |
| Image#15       | 4.6906 | 80.1667 | 30.5145 | 0.9332 | 0.7913 | 0.9924 | 2.7882 | 9.1287 |
| complex         | Image#21| 6.5677 | 31.9382 | 43.4581 | 0.9447 | 0.8466 | 0.9779 | 2.0631 | 6.5132 |
| Image#22       | 4.6901 | 20.8125 | 43.4582 | 0.9083 | 0.7958 | 0.9440 | 2.2922 | 6.6052 |
| Image#23       | 6.9667 | 84.3539 | 45.3280 | 0.8755 | 0.6754 | 0.9454 | 2.8011 | 8.7770 |
| Image#24       | 7.2448 | 84.5157 | 44.5202 | 0.8789 | 0.6781 | 0.9472 | 2.8120 | 8.8777 |
| Image#25       | 5.8908 | 81.3370 | 43.0126 | 0.8800 | 0.6710 | 0.9542 | 2.8183 | 8.8722 |

FIGURE 10. A medium-scale texture reference image and its filtered test images. (a) Reference image. (b)-(f) Five filtered test images#11 - #15. The area in each blue box represents the texture region, and the area in each red box contains typical structure details. The two corresponding local areas are zoomed on the right of each image.

zoomed some local details in each selected image as shown in Fig. 9 - 11. We could find from the local zoomed textures and structures that, the visual results of de-texturing images (b) & (c) in Fig. 9 - 11 are relatively poor, and (d) & (f) in Fig. 9 - 11 are difficult to compare intuitively.

Since the range of each evaluation is not uniform, we normalized eight groups of evaluation values to the interval [0.1, 0.9] for better graphical comparison. Each group of evaluation results is exhibited by five bar graphs (see Fig. 12) and the height of bar represents the texture IQA. Comparing with the heights of bars, we can find that no-reference quality assessment algorithms: NIQE, PIQE and BRISQUE are obviously superior to the methods (SFF, ISF, QILV) with reference when performing on simple (see the 1st - 3rd groups of color-bars in Fig. 12 (a) ). Due to the unilateral measurement of structures, the height order of bars shows full-reference metrics have converse results to User Study as the 4th - 5th and 8th group of color-bars shown in Fig. 12 (a) - (c).

For the medium and complex texture filtered images, no-reference quality assessment algorithms are almost
difficult to obtain the effective measurement (see the 1st - 3rd groups of color-bars Fig. 12 (b) & (c)). Conversely, the seventh and eighth groups of bars in Fig. 12 (a) - (c) show that, no matter what scale texture images, our T3SI is close to the results of User Study for evaluating texture filtered images.

3) PERFORMANCE ON NOISY IMAGES

In order to perform T3SI on noisy images, we consider the noisy image as the reference for calculating EPI and the clean image as the reference input of SSIM. Choosing five images with a large flat region from the LIVE database [39], [40], we perform filtering methods on the noisy images and obtain denoised image as shown in Fig. 13 (c) - (e). When we regard the noise in flat areas as the texture, the cleanliness of flat areas is similar to the texture smoothness. Hence, we can select some patches inside flat regions to measure the removal of noise, and choose some patches containing large edge features to evaluate the retention of structures. Based on the attribute of the two measurements, T3SI could be used to evaluate the noisy image. We compared our T3SI with the metrics: NIQE [15], PIQE [17], BRISQUE [18], SFF [32], IFS [33], QILV [31], PSNR [23] and SSIM [11]. The results of evaluation are detailed in Tab. 3. Since PSNR and SSIM are generally used to evaluate denoised images quality, the results of other metrics are compared with that of them. We could observe from PSNR and SSIM that the three denoised results of each noisy image show: (e) is the best, (d) is the second, and (c) is the worst. Compared these metrics, the full-reference methods SFF [32], IFS [33], QILV [31] and our approach for

![FIGURE 11. A complex texture reference image and its filtered test images. (a) Reference image. (b)-(f) Five filtered test images #21-#25. The area in each yellow box represents the texture region, and the area in red each box contains typical structure details. The two corresponding local areas are zoomed on the right of each image.](image1)

![FIGURE 12. Comparisons between various assessment methods. Each group of bars represents the evaluation results of five filtered images by one evaluation method.](image2)
FIGURE 13. Noisy images evaluation of the test set (Womanhat, Sailing, Plane, Lighthouse, Carnivaldolls). (a) Original images. (b) Noisy images. (c) - (e) Denoised images by using different filtering methods. The areas in red boxes are selected to measure the removal of noise, and the areas in green boxes are used to detect the structural fidelity.

TABLE 3. The measurement for the output of each noisy image (see Fig. 13 (c) - (e)). The data in the table is the assessment values of various metrics: NIQE [15], PIQE [17], BRISQUE [18], SFF [32], IFS [33], QILV [31], PSNR [23], SSIM [11] and T3SI.

| Image     | Denoising | NIQE | PIQE | BRISQUE | SFF | IFS | QILV | PSNR | SSIM | T3SI |
|-----------|-----------|------|------|---------|-----|-----|------|------|------|------|
| Womanhat  | Output#1  | 7.5031 | 8.7084 | 18.3352 | 0.9565 | 0.9093 | 0.9935 | 27.8594 | 0.9177 | 1.9113 |
|           | Output#2  | 3.3478 | 13.1328 | 33.0687 | 0.9524 | 0.9103 | 0.9946 | 27.9111 | 0.9382 | 2.2407 |
|           | Output#3  | 2.2940 | 37.6532 | 23.5756 | 0.9688 | 0.9355 | 0.9972 | 31.3222 | 0.9700 | 2.2846 |
| Lighthouse| Output#1  | 6.2064 | 10.8026 | 32.0661 | 0.9523 | 0.8977 | 0.9932 | 25.8620 | 0.8089 | 1.9145 |
|           | Output#2  | 3.7559 | 39.9249 | 34.7371 | 0.9657 | 0.9140 | 0.9938 | 25.8821 | 0.8510 | 2.2052 |
|           | Output#3  | 2.1612 | 32.1436 | 19.4153 | 0.9748 | 0.9395 | 0.9976 | 30.0393 | 0.9207 | 2.2663 |
| Carnivaldolls| Output#1 | 6.7806 | 17.4329 | 32.7559 | 0.9655 | 0.9135 | 0.9966 | 25.7291 | 0.7715 | 1.8658 |
|           | Output#2  | 4.6225 | 41.1823 | 33.6745 | 0.9736 | 0.9265 | 0.9941 | 25.0080 | 0.8536 | 2.0963 |
|           | Output#3  | 2.6987 | 43.4437 | 23.8822 | 0.9781 | 0.9479 | 0.9988 | 30.1717 | 0.9435 | 2.1114 |
| Sailing   | Output#1  | 7.8553 | 10.4716 | 34.4375 | 0.9546 | 0.9081 | 0.9907 | 27.6131 | 0.7150 | 1.9132 |
|           | Output#2  | 3.2822 | 46.5387 | 34.3888 | 0.9627 | 0.9194 | 0.9945 | 28.2844 | 0.8319 | 2.2159 |
|           | Output#3  | 2.9737 | 45.4075 | 21.9092 | 0.9776 | 0.9498 | 0.9971 | 32.8416 | 0.9160 | 2.2869 |
| Plane     | Output#1  | 7.6628 | 10.3104 | 26.9057 | 0.9654 | 0.9224 | 0.9988 | 28.3226 | 0.8399 | 1.8592 |
|           | Output#2  | 3.6368 | 33.8626 | 36.0644 | 0.9733 | 0.9222 | 0.9915 | 26.5656 | 0.8837 | 2.0411 |
|           | Output#3  | 2.2882 | 41.1789 | 20.6271 | 0.9823 | 0.9503 | 0.9998 | 32.6353 | 0.9485 | 2.2283 |

the same noisy image are consistent with PSNR/SSIM, while the no-reference methods perform poorly. Actually, we could intuitively find that the denoising effects in (d) & (e) are obviously superior to that of in (c). Our T3SI is in line with subjective evaluation, and experimental results demonstrate that our method has a good subjective applicability for IQA.
VI. CONCLUSION

We present an interactive assessment method T3SI for measuring texture filtered images that are difficult to evaluate by existing IQA metrics. T3SI constructed by using an exponential cross entropy formulation can comprehensively evaluate the texture smoothness and the structure retention, which could make objective evaluation consistent with subjective practical applications. Especially, the local patches selected by human perception has advantage of the accurate identification of textures, by which T3SI can achieve the unity of the local and overall assessment, the objective and subjective evaluation. T3SI is significant for image texture quality assessment, and numerous experiments demonstrate that T3SI is stable, comprehensive and quantitative in IQQA.

In our future work, we attempt to improve the adaptive selection of patches according to different texture patterns, such that our proposed T3SI can be used to assess a mass of texture filtered images.

REFERENCES

[1] T. F. Chan, S. H. Kang, and J. Shen, “Total variation denoising and enhancement of color images based on the CB and HSV color models,” J. Vis. Commun. Image Represent., vol. 12, no. 4, pp. 422–435, Dec. 2001.
[2] S. S. Yaghoubi, M. A. Couto, C. C. Chen, L. Polavaram, and S. S. Ghanbari, “Preclinical safety evaluation of 18F-FHBG: A pet reporter probe for imaging herpes simplex virus type I thymidine kinase (HSV1-tk) or mutant HSV1-sr39tk’s expression,” J. Nuclear Med., vol. 47, no. 4, pp. 706–715, 2006.
[3] H. R. Sheikh, A. C. Bovik, and G. de Veciana, “An information fidelity criterion for image quality assessment using natural scene statistics,” IEEE Trans. Image Process., vol. 14, no. 12, pp. 2117–2128, Dec. 2005.
[4] Q. Wu, Z. Wang, and H. Li, “A highly efficient method for blind image quality assessment,” in Proc. IEEE Int. Conf. Image Process. (ICIP), 2015, pp. 339–343.
[5] F. Sattar, L. Floreby, G. Salomonsson, and B. Lovstrom, “Image enhancement based on a nonlinear multiscale method,” IEEE Trans. Image Process., vol. 6, no. 6, pp. 888–895, Jun. 1997.
[6] A. K. Moorthy, C.-C. Su, A. Mittal, and A. C. Bovik, “Subjective evaluation of stereoscopic image quality,” Signal Process., Image Commun., vol. 28, no. 8, pp. 870–883, Sep. 2013.
[7] J. Wang, K. Zeng, and Z. Wang, “Quality prediction of asymmetrically distorted stereoscopic images from single views,” in Proc. IEEE Int. Conf. Multimedia Expo (ICME), Jul. 2014, pp. 1–6.
[8] A. Rubel, V. Lukin, M. Uss, B. Vozel, O. Pogrebnyak, and K. Egiazarian, “Efficiency of texture image enhancement by DCT-based filtering,” Neurocomputing, vol. 175, pp. 948–965, Jan. 2016.
[9] H. Tang and L. W. Cahill, “A new criterion for the evaluation of image restoration quality,” in Proc. 21st Technol. Enabling Tomorrow, Comput., Commun. Automat. (TENCON), 1992, pp. 573–577.
[10] N. Liu and G. Zhai, “Free energy adjusted peak signal to noise ratio (FEA-PSNR) for image quality assessment,” Sens. Imaging, vol. 18, no. 1, p. 11, Dec. 2017.
[11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.
[12] F. Chen, X. Huang, and W. Chen, “Texture-preserving image deblurring,” IEEE Signal Process. Lett., vol. 17, no. 12, pp. 1018–1021, Dec. 2010.
[13] S. Tebini, H. Seddik, and E. Ben Braiek, “Efficient image texture and structure preservation via a rapid diffusion function with accelerator,” in Proc. 4th Int. Conf. Adv. Technol. Signal Image Process. (ASTISP), Mar. 2018, pp. 1–6.
[14] C. Song, C. Xiao, X. Li, J. Li, and H. Sui, “Structure-preserving texture filtering for adaptive image smoothing,” J. Vis. Lang. Comput., vol. 45, pp. 17–23, Apr. 2018.
[15] 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.
[16] B. Appina, S. Khan, and S. S. Channappayya, “No-reference stereoscopic image quality assessment using natural scene statistics,” Signal Process., Image Commun., vol. 43, pp. 1–14, Apr. 2016.
[17] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani, “Blind image quality evaluation using perception based features,” in Proc. 21st Stereoscopic Displays and Applications, Feb. 2010, Art. no. 75240T.
[18] M.-J. Chen, L. K. Shih, and A. C. Bovik, “No-reference quality assessment of natural stereopairs,” IEEE Trans. Image Process., vol. 22, no. 9, pp. 3379–3391, Sep. 2013.
[19] F. Shao, K. Li, W. Lin, G. Jiang, and Q. Dai, “Learning blind quality evaluator for stereoscopic images using joint sparse representation,” IEEE Trans. Multimedia, vol. 18, no. 10, pp. 2104–2114, Oct. 2016.
[20] Q. Jiang, F. Shao, W. Gao, Z. Chen, G. Jiang, and Y.-S. Ho, “Unified no-reference quality assessment of singly and multiply distorted stereoscopic images,” IEEE Trans. Image Process., vol. 28, no. 4, pp. 1866–1881, Apr. 2019.
[21] A. Tanchenko, “Visual-PSNR measure of image quality,” J. Vis. Commun. Image Represent., vol. 25, no. 5, pp. 874–878, Jul. 2014.
[22] A. M. Eskicioglu and P. S. Fisher, “Image quality measures and their performance,” IEEE Trans. Image Commun., vol. 43, no. 12, pp. 2959–2965, Dec. 1995.
[23] Q. Huynh-Thu and M. Ghanbari, “Scope of validity of PSNR in image/video quality assessment,” Electron. Lett., vol. 44, no. 13, p. 800, 2008.
[24] W. Xue, X. Mou, L. Zhang, and X. Feng, “Perceptual fidelity aware mean squared error,” in Proc. IEEE Int. Conf. Comput. Vis., Dec. 2013, pp. 705–712.
[25] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, and W. Zhang, “Content-weighted mean-squared error for quality assessment of compressed images,” Signal Image Video Process., vol. 10, no. 5, pp. 803–810, Jul. 2016.
[26] Z. Wang, E. P. Simoncelli, and A. C. Bovik, “Multiscale structural similarity for image quality assessment,” in Proc. 37th Asilomar Conf. Signals, Syst. Comput., 2003, pp. 1398–1402.
[27] Z. Wang and Q. Li, “Information content weighting for perceptual image quality assessment,” IEEE Trans. Image Process., vol. 20, no. 5, pp. 1185–1198, May 2011.
[28] L. Zhang, L. Zhang, and X. Mou, “RFSIM: A feature based image quality assessment metric using Riesz transforms,” in Proc. IEEE Int. Conf. Image Process., Sep. 2010, pp. 321–324.
[29] S. Aja-Fernandez, R. S. J. Estepar, C. Alberola-Lopez, and C.-F. Westin, “Image quality assessment based on local variance,” in Proc. Int. Conf. IEEE Eng. Med. Biol. Soc., Aug./Sep. 2006, no. 1, pp. 4815–4818.
[30] H. H. Chang, H. Yang, Y. Gan, and M. H. Wang, “Sparse feature fidelity for perceptual image quality evaluation,” IEEE Trans. Image Process., vol. 22, no. 10, pp. 4007–4018, Jan. 2013.
[31] H.-W. Chang, Q.-W. Zhang, Q.-G. Wu, and Y. Gan, “Perceptual image quality assessment by independent feature detector,” Neurocomputing, vol. 151, pp. 1142–1152, Mar. 2015.
[32] Q. Zhang, X. Shen, L. Xu, and J. Jia, “Rolling guidance filter,” in Proc. Eur. Conf. Comput. Vis., vol. 5691, 2014, pp. 815–830.
[33] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in Proc. 6th Int. Conf. Comput. Vis., 1998, pp. 839–846.
[34] L. Xu, Q. Yan, Y. Xia, and J. Jia, “Structure extraction from texture via relative total variation,” ACM Trans. Graph., vol. 31, no. 6, pp. 1–10, Nov. 2012.
[35] H. Cho, H. Lee, H. Kang, S. Lee, “Bilateral texture filtering,” ACM Trans. Graph., vol. 33, no. 4, pp. 128:1–128:14, 2014.
[36] B. Cai, X. Xing, and X. Xu, “Edge/structure preserving smoothing via relative-to-Gaussian,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2017, pp. 250–254.
[37] L. C. H. R. Sheikh, Z. Wang, and A. Bovik, Live Image Quality Assessment Database Release 2. [Online]. Available: http://live.ece.utexas.edu/research/quality
[38] 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.
CHONG LIU received the bachelor’s degree in information and computing from the Harbin University of Technology, in 2002. He is currently pursuing the Ph.D. degree with the Nanjing University of Aeronautics and Astronautics, China. Since 2005, he has been a Lecturer with Anqing Normal University. His research interests include numerical computation, image processing, and machine learning.

CUI YANG received the bachelor’s degree in mathematics and computing science from Anqing Normal University, in 2002, and the master’s degree in applied mathematics from the Beijing University of Chemical Technology, in 2005. Since 2005, she has been a Lecturer with Anqing Normal University. Her research interests include numerical computation and machine learning.

MINGQIANG WEI received the Ph.D. degree in computer science and engineering from The Chinese University of Hong Kong (CUHK), in 2014. He is currently an Associate Professor with the School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics (NUAA). His research interests focus on computer vision, computer graphics, and computer-aided art design. He received the CUHK Young Scholar Thesis Awards 2014.

JUN WANG received the bachelor’s and Ph.D. degrees in computer-aided design from the Nanjing University of Aeronautics and Astronautics (NUAA), China, in 2002 and 2007, respectively. From 2010 to 2013, he worked as a Senior Research Engineer at Leica Geosystems Inc., USA. He is currently a Professor with NUAA. His research interests include geometry processing and geometric modeling, especially large-scale LiDAR point data capturing, management, processing, and analysis.