This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

Reference-free Quality Assessment of Sonar Images via Contour Degradation Measurement

Weiling Chen, Ke Gu, Weisi Lin, *Fellow, IEEE,* Zhifang Xia, Patrick Le Callet, *Fellow, IEEE,* and En Cheng

***Abstract*—Sonar imagery plays a significant role in oceanic applications since there is little natural light underwater, and light is irrelevant to sonar imaging. Sonar images are very likely to be affected by various distortions during the process of transmission via the underwater acoustic channel for further analysis. At the receiving end, the reference image is unavailable due to the complex and changing underwater environment and our unfamiliarity with it. To the best of our knowledge, one of the important usages of sonar images is target recognition on the basis of contour information. The contour degradation degree for a sonar image is relevant to the distortions contained in it. To this end, we developed a new no-reference contour degradation measurement (NRCDM) for perceiving the quality of sonar images. The sparsities of a series of transform coefficient matrices, which are descriptive of contour information, are first extracted as features from the frequency and spatial domains. The contour degradation degree for a sonar image is then measured by calculating the ratios of extracted features before and after filtering this sonar image. Finally, a bootstrap aggregating (bagging)-based support vector regression (SVR) module is learned to capture the relationship between the contour degradation degree and the sonar image quality. The results of experiments validate that the proposed metric is competitive with state-of-the-art reference-based quality metrics and outperforms the latest reference-free competitors.**

***Index Terms*—Image quality assessment (IQA), reference-free, sonar image, degradation measurement, bagging**

I. Introduction

With the expanding scale of ocean-related technology development, oceanic applications, such as ocean exploration, underwater navigation, and underwater acoustic communication, are growing rapidly. As an important carrier of oceanic information, sonar images can reflect underwater scenes in darkness because sonar imaging is independent of illumination. Both the process of acquisition and transmission may introduce distortion into sonar images. In the acquisition process, sonar images may contain radiometric distortion, speckle noise, and other typical artifacts due to the propagation characteristics of the acoustic field and the unpredictable underwater environment. Sonar images are often captured underwater but then analyzed by a human in a remote location above water because of the limitation of the processing capacity of underwater equipment. The transmission of sonar images via underwater acoustic channels is very likely to contain distortions since the condition of the underwater acoustic channel is much worse than that of the terrestrial channel. In practice, distorted sonar images can truly affect the behavior of their applications. Under these circumstances, image quality assessment (IQA) plays an essential role in quality control and service monitoring. In this paper, we mainly discuss the sonar IQA metric, aiming at distortions that are caused in the transmission process and assuming a reasonable quality of the sonar images selected for transmission.

Diversified contents are included in the field of IQA methods designed for natural scene images (NSIs). Some of these methods are dependent on reference information [1]-[3] and evaluate image quality according to the measurement of fidelity or distortion. Another important type of IQA method does not need a reference at all. Most of these methods tend to establish links between image characteristics (e.g., natural scene statistic (NSS) model) and image quality that rely on a learning-based predictive model [4]-[12]. In practice, NSIs are usually captured for the general interpretation or representation of a visual scene, and thus, most viewers of NSIs are people without any expertise. They can judge the perceptual quality of a NSI according to their common sense and aesthetics. The perceptual quality of a NSI is associated with the perceptual similarity of the test image to a reference (the reference may be either an external or an internal reference). It can be influenced by contrast, color saturation, sharpness, and many other image components. In contrast, the viewers of sonar images are usually experts since sonar images all have specific applications. Among these applications, detection and recognition are mostly concerned in this paper because of their wide utilization in different sonar image-related scenarios. Considering sonar image applications, utility quality is perceived by viewers of sonar images. The evaluation of the utility quality of a sonar image is determined by its performance in specific applications and cannot be replaced by perceptual quality. Specifically, sonar images with a clear and complete macroscopic structure perform better in detection and recognition. Most NSIs are

Manuscript received February 24, 2018; revised November 23, 2018; accepted March 27, 2019. This work was supported in part by the National Science Foundation of China under Grant 61703009, 61871336, the Young Elite Scientist Sponsorship Program by China Association for Science and Technology under Grant 2017QNRC001, and the Young Top-Notch Talents Team Program of Beijing Excellent Talents Funding under Grant 2017000026833ZK40. (Corresponding author: Ke Gu.)

W. Chen is with the College of Physics and Information Engineering, Fuzhou University, China (email: weiling.chen@fzu.edu.cn).

K. Gu is with the Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China (email: guke.doctor@gmail.com).

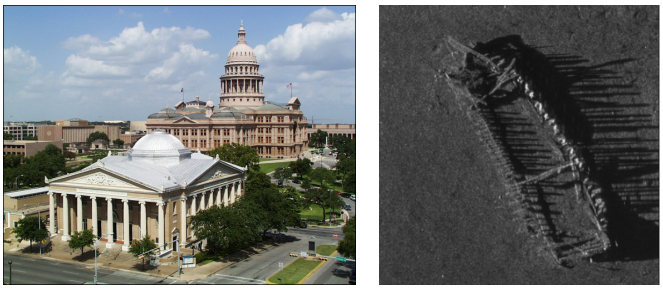
W. Lin is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore, 639798 (email: wslin@ntu.edu.sg).

Z. Xia is with the State Information Center of P.R.China, Beijing, China, and also with Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China (email: spidergirl21@163.com).

P. Le Callet is with Luman Université, Université de Nantes, IRCCyN UMR CNRS 6597, Polytech Nantes, France (email: patrick.lecallet@univ-nantes.fr).

E. Chen is with Key Laboratory of Underwater Acoustic Communication and Marine Information Technology (Xiamen University), Ministry of Education, Xiamen University, China, 361005 (e-mail: chengen@xmu.edu.cn).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing



(a) (b)

Fig. 1. Examples of NSI and sonar image. These images are substantially different, and thus, new models for IQA are required for sonar images. (a) A NSI captured by a camera; (b) A sonar image captured by side-scan sonar.

created using a camera, which uses a lens to focus the scene's visible wavelengths of light onto a copy of what the human eye would see. For a sonar image, it is captured by sonar, which emits pulses of sounds and listens for echoes; the sonar image is then formed according to the arrival time and intensity of echoes. In addition, the characteristics of sonar images and NSIs are different. An example of a NSI taken by a digital camera and a sonar image captured by side-scan sonar is shown in Fig. 1(a) (b). This example depicts the characteristics of most NSIs and sonar images, that is, 1) NSIs are usually colorful and have rich color variations, thick lines, more details and complex texture content, and 2) sonar images are mostly characterized by small variation of pixel values, low contrast, less details and gray levels [13]-[14]. Based on the above discussion, it can be inferred that the existing IQA methods designed for NSIs are not suitable for sonar images. A performance comparison of 9 existing classical and state-of-the-art reference-based IQA models was conducted on the sonar image quality database (SIQD) in [15]. The experimental results illustrate that despite the superior performance of the aforementioned reference-based IQA models on NSIs, they fail on sonar IQA. To the best of our knowledge, the research on sonar IQA is very limited. Moreover, almost no reference-free IQA method designed for acoustic lens and side-scan sonar images currently exists, such as the approach used in this paper. In many underwater applications, such as underwater acoustic transmission and transmitted image restoration, a reference image is unavailable. Therefore, it is necessary to establish a reference-free sonar IQA method.

With the goal of solving the challenging problem above, we propose a no-reference contour degradation measurement (NRCDM) for sonar image quality assessment in this paper. One can deduce whether there is a target or what the terrain is from the macroscopic structure of a sonar image. Contour is an important form of macroscopic structure. Therefore, the integrality of the contour to a greater extent determines the utility quality of a sonar image [16]-[18]. From another perspective, the quality of an image is dependent on the response of the human visual system (HVS) to different frequency components [19]-[21]. The contour information, which is important for sonar image applications, makes up the intermediate frequency components, while most distortion types are relevant to high-frequency or low-frequency components. According to reference [22], the addition of low-frequency distortion generally slows the contour degradation degree, and vice versa. Since the sparsity of different transform coefficient matrices can represent the energy of macroscopic structures of an image [23]-[26], the sparsities of three types of transform coefficient matrices are extracted as features in this paper. Then, the test image is first filtered by a degradation model, and the same features are extracted from the filtered version of the test image. This kind of framework is employed according to [22], [27], which demonstrated that the addition of distortion has a real impact on the contour degradation degree. Finally, the ratios of features before and after degradation are converted into a quality score by a **b**ootstrap **agg**regat**ing** (bagging)-based support vector regression (SVR) module, which can avoid the unsteadiness and overfitting caused by a small-sized training set [28].

The major contributions of this work are summarized as three points. As one of the first teams dedicated to sonar IQA, we conduct many experiments and have conclude that existing quality assessment methods designed for NSIs or screen content images are not suitable for sonar images. Second, we provide a reference-free unified framework for sonar image quality assessment, which is the first no-reference quality evaluation algorithm for acoustic lens sonar and side-scan sonar images. Third, we show that macroscopic structure is important for relevant types of sonar images via experiments and analysis and design corresponding quality assessment methods based on this conclusion. The experimental results show that the proposed work has better performance, robustness, and efficiency compared to the existing methods.

We arrange the rest of this paper as follows. Section II provides a review for IQA, and the description of sonar images, including the introduction of the SIQD database and a straightforward illustration of the differences between NSIs and sonar images. Details about the feature extraction and the bagging-based SVR module are presented in Section III. In Section IV, performance comparisons of the proposed NRCDM metric with the existing state-of-the-art reference-based and reference-free quality models are evaluated using the SIQD database to validate the superior performance of the NRCDM metric. The motivation, novelty and some important information about this paper are summarized in Section V. This paper is concluded in Section VI.

II. Related Work

A. Previous work

Since there are many mature works about reference-free IQA algorithms, the previous work closely related to this paper is briefly reviewed in this section. Most sonar images are in grayscale, and thus, color-to-gray is not a required step in the sonar IQA process [29]. Furthermore, it has been proven in [16] and [30] that image structure is adequate to correctly evaluate image utility. In addition, image structure is also one of the most important image components. Over the past decades, structure-based IQA algorithms have become

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

an active research topic. The gradient similarity (GSM) [3] is one of the most typical algorithms that evaluates structure and contrast change between reference and distorted images by gradient similarity under the framework of the popular structural similarity index (SSIM) [1]. In the gradient magnitude similarity deviation (GMSD) [31], the perceptual image quality is predicted by gradient magnitude similarity combined with the standard deviation of the gradient magnitude similarity map. GSM and GMSD are implemented on the basis of the gradient-relevant map, which depicts both macroscopic structure (such as image contour) and microscopic structure (such as image details and delicate textures). For sonar images, macroscopic structure plays a main role in its applications. Furthermore, Rouse has made efforts to perceive the image quality utilizing dominant structures. Rouse [16] exploited both single- and multiscale contour degradation information to accurately estimate utility quality. Though this work performs well for NSIs, it is designed in the full-reference mechanism. Considering the applications of sonar images, the reference-free methodology may be more suitable.

Without access to any reference information, researchers usually transform IQA to a regression problem, where the regressor is trained to establish the link between NSS- and/or quality-relevant features and image quality. Based on the features, previous reference-free work generally follows one of two trends: 1) NSS-based approaches and 2) signal-based approaches. NSS features have been commonly used for IQA in recent years. NSS features have been extracted from the wavelet domain [32]-[33], the discrete cosine transform (DCT) domain [11] and the spatial domain [40] and are assumed to contain certain statistical properties, which may be affected by the presence of distortion. Since the existing NSS models designed for NSIs are not suitable for sonar images, which will be verified subsequently, we would prefer a signal based method. The signal-based approaches usually rely on features that can represent the quality of images and a quality prediction function. The feature-based approaches can be classified into three main categories according to the regression methods selected. The first category is composed of codebook-based methods [8], which encode the characteristics of images using a trained codebook, and the relationship between codewords and image qualities are established. The second category is neural network-based methods [7], [9]-[10] that use neural networks to learn the image quality from the feature space. The last category employs shallow learner-based methods [27], [34]. These methods are all dependent on a vast amount of reliable training data to avoid overfitting and to cover corresponding quality factors as much as possible.

The proposed work differs from previous related works in three aspects. First, the proposed work focuses on the statistical characteristics of macroscopic structure instead of describing sketches, color and details, as most previous works do. Due to the limited resolution of sonar images, macroscopic structure, such as contour, is the most important element for image understanding. In the proposed work, the extraction of the main structure (i.e., macroscopic structure) will avoid the interference caused by details, such as texture. Second, a guided image filter [35] is employed as a degradation model in this paper, and Gu *et al.* has proposed a similar framework by utilizing a degradation model combining both an autoregressive model and a bilateral filter in [27]. Compared with this work, the guided image filter has a better edge-preserving ability and is computationally simpler. Third, since it is not easy to access a vast amount of sonar images, the scale of sonar image databases is generally not large. Commonly used learning machines as mentioned above, such as neural networks and probabilistic models, require a massive amount of training data to maintain the balance between bias and variance. To address this problem, ensemble learning is employed. We utilize a bagging-based SVR module to improve the generalization ability of the model by bootstrap sampling.

B. The description of sonar images

1) Image Database: To our knowledge, there is currently no a standard quality database for sonar images. We have established the SIQD database, which is made up of 840 sonar images and their mean opinion scores (MOSs) of utility quality [15]. Unlike the database of NSIs, subjective viewers of the SIQD database should be experts. That is, viewers of the SIQD database should be experienced in work related to underwater detection and recognition The difference between experts and nonexperts is that experts have prior knowledge about the possible content and applications of corresponding sonar images. It is especially critical for the assessment of utility quality since the utility quality should be considered under specific applications, which are unfamiliar to nonexperts. In contrast, the perceptual quality is independent of specific applications and professional backgrounds. It is usually evaluated according to life experiences and aesthetic standards. In the subjective testing process, viewers are asked to evaluate the quality of the test image by considering whether it will perform well in practical applications. In the SIQD database, the applications mainly refer to object recognition and detection.

Forty reference sonar images are included in the SIQD database. They are collected by side-scan sonar and acoustic lens sonar, both of which provide relatively high resolution and are applied widely in image acquisition. There are 800 distorted sonar images in the SIQD database, and four categories of distortions are covered in the SIQD database. They are distortions made by ComGBR coding [36] and SPIHT coding [37], and manmade bit errors in bitstreams of SPIHT coding and ComGBR coding, where each has four to five distortion levels. Due to the narrow bandwidth and complicated condition of underwater acoustic channels, the good robustness provided by ComGBR coding and the high compressibility provided by SPIHT coding meet different underwater acoustic transmission requirements [38]. In addition, the bit error ratios (BERs) used in the SIQD database are selected according to recent achievements in underwater acoustic communication [39]. As illustrated in Fig. 2, ComGBR coding and SPIHT coding both introduce some kinds of blur into the image, which are marked as ‘CC’ and ‘SC’, and the results of bit errors in the ComGBR coding stream and SPIHT coding stream are noise and eccentricity, respectively, which are indicated by ‘TC’ and ‘TS’. More details about the SIQD database are tabulated in Table I.

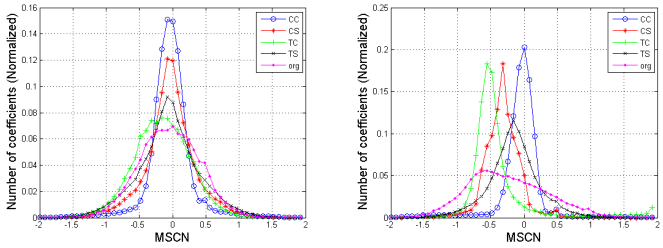
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

(a) (b) (c) (d) (e)

Fig. 2. Examples of sonar images in the SIQD database, which demonstrate the typical distortion types contained in the SIQD database. (a) Reference image; (b) Distorted sonar image with 'CC', MOS=35.15; (c) Distorted sonar image with 'CS', MOS=32.34; (d) Distorted sonar image with 'TC', MOS=34.21; and (e) Distorted sonar image with 'TS', MOS=35.01.

TABLE I
DETAILS ABOUT THE SIQD DATABASE.

Characteristics	Information
Number of reference images	40
Number of distorted images	800
Number of different types of distortions	4
Image resolution	320×320
Rating scales	5-category discrete scale
Evaluation method	Single stimulus with multiple repetition
Number of subjective viewers per image	25



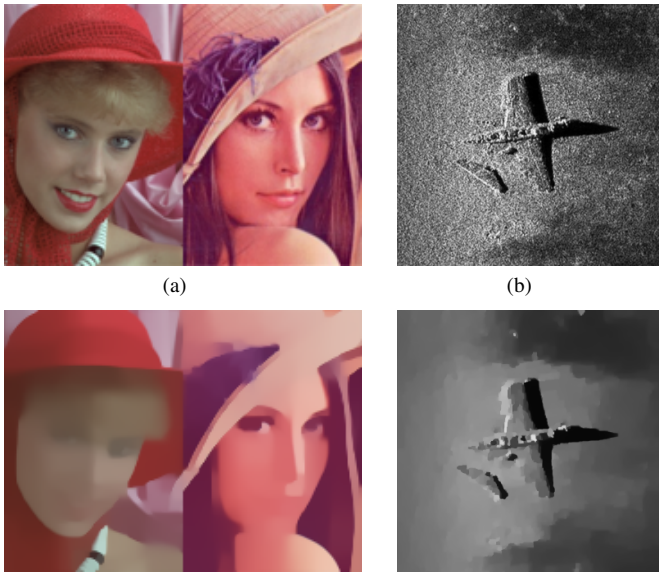
(a) (b)

Fig. 3. The histogram of MSCN coefficients of original and distorted versions of (a) Fig. 1(a) and (b) Fig. 1(b), which describe the differences between NSIs and sonar images statistically.

2) Differences between NSIs and sonar images: Fig. 1 shows that NSIs and sonar images are different from the visual perspective. Furthermore, as mentioned above, NSS models are inadequate compared to sonar images, although the scenes that sonar images reflect are also the results of a natural source. To illustrate it clearly, the mean subtracted contrast normalized (MSCN) coefficients are employed to reflect NSS characteristics. The statistical properties of MSCN coefficients have been widely used to establish a NSS model for NSIs since they vary with the existence of distortions contained in the image. In particular, the MSCN coefficients of original NSIs exhibit a Gaussian-like appearance, and the border spectrum of distorted NSI statistics can be captured by a generalized Gaussian distribution (GGD) [40], as Fig. 3(a) shows. Fig. 3(a) depicts the histogram of MSCN coefficients of Fig. 1(a), which is a NSI and its distorted versions. The distortions in this example are in the four categories contained in the SIQD database, as mentioned in Section II-B1, i.e., 'CC', 'CS', 'TC', and 'TS'. Then, we plot the histogram of MSCN coefficients of a sonar image with the aforementioned four distortions in Fig. 3(b). However, as shown in Fig. 3(b), the original sonar image exhibits a Rayleigh-like MSCN distribution, while the 'TS' creates a Laplacian distribution, and 'CC', 'CS', and 'TC' make the histogram of MSCN coefficients similar to an asymmetric generalized Gaussian distribution (AGGD). Thus, the NSS models for NSIs are not fully applicable to the quality evaluation of sonar images, and the features designed for NSIs might be unsuitable for sonar images.

From the application perspective, perceptual quality is important for NSIs, while utility quality is considered more for sonar images. NSIs are primarily for human consumption, where the individual differences in aesthetic standards will influence the quality evaluation results to a certain extent. All of these influences will be taken into account in the feature extraction and training process of NSI quality assessment. However, most sonar images are used for practical uses, for example, rescuing, underwater searching, underwater biological detection, and seabed mapping. Most of these applications are relevant to recognition and detection, which are also the key considerations of this paper. In the subjective quality assessment step of sonar images, the evaluation will not be affected by any aesthetic element; instead, viewers perceive the utility quality. The experimental results have shown that a perceptual quality score is not an alternative for a utility score. According to the evaluation criteria of utility quality mentioned previously, macroscopic structure plays a fundamental role in utility quality assessment based on recognition and detection. Furthermore, contour, as an important form of macroscopic structure, is important to the HVS of object perception[16], [17]-[18]. Accordingly, the microscopic structure, such as details and complex textures, is almost negligible in the utility quality assessment of sonar images. Most existing quality evaluation algorithms aim to generate scores for the perceptual quality assessment task. The macroscopic structure and microscopic structure are of equal importance to perceptual quality assessment. Hence, existing structure-based IQA methods take both microscopic and macroscopic structures into account. It might be one of the reasons why existing IQA models for NSIs fail in sonar IQA.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing



(a) (b) (c) (d)

Fig. 4. A sonar image and a NSI without detailed information, which shows the importance of macroscopic structure in a sonar image. (a) Original NSI "Woman"; (b) Original sonar image "Aircraft Wreckage"; (c) "Woman" without microscopic structures; and (d) "Aircraft Wreckage" without microscopic structures.

For comparison, we remove the detailed textures from a sonar image and a NSI by utilizing the structure extraction approach proposed in [41]. Fig. 4(a) and Fig. 4(b) are the original NSI and sonar image, respectively, and Fig. 4(c) and Fig. 4(d) are the macroscopic structures of Fig. 4(a) and Fig. 4(b), respectively, employing the same contour extraction preferences. There are two faces of women in Fig. 4(a), and aircraft wreckages are the main content of Fig. 4(b). When detailed structures are removed from a NSI, as Fig. 4(c) shows, it is obvious that only facial contours remain in the image. The important information, that is, the faces of women in this example, has been lost, which will significantly affect the perceptual quality of NSI "Woman". For the sonar image, the main object can be recognized from the remaining dominant structures. It can be concluded from Fig. 4 that macroscopic structures such as image contour provide the information for object identification, which is the important step in the abovementioned sonar image applications, and the number of details contained by sonar images are fewer than those contained in NSIs. Therefore, the contour contains the most useful information of a sonar image, while the useful information of a NSI can include color, contrast, brightness, texture, or other image components.

III. Methodology

A. Algorithm overview

Depending on the aforementioned differences between NSIs and sonar images, we noted that considerations of IQA for sonar images are different from considerations applied in NSIs. First, although the scenes reflected in sonar images are also a kind of natural scene, they cannot be described using existing NSS models. That is, NSS-based no-reference IQA methods do not work efficiently for sonar images. Due to the differences in visual characteristics, microscopic structures in sonar images are much fewer than those in NSIs. Hence, the influence of microscopic structures, such as details, thick lines and complex texture content, is very limited for sonar IQA. In contrast, they cannot be ignored regarding the perceptual quality evaluation for NSIs. For differences in applications, the utility quality is often discussed for sonar images. The sonar image utility quality in detection and recognition is highly relevant to macroscopic structures. However, the quality of NSIs generally refers to perceptual quality, which is evaluated from the view of both microscopic and macroscopic structures. To sum up, macroscopic structures are the main consideration of sonar IQA, while the microscopic structures can be neglected for a more efficient quality assessment. In this paper, the contour information, which is the main component of macroscopic structure, is extracted as the main feature indicating the quality of the sonar image.

To clearly describe the importance of macroscopic structures for sonar images, we have added a subjective test in the revised paper. In this test, 50 sonar images, which are not included in the SIQD database, are selected as test samples. Then, the macroscopic structures of these images are extracted according to the approach proposed in [41]. One original image and its macroscopic structure were displayed in the user surface at every turn. The viewers were asked to provide their personal opinions regarding whether they can recognize the target from the original image and the corresponding macroscopic structure. After providing their opinions on one image pair, the viewers were able to continue evaluating the next pair of images. For 50 image pairs, each viewer took approximately 10-20 minutes to accomplish the subjective testing. We collected the subjective opinions from 25 viewers. They all mastered the basic knowledge about target recognition from sonar images. The evaluation for each pair of images can be considered as an event. These events can be divided into four categories: (1) both the target in the original image and the macroscopic structure can be recognized, and the cost is indicated by C_{11}; (2) both the target in the original image and the macroscopic structure cannot be recognized, and the cost is indicated by C_{00}; (3) the target in the original image can be recognized, but the target in the macroscopic structure cannot be recognized, and the cost is represented by C_{10}; and (4) the target in the original image cannot be recognized, but the target in the macroscopic structure can be recognized, and the cost is represented by C_{01}. The hypothesis in this experiment is that the important information for target recognition remains in the macroscopic structure of a sonar image. Under the minimal error probability criterion ($C_{11} = C_{00} = 0$ and $C_{10} = C_{01} = 1$), the average cost for each event can be calculated as 1.93×10^{-4}, which is very close to 0. It is obvious that this result can support our hypothesis; thus, the proposed metric is designed based on the importance of macroscopic structure for sonar images.

Fig. 5 depicts a general framework of the NRCDM metric considering the utility of sonar images. The degradation model degrades image contour by filtering some specific frequency

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

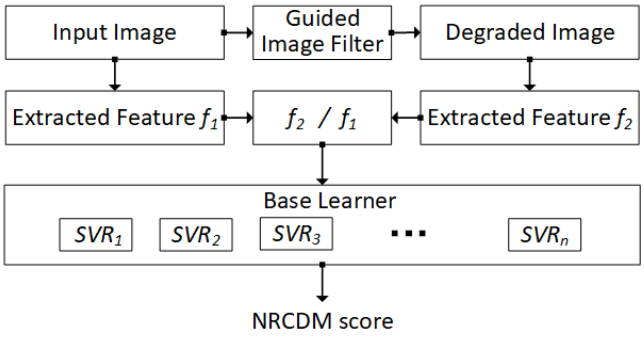


Fig. 5. A general framework of the proposed NRCDM metric.

components, and the addition of distortions will influence the effect of the degradation model [22]-[27]. This motivates us to explore the probability of perceiving distortion by measuring the degradation degree. We first extract features that reflect contour information, then measure the contour degradation degree by comparing features between the test image I and its filtered version I', and finally evaluate the quality of the test image using such a contour degradation degree in the quality evaluation module. According to this framework, in Section III, we describe the proposed methodology from three types of considerations: (1) the features are extracted in Section III-B; (2) in Section III-C, we introduce a degradation model, which is used to degrade test image, and then measure the contour degradation between the test image and its filtered version by comparing features extracted before and after degradation; and (3) in Section III-D, the bagging-based SVR module is trained to build the relationship between the contour degradation measured in Section III-C and the subjective quality values of each sonar image.

B. Feature extraction

The hypothesis that spatial frequency decomposition occurs in the HVS has been supported in many papers, i.e., the quality of an image is determined by the response of HVS to different frequency components [19]-[20]. Furthermore, such a premise has been successfully used in a simple pattern recognition system and a model of HVS based on spatial frequency [21]. Since the utility quality evaluated for sonar images in this paper is employed on the basis of recognition and detection, this premise can also be used here. Accordingly, different components in the image correspond to different frequencies. The image contour, whose importance to sonar image applications has been proven above, corresponds to the intermediate frequency components. The high-frequency and low-frequency components are made up of most distortion types. An example is illustrated in Fig. 6, where the distortions destroy the contour of the sonar image. It can be observed that distortions corresponding to the increase of high-frequency components make the frequency spectrum brighter. However, distortions that add the low-frequency components to the image make the frequency spectrum darker. The addition of high-frequency or low-frequency components has a substantial impact on the image. According to [22]-[27], high-frequency

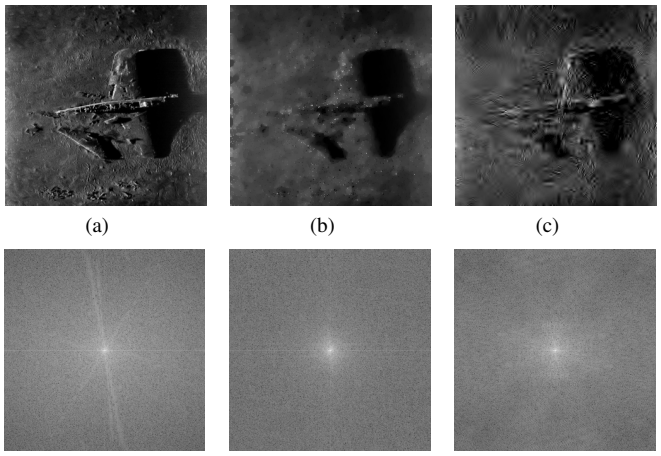


Fig. 6. Illustration of the relationship between the distortion and frequency component. The distortions destroy the contour of the sonar image and make the frequency spectrum become brighter or darker. (a) Reference image; (b) Distorted sonar image with 'CC'; (c) Distorted sonar image with 'TS'; and (d)-(f) Frequency spectrum of (a)-(c).

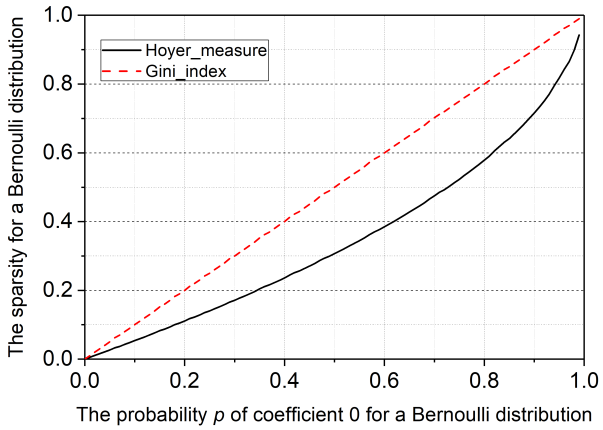


Fig. 7. Sparsity *vs* p for a Bernoulli distribution with coefficients equal to 0 with probability p or 1 with probability $1-p$.

or low-frequency components can change the degradation degree of the contour under the same degradation model. In detail, the contour of a sonar image with additional low-frequency distortion generally has a lower degradation speed, and vice versa. In this paper, we analyze the quality of sonar images by the contour degradation measurement.

We collect the contour information from the frequency domain by extracting features using the DCT and Cohen-Daubechies-Feauveau 9/7 wavelet transform (C-D-F 9/7). Since singular values of an image also contain the information of the main contour [26], the singular value decomposition (SVD) is employed to obtain image contour information from the spatial domain. The features selected for sonar images are extracted in the following way. Given a test sonar image I, we first transform it into the aforementioned domains. We denote D and C as coefficient matrices of the DCT and the C-D-F 9/7 wavelet transform and S as the diagonal matrix

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

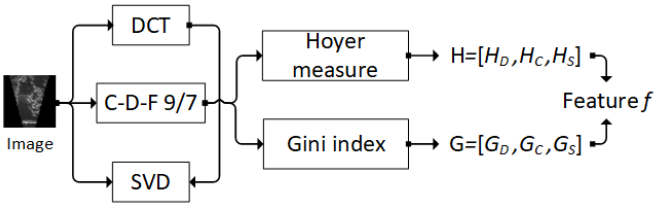


Fig. 8. The flow chart of the feature extraction process.

of the SVD transform. Coefficients with large values of D, C and S are representative of contour, which is the main part of the macroscopic structure. These three matrices are all sparse, and their sparsity depends on the coefficients with large values. Therefore, the sparsity of different transform coefficient matrices can be considered as the representation of contour information. Specifically, D, C and S will either become more sparse when low-frequency components are increased or less sparse when high-frequency components are added.

A sparsity measure should satisfy six criteria: *Scaling*, *Bill Gates*, *Rising Tide*, *Robin Hood*, *Babies*, and *Cloning*. When the number of coefficients is fixed, the Hoyer measure and Gini index satisfy all six criteria. Both the Gini index and Hoyer measure outperform other sparsity measures, such as l^0, $-l^1$, $\frac{l^2}{l^1}$ [23]. An experiment is conducted, wherein we test the sparsity of a set whose coefficients correspond to a Bernoulli distribution using the Gini index and Hoyer measure. The coefficient of this set is either 1 with probability $1-p$ or 0 with probability p. As Fig. 7 shows, the sparsity measure increases when p grows increasingly closer to 1, and we note that this is the case in both the Gini index and Hoyer measure. When p increases linearly, the Gini index also increases linearly; however, the sparsity increases faster when p is close to 1 for the Hoyer measure. This indicates that different sparsity measures have varying levels of sensitivity to sparsity, i.e., the sensitivity of different sparsity measures to contour distortion is different. Therefore, in this paper, we make use of both the Gini index and Hoyer measure [23] to more comprehensively describe the contour information. The Hoyer measure and Gini index are given by the following:

$$H(\mathbf{c}) = \frac{\sqrt{N} - (\sum_{i=1}^{N} \mid c_i \mid)/\sqrt{\sum_{i=1}^{N} c_i^2}}{\sqrt{N} - 1} \tag{1}$$

$$G(\mathbf{c}) = 1 - 2\sum_{k=1}^{N} \frac{c_{(k)}}{\parallel \mathbf{c} \parallel_1}\left(\frac{N-k+\frac{1}{2}}{N}\right) \quad for \quad c_{(1)} \le c_{(2)} \le \cdots \le c_{(N)} \tag{2}$$

where $\mathbf{c} = [c_1, c_2, \cdots, c_N]$ is a 1D vector that is converted from 2D transform coefficient matrix. We order $\mathbf{c}$ from smallest to largest, $c_{(1)} \le c_{(2)} \le \cdots \le c_{(N)}$, where $(1), (2), \cdots, (N)$ are the new indices after the sorting operation.

Fig. 8 shows the flow chart of the feature extraction process for sonar images, where the extracted features, $[H_D, H_C, H_S]$ and $[G_D, G_C, G_S]$, are the Hoyer measure and Gini index of D, C and S, respectively. Most of the contour information of a sonar image is contained in these features.

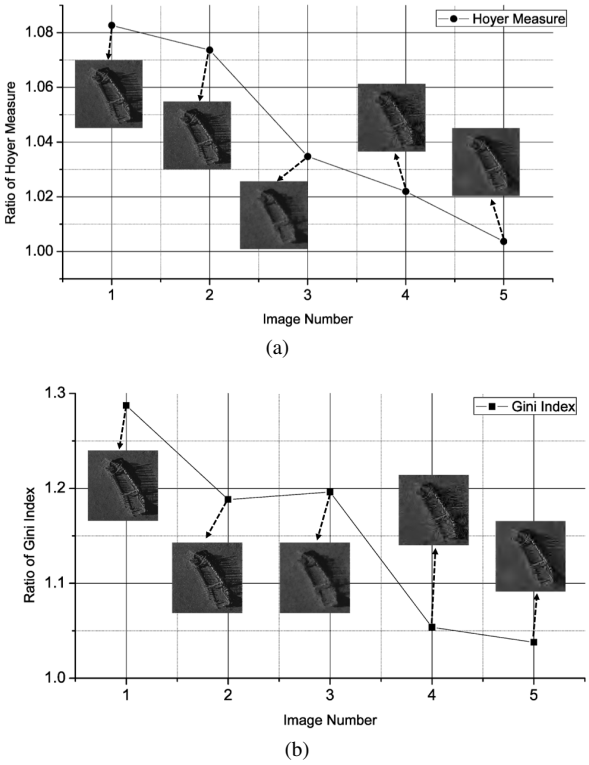


Fig. 9. The variation of (a) $\frac{H'_S}{H_S}$ and (b) $\frac{G'_S}{G_S}$ when distortion is added to the image (the vertical axes correspond to the ratio of (a) Hoyer measure and (b) Gini index of the diagonal matrix of the SVD transform before and after degradation, and the horizontal axes correspond to the image number.).

C. Contour degradation measurement

To measure the degradation degree, a degradation model is employed in this paper. The guided image filter [35], which degrades images by smoothing, is applied as the degradation model. The kernel weights of the guided image filter can be clearly expressed as follows:

$$W_{ij}(P) = \frac{1}{\mid \omega \mid^2} \sum_{k:(i,j)\in\omega_k} \left(1 + \frac{(P_i - \mu_k)(P_j - \mu_k)}{\sigma_k^2 + \epsilon}\right) \tag{3}$$

where ω_k is a window, $\mid \omega \mid$ is the number of pixels in ω_k, μ_k and σ_k^2 are the mean and variance of the guidance image P in ω_k, respectively, and ϵ is a regularization parameter that determines whether the current pixel should be the average of the pixels nearby or preserved. The degradation intensity of the guided image filter can be controlled by the regularization term ϵ, which has a positive correlation with $[H_D, H_C, H_S]$ and $[G_D, G_C, G_S]$. Through the theoretical analysis and experimental check, it can be observed that the sparsity of different transform coefficient matrices increases as ϵ grows, i.e., the sparsity of different transform coefficient matrices is relevant to the degradation caused by the guided image filter.

The guided image filter should be fixed so that the relationship between the degradation degree and distortions contained in the test image (or qualities of the test images) can be built by a regression model. In this paper, we set ϵ at a fixed value

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

($\epsilon = 0.01$) and measure the degradation degree, as Eq. 4 shows:

$$F = \frac{f_2}{f_1} = [\frac{H'_D}{H_D}, \frac{H'_C}{H_C}, \frac{H'_S}{H_S}, \frac{G'_D}{G_D}, \frac{G'_C}{G_C}, \frac{G'_S}{G_S}] \quad (4)$$

where H'_D, H'_C, H'_S and G'_D, G'_C, G'_S represent the Hoyer measure and Gini index of three transform coefficients of the filtered test sonar image I', respectively. Since distortion might exist in the test sonar image I, the effect of distortion will be reflected in the degradation degree of I, that is, the effect of distortion can be reflected in F. To demonstrate this supposition, Fig. 9 shows the variation of $\frac{H'_S}{H_S}$ and $\frac{G'_S}{G_S}$ when distortion is added to the test sonar image, where image 1 is the pristine sonar image without any distortion and image 2 to image 5 are the distorted versions of image 1. In detail, image 2 contains 'TC', image 3 is afflicted with 'CC', image 4 is distorted by 'TS', and there is 'CS' in image 5. As seen from Fig. 9, $\frac{H'_S}{H_S}$ and $\frac{G'_S}{G_S}$ are related to the distortion contained in a sonar image. In the next section, we try to determine such a relation with the help of bagging-based SVR.

D. Bagging-based SVR module

Conventional quality evaluation models connecting features to evaluated quality only include a single model trained on a subspace of a test database. However, it is easy to prove that the employment of multiple models will improve the performance of the quality evaluation module. To convert the extracted features into a quality index of a sonar image, we employ bagging to build the relationship between extracted features and qualities of sonar images. Bagging is one of the ensemble learning algorithms for statistical classification and regression. In this case, bagging provides multiple versions of a base learner of a regression task and uses these learners to obtain an aggregated result [42]. Theoretical research and experimental results indicate that it can improve the stability and accuracy of machine learning algorithms. In addition, it can also reduce variance and help to avoid overfitting caused by a small-sized training set [42]. For a standard training set T of size N, bagging generates n new training sets, which are denoted by T_i, each of size N', by sampling from T uniformly and with replacement. This implementation may lead to the repetition of some observations in T_i. When $N = N'$, this kind of sample is called a bootstrap sample. The base learners $S_1(\cdot), S_2(\cdot), \cdots, S_n(\cdot)$ are fitted using the above n bootstrap samples and combined by averaging the outputs:

$$NRCDM\ Score = \frac{1}{n} \cdot \sum_{i=1}^{n} S_i(t) \quad (5)$$

Our training set based on the SIQD database includes 840 sonar images, that is, $N = 840$, among which approximately 672 samples will be used for model training (80% of the SIQD database contents), and the other 168 (20% of the SIQD database contents) are the testing samples. There is no overlap between the training set and the testing set. In this paper, SVR is selected as the base learner due to its superior performance. We will discuss the selection of the number of SVR models in the following section since the number of SVR models is relevant to the performance.

IV. Experimental Results and Analysis

A. Testing Metric and Evaluation Protocols

1) Criteria: To assess the performance of the proposed NRCDM metric, the correlation results between the evaluated quality scores and human opinions are reported using four representative criteria. These criteria can be divided into three categories: 1) prediction monotonicity index, including Spearman rank order correlation coefficient (SROCC) and Kendall's rank order correlation coefficient (KROCC); 2) prediction accuracy measure, that is, Pearson linear correlation coefficient (PLCC); and 3) prediction consistency index, which consists of the root mean square error (RMSE). Since there is nonlinearity between original subjective score (MOSs) and objective score process which is caused by the subjective rating, a logistic procedure is applied to remove the nonlinearity, as is defined in Eq. 6:

$$Q_p = \beta_1(\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(Q_o - \beta_3))}) + \beta_4 Q_o + \beta_5 \quad (6)$$

where the objective score is represented by Q_o, and Q_p denotes the IQA score after regression. Additionally, β_1 to β_5 are parameters of the regression model, which are determined by minimizing the sum of squared differences between the mapped Q_p and MOS. The SROCC and KROCC are measured between MOS and Q_o, while PLCC and RMSE are evaluated between MOS and Q_p. A superior correlation with the subjective scores should lead to high values in SROCC, PLCC and KROCC (close to 1) and low values (close to 0) in RMSE.

Since the uncertainty of subjective scores is overlooked by the above stated performance evaluation metrics, a novel performance evaluation method considering the statistical significance of the subjective scores is employed as a supplement to the aforementioned performance evaluation metrics [43]. It can easily combine data from different experiments without considering their source, range, or format methods. First, the z-score for each pair of stimuli (i, j) is calculated using Eq. 7:

$$z(i,j) = \frac{\mid MOS(i) - MOS(j) \mid}{\sqrt{\frac{var(i)}{N(i)} + \frac{var(j)}{N(j)}}} \quad (7)$$

where $var(\cdot)$ is the variance of collected subjective scores, and N is the number of subjective viewers. Then, the cumulative distribution function (cdf) of the normal distribution is employed to measure the probability that the stimuli are different, as Eq. 8 shows:

$$p_s = cdf(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp(\frac{z^2}{2}) dz. \quad (8)$$

When $p_s(i,j)$ is greater than 0.95, stimuli (i,j) is considered to be significantly different. Then, the predicted qualities provided by the selected IQA methods will be processed as

$$\Delta_m(i,j) = q_m(i) - q_m(j) \quad (9)$$

where $q_m(\cdot)$ is the objective quality predicted by a particular IQA method m. The percentage of correctly recognized stimuli

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

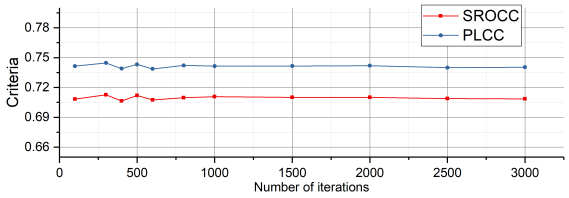


Fig. 10. The influence of the number of repetitions on performance.

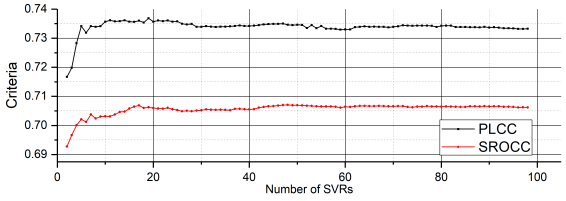


Fig. 11. The influence of the number of SVR models on performance.

of higher quality from the significantly different pairs selected by the IQA methods is measured as

$$C_0 = \frac{\Psi_m}{\Psi_{(i,j)}} \tag{10}$$

where $\Psi_{(i,j)}$ and Ψ_m represent the number of stimuli of higher quality counted from the MOS values and the number of stimuli of higher quality that can be recognized by the IQA method m, respectively. The initial purpose of this method is to evaluate whether an objective model can reliably recognize the better/worse one from two stimuli. The behaviors of different objective methods are compared by contrasting their percentages of correct classification (C_0) with statistical significance, where higher C_0 (close to 1) corresponds to better performance.

2) Algorithms: For the purpose of evaluating the effectiveness of the proposed NRCDM metric, two classes of IQA models (metrics) that have been demonstrated to be performance-efficient were chosen for comparison. The first class is composed of seven reference-based IQA metrics: LESQP [15], SSIM [1], VSNR [2], CPCQI [44], GSM [3], SVQI [45], PSNR. The second class consists of eleven reference-free IQA algorithms: BLIINDS II [11], BQMS [22], ASIQE [27], BRISQUE [40], SISBLIM [46], IL-NIQE [47], CurveletQA [48], NFERM [49], ARISM [50], BPRI [51], HOSA [52].

3) Model training: All experiments presented in this section are based on the bagging-based SVR module. In this paper, we learn the bagging-based SVR module in a training set that includes 63.2% of the sonar images from the SIQD database, and the module is tested in a testing set consisting of the remaining 36.8% of the sonar images. There is no overlap between the training set and the testing set. In this way, we can prevent the occurrence of experimental results that depend on features extracted from known content. In addition, to exclude the performance bias, random choice of training and testing sets is repeated for several iterations. Since the influence of the number of repetitions on the performance is very small, as Fig. 10 shows, the average of the performance criteria across 100 repetitions is reported as the results in this paper.

TABLE II
PERFORMANCE COMPARISON OF FEATURES RELATED TO THE HOYER MEASURE, FEATURES RELATED TO THE GINI INDEX AND THE COMBINATION OF THESE TWO KINDS OF FEATURES.

Features	SROCC	KROCC	PLCC	RMSE
Hoyer measure-related features	0.562	0.385	0.585	11.113
Gini index-related features	0.478	0.325	0.499	11.898
Combination	**0.709**	**0.514**	**0.734**	**9.498**

TABLE III
PERFORMANCE COMPARISON OF FEATURES EXTRACTED FROM DIFFERENT TRANSFORM DOMAINS.

Features	SROCC	KROCC	PLCC	RMSE
$[\frac{H'_D}{H_D}, \frac{G'_D}{G_D}]$	0.583	0.406	0.605	10.827
$[\frac{H'_C}{H_C}, \frac{G'_C}{G_C}]$	0.474	0.327	0.508	11.891
$[\frac{H'_S}{H_S}, \frac{G'_S}{G_S}]$	0.506	0.348	0.542	11.567
$[\frac{H'_D}{H_D}, \frac{G'_D}{G_D}, \frac{H'_S}{H_S}, \frac{G'_S}{G_S}]$	0.562	0.388	0.583	11.105
$[\frac{H'_D}{H_D}, \frac{G'_D}{G_D}, \frac{H'_C}{H_C}, \frac{G'_C}{G_C}]$	0.510	0.350	0.536	11.551
$[\frac{H'_S}{H_S}, \frac{G'_S}{G_S}, \frac{H'_C}{H_C}, \frac{G'_C}{G_C}]$	0.529	0.363	0.563	11.421
All features	**0.709**	**0.514**	**0.734**	**9.498**

B. Parameter and Feature Selection

The bagging-based SVR module is applied in the NRCDM metric to improve the performance, and the number of SVR models is related to its performance. Fig. 11 shows the influence of the number of SVR models, where the horizontal axis corresponds to the number of SVR models, and the vertical axis indicates the corresponding performance of the bagging-based SVR module. As seen from Fig. 11, the performance increases first and then tends to stabilize gradually as the number of the used SVR models increases. The number of SVR models is selected as 20 since it can nearly provide the best performance.

We have chosen six features, i.e., two sparsity measures of three image transform coefficient matrices, to establish the NRCDM metric. First, to verify the contribution of the combination of two sparsity measures, we examine the performance of the Hoyer measure and Gini index. The quality prediction performances with the features related to the Hoyer measure ($[\frac{H'_D}{H_D}, \frac{H'_C}{H_C}, \frac{H'_S}{H_S}]$) and the features related to the Gini index ($[\frac{G'_D}{G_D}, \frac{G'_C}{G_C}, \frac{G'_S}{G_S}]$) are tabulated in Table II by SROCC, KROCC, PLCC, and RMSE. The best performance is indicated in bold. As observed in Table II, employing one sparsity measure can only provide poor performance, and the combination of two sparsity measures enhances the performance, which verifies the effect of the combination that we explained in Section III-B. We have also built the bagging-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

TABLE IV
PERFORMANCE COMPARISON WITH AND WITHOUT "BAGGING".

Without bagging	SROCC	KROCC	PLCC	RMSE
Training set	0.736	0.533	0.753	9.206
Testing set	0.682	0.484	0.709	9.681
With bagging	SROCC	KROCC	PLCC	RMSE
Training set	0.738	0.542	0.759	9.146
Testing set	0.709	0.514	0.734	9.498

TABLE V
PERFORMANCE COMPARISON OF THE NRCDM METRIC AND REFERENCE-BASED IQA METHODS ON THE SIQD DATABASE.

IQA	SROCC	KROCC	PLCC	RMSE
SSIM	0.654	0.469	0.673	10.345
VSNR	0.433	0.299	0.476	11.990
PSNR	0.622	0.443	0.639	10.760
GSM	0.642	0.455	0.658	10.533
SVQI	0.666	0.473	0.691	10.110
CPCQI	0.549	0.377	0.567	11.517
LESQP	**0.785**	**0.593**	**0.796**	**8.474**
NRCDM	0.709	0.514	0.734	9.498

TABLE VI
COMPARISON OF THE STATISTICAL SIGNIFICANCE OF THE NRCDM METRIC AND REFERENCE-BASED IQA METRICS ON THE SIQD DATABASE

IQA	SSIM	VSNR	PSNR	GSM	SVQI	CPCQI	LESQP
Index	$+1$	$+1$	$+1$	$+1$	0	$+1$	0

based SVR modules for features extracted from each domain and the combination of two of three domains. The performance comparison is tabulated in Table III. In Table III, 'All features' denotes the features extracted from the DCT domain, C-D-F 9/7 wavelet transform domain, and SVD domain, that is, $[\frac{H'_D}{H_D}, \frac{H'_C}{H_C}, \frac{H'_S}{H_S}, \frac{G'_D}{G_D}, \frac{G'_C}{G_C}, \frac{G'_S}{G_S}]$. We find that the combination of features extracted from all three domains obtains a significant performance gain compared with features extracted from only one or two domains.

To prove the effectiveness of 'bagging', we first compare the performance of the proposed method on the training set and testing set. Then, we replace the bagging-based SVR with the normal SVR. The experiments are repeated 100 times in 100 randomly divided training-testing sets (completely nonoverlapping). The average experimental results are tabulated in Table IV. The following conclusions can be obtained from Table IV: (1) since the performances of the bagging-based SVR and normal SVR are close to each other in the training set, the extracted features are a good representation of the characteristics of sonar image samples, and (2) it is obvious that the addition of the bagging-based SVR increases the performance of the proposed method in the testing set, which can be a good justification of the efficiency of 'bagging'.

C. Performance Evaluation

1) Performance Comparison with Existing IQA Methods: We compare the performance of the proposed NRCDM metric with the classical and state-of-the-art reference-based IQA methods. Since reference information is available for the reference-based IQA method, most of them show better performances than reference-free methods. We tabulate the performances of seven reference-based IQA methods in Table V and then highlight the best and the second best performance in bold font and underline, respectively. Among the selected IQA methods, the LESQP metric, which was designed for sonar images, is approximately 10% better than the proposed NRCDM metric, but the NRCDM metric needs no reference information, while the LESQP is reference-dependent. Except for LESQP and NRCDM, the SVQI metric shows superior performance compared to the rest of the reference-based IQA methods, and the SSIM, GSM and PSNR behave similarly but slightly worse than SVQI. All the selected methods except LESQP show good performances for NSIs or screen content images, but their performances are worse than or close to the NRCDM metric, as Table V shows. In addition, to statistically compare the performance of the selected algorithms and the proposed NRCDM metric, the F-test is employed here [52]-[53]. First, the ratio of the prediction residual variances of the regressed objective qualities (using the five-parameter logistic nonlinear regression function) and MOS values is denoted as F_t. Then, the critical threshold, which is decided by the number of residuals and the confidence level, is denoted by F_{ct}. In the case of $F_t > F_{ct}$, the performance of these two testing IQA methods can be regarded as significantly distinct; then, by considering Table V, the statistically superior one can be decided between the two testing IQA methods. In this paper, the confidence level is assigned to be 95%. The results of the significance comparison are tabulated in Table VI, where '$+1$' represents that the proposed NRCDM metric is significantly superior to the associated metric, and '0' or '-1' means that our NRCDM metric is significantly equivalent or worse than the comparable IQA method. As seen from Table VI, the proposed NRCDM metric is significantly better than most of the selected reference-based IQA methods. Only the SVQI and LESQP, which are full-reference methods, are statistically indistinguishable from our NRCDM metric.

Having compared the NRCDM metric with the reference-based IQA methods, we now demonstrate the performance of the NRCDM metric for different types of distortion. Fig. 12 shows the performance comparison of the NRCDM metric for different distortions on the SIQD database, where the horizontal axis represents the distortion types contained in the SIQD database, and the vertical axis represents the value of the corresponding criteria. It can be observed that the NRCDM metric shows the best prediction monotonicity and accuracy for distorted sonar images with 'TC'. The best prediction consistency is obtained for distorted sonar images with 'CS'. The distortion 'CC', which is relevant to blur, makes the NRCDM

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

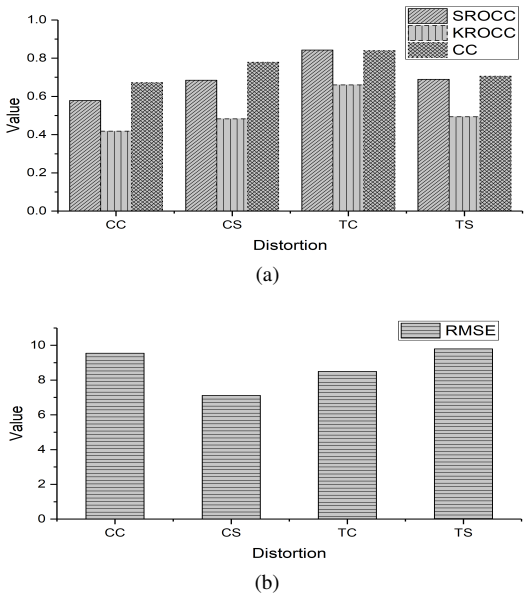


Fig. 12. Performance comparison for different distortions contained in the SIQD database. (a) SROCC, PLCC and KROCC; (b) RMSE.

metric perform poorly on prediction monotonicity and accuracy. The worst prediction consistency appears on sonar images with 'TS', which contains messy, unnatural marks. Then, we compare the performance of the NRCDM metric and the state-of-the-art reference-free IQA methods. We bold the algorithm with the best performance and underline the algorithm with the second best performance. It can be concluded from Table VII that the proposed NRCDM metric is competitive with all the selected reference-free IQA algorithms. The performance gain of the proposed NRCDM metric is approximately 16% in terms of SROCC, 21% in terms of KROCC and over 19% and 14% in terms of PLCC and RMSE, respectively, in comparison with the second-ranking reference-free IQA metrics, which indicates the improvement in performance accuracy, consistency and monotonicity. Only the proposed NRCDM metric achieves a performance greater than 0.7 for SROCC and PLCC and greater than 0.5 for KROCC but lower than 10 for RMSE. Similarly, the F-test is also implemented between the NRCDM metric and the selected reference-free IQA methods, and the results are tabulated in Table VIII. It is obvious that the proposed NRCDM metric is significantly superior to the other eleven reference-free IQA methods.

To compare the performances of the selected reference-free methods by C_0, we first numbered eleven selected methods and proposed NRCDM methods according to Table IX. The results of C_0 for different IQA methods are shown in the left of Fig. 13. In addition, to statistically compare C_0 among multiple IQA methods, Fisher's exact test and a Benjamini-Hochberg procedure are employed here [43]. The statistical analysis of C_0 is supplemented in the right of Fig. 13, where the gray boxes in the significant plots correspond to the cases where the model in the column performs similar to the model in the row. When the model in the column outperforms the model in the row, the corresponding box is black; otherwise, the box is white. From Fig. 13, we can observe that the NRCDM metric (#1) performs statistically better than the other selected reference-free IQA methods. Apart from the NRCDM metric, ARISM (#6), ASIQE (#7) and CurveletQA (#10) also show competitive performance, considering the ability to recognize the better/worse one from two stimuli.

2) Computational Time Comparison: The comparison of computational time is implemented by using the software platform of MATLAB R2014a on a computer with 3.60 GHz CPU processor and 8.00 GB RAM. We obtained the source codes of the selected IQA metrics from the authors or websites. For each IQA method, we first compare the time consumption for all images on the SIQD database and then report their average in the last column of Table VII. Although the NRCDM metric is not the fastest one, its performance is 16% better than the fastest competitor: BRISQUE. Additionally, it takes less processing time than most of the selected methods and is

TABLE VII
PERFORMANCE COMPARISON OF REFERENCE-FREE IQA ALGORITHMS ON TEST SONAR IMAGES

IQA	SROCC	KROCC	PLCC	RMSE	run-time (s)
BLIINDS II	0.441	0.302	0.448	12.505	14.70
BRISQUE	<u>0.607</u>	<u>0.425</u>	<u>0.614</u>	<u>11.036</u>	$\mathbf{3.00 \times 10^{-4}}$
IL-NIQE	0.553	0.394	0.591	11.282	<u>6.53×10^{-2}</u>
ARISM	0.451	0.308	0.465	12.384	2.51
NFERM	0.423	0.290	0.441	12.552	14.81
SISBLIM	0.421	0.296	0.515	11.985	8.68×10^{-1}
ASIQE	0.308	0.211	0.508	12.043	2.91×10^{-1}
BQMS	0.374	0.257	0.444	12.533	14.49
CurveletQA	0.424	0.289	0.504	12.075	6.64×10^{-1}
BPRI	0.445	0.309	0.474	12.314	4.01×10^{-1}
HOSA	0.567	0.392	0.564	11.548	3.17×10^{-1}
NRCDM	**0.709**	**0.514**	**0.734**	**9.498**	7.78×10^{-2}

TABLE VIII
COMPARISON OF THE STATISTICAL SIGNIFICANCE OF THE NRCDM METRIC AND REFERENCE-FREE IQA METRICS ON THE SIQD DATABASE

IQA	BIIINDS II	BRISQUE	IL-NIQE	ARISM	NFERM
Index	+1	+1	+1	+1	+1
SISBLIM	ASIQE	BQMS	CurveletQA	BPRI	HOSA
+1	+1	+1	+1	+1	+1

TABLE IX
NUMBER OF SELECTED REFERENCE-FREE METHODS

Number	1	2	3	4
IQA	NRCDM	BQMS	IL-NIQE	NFERM
Number	5	6	7	8
IQA	SISBLISM	ARISM	ASIQE	BLIINDS II
Number	9	10	11	12
IQA	BPRI	CurveletQA	HOSA	BRISQUE

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

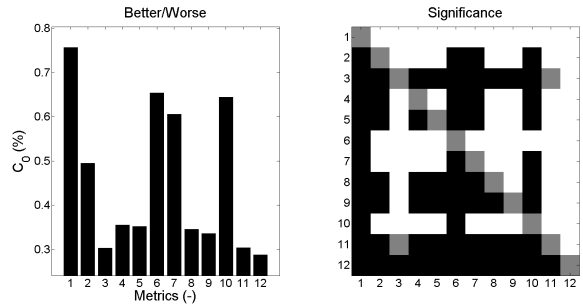


Fig. 13. The results and statistical analysis of C_0 (%) for the SIQD database.

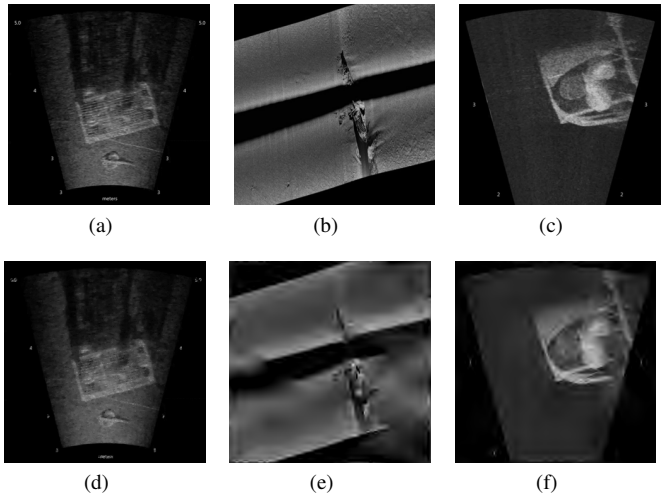


Fig. 14. Examples of the quality predictions provided by the NRCDM metric, which shows the good performance of the proposed NRCDM metric. (a)-(c) Reference sonar images; (d) Distorted version of (a), MOS=55.24, NRCDM score=56.95; (e) Distorted version of (b), MOS=36.31, NRCDM score=37.54; and (f) Distorted version of (c), MOS=36.37, NRCDM score=38.12.

approximately 5 times faster than BPRI and HOSA, 10 times faster than SISBLIM and CurveletQA, 30 times faster than ARISM and approximately 100 times faster than BLIINDS II, NFERM and BQMS.

3) Intuitive Comparison: For visual comparison, Fig. 14 shows examples of quality predictions provided by the NRCDM metric, which are very close to the MOS value provided by the SIQD database. In addition, scatter plots between MOSs and predictive qualities yielded by the eight selected reference-free IQA metrics and the NRCDM metric are presented in Fig. 15. In each scatter plot, distinct symbols are used to discriminate different distortion types: red circle for 'CC', green square for 'CS', dark-blue diamonds for 'TC', blue triangle for 'TS', purple inverted triangle for reference images without any distortion, and the red line is the five-parameter logistic mapping curve between predictive qualities and MOSs. Then, we provide scatter plots between the regressed predictive qualities and MOSs in Fig. 16. Regardless of considering either Fig. 15 or in Fig. 16, the quality scores predicted by our NRCDM metric show the best correlation to MOSs when compared with scatter plots of the other reference-free IQA metrics.

V. Summary

This paper proposes a contour degradation measurement for sonar IQA. The motivation of this work includes the following two points. First, due to the poor condition of the underwater acoustic channel, quality monitoring during compression and transmission is of significance to system optimization. To the best of our knowledge, little has been devoted to reference-free sonar IQA; therefore, it is a pivotal but unsolved topic. Second, the proposed method in this paper is motivated by the characteristics and applications of sonar images. In detail, object detection and recognition are two important applications of sonar images. They can be implemented by extracting macroscopic structures, typically in the form of contour measurements. The sonar images are characterized by less details and the grayscale, which indicates that microscopic structures can be negligible and of less importance compared to macroscopic structures. Inspired by the above two points, macroscopic structure information is extracted to represent the distortion of sonar images in this paper.

As mentioned earlier, contour is one of the important forms of macroscopic structures; consequently, we use contour information to represent macroscopic structures. The contour information is extracted from the sparsity of different transforms of sonar images. Then, the distortion of the sonar image is measured by the contour degradation degree after filtering. We have adopted an efficient approach in this paper for extracting contour information. Because of the complexity of the acquisition environment, it is difficult to obtain a large amount of sonar images. Thus, the size of the SIQD database is small. Considering the small scale of the SIQD database, "bagging" is employed to avoid the overfitting problem. On the one hand, the proposed mixture of different techniques takes both application requirements and database conditions into account. Furthermore, its effectiveness has been verified by experiments. On the other hand, a new technology for sonar IQA is one of our important research interests in our future work.

The novelty of this work can be reflected in the two points below. First, sonar IQA is a novel and imperative topic. To date, few efforts have been devoted to sonar IQA, which might be due to the following reasons: 1) the underwater environment for image acquisition and transmission is complicated, and 2) many differences between sonar images and typical visual images exist. To the best of our knowledge, we are one of the first teams to study sonar image quality assessment considering underwater acoustic transmission and have presented a simple and efficient strategy for addressing this difficult issue to a certain extent. Second, the importance of macroscopic structure represented by contour measurement has been employed to capture the effects of distortion in sonar image applications. In addition, its feasibility has been demonstrated through theoretical analysis and synthetic experiments.

VI. Conclusion

Sonar images carry important underwater information since they can reflect underwater scenes in darkness. Due to the poor condition of underwater acoustic channels, transmitted

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

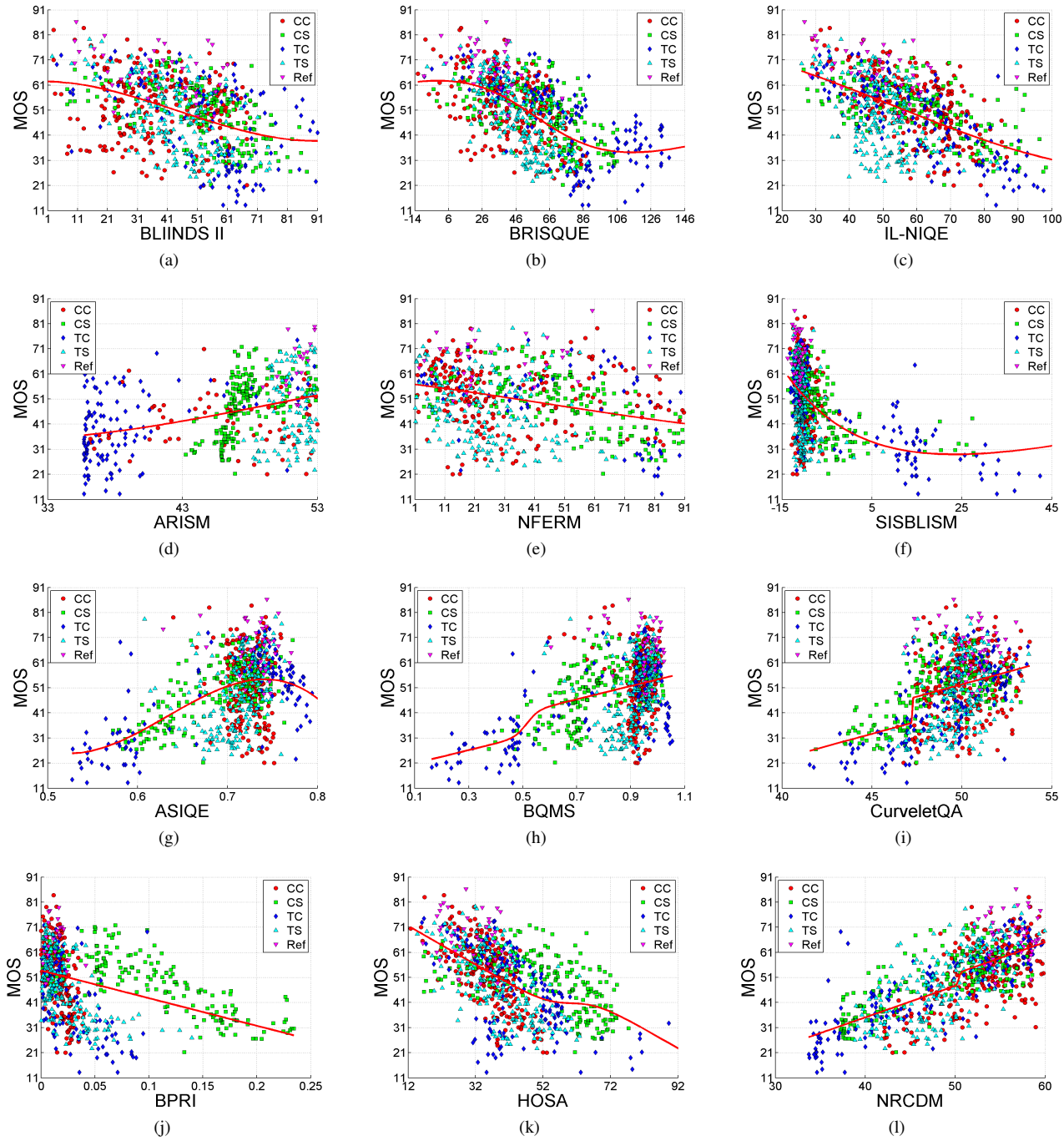


Fig. 15. Scatter plots of MOS versus model prediction. Each sample point represents one test image. (a) BLIINDS II; (b) BRISQUE; (c) IL-NIQE; (d) ARISM; (e) NFERM; (f) SISBLIM; (g) ASIQE; (h) BQMS; (i) CurveletQA; (j) BPRI; (k) HOSA; and (l) NRCDM.

sonar images often contain different distortions after lossy compression and underwater transmission. It is necessary to build a sonar image quality evaluation model for transmission monitoring. Although there have been many well-established IQA models for NSIs, they do not work for sonar images. The reason can be summarized by the following two points. First, viewers perceive utility quality in the subjective quality assessment step for sonar images due to its application, while most existing IQA algorithms generate a perceptual quality score, which is not an alternative for utility qualities. Second, the macroscopic structures, which contain the most important information for sonar image applications, are relevant to the utility quality of sonar images, but most existing structure-based IQA models for NSIs take both macroscopic and microcosmic structures into account. To address these problems, we have established a degradation measurement-based IQA metric (NRCDM) for sonar images. First, we transform an image into three domains: the DCT domain, C-D-F 9/7 wavelet transform domain and SVD domain. Then, the sparsity of coefficient matrices of these three domains is measured using the Hoyer measure and the Gini index as features. We extract the same features after degrading the test image by blur

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2019.2910666, IEEE Transactions on Image Processing

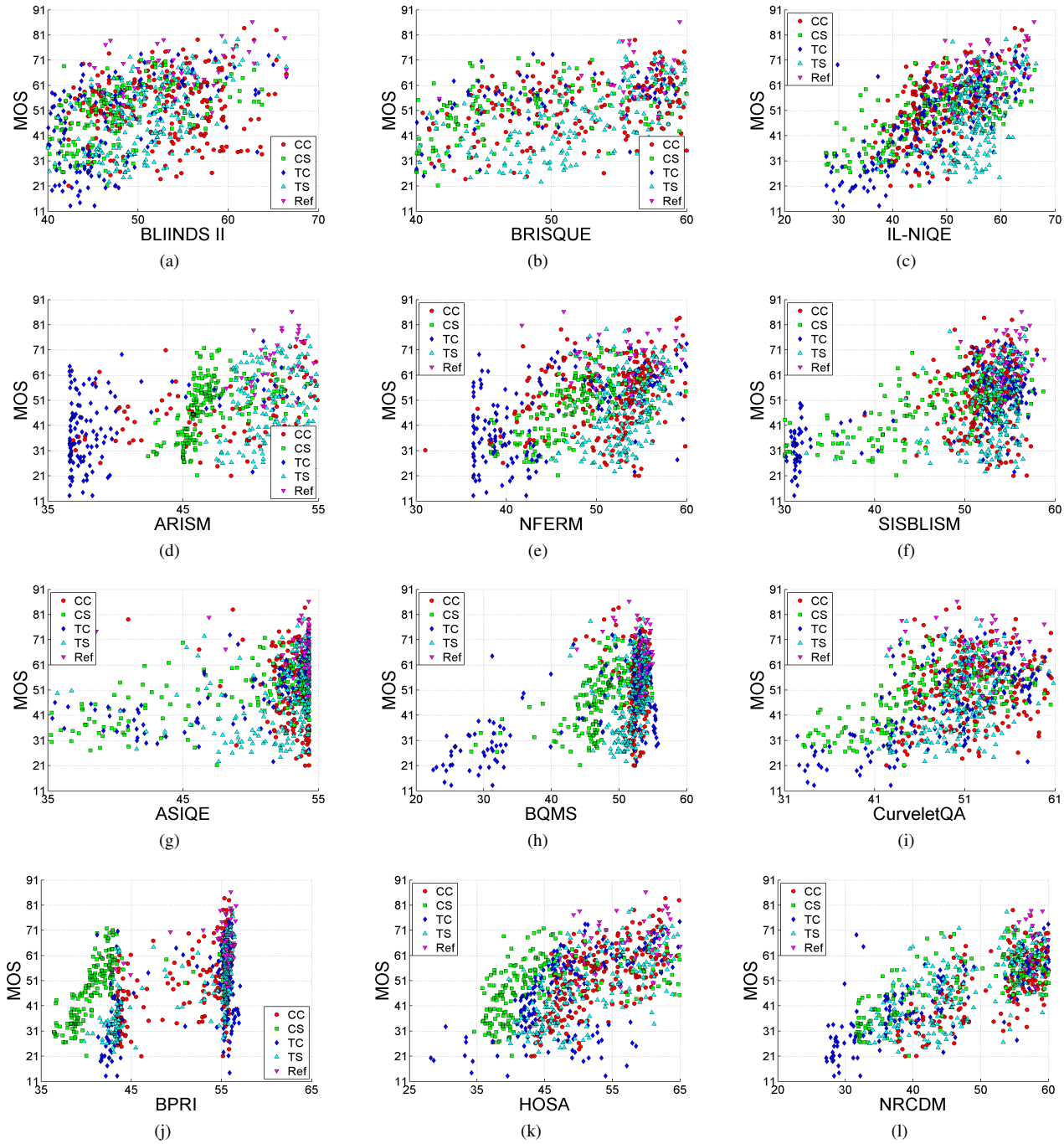


Fig. 16. Scatter plots of MOS versus regressed model prediction. Each sample point represents one test image. (a) BLIINDS II; (b) BRISQUE; (c) IL-NIQE; (d) ARISM; (e) NFERM; (f) SISBLIM; (g) ASIQE; (h) BQMS; (i) CurveletQA; (j) BPRI; (k) HOSA; and (l) NRCDM.

degradation using a guided image filter. To establish the correlation between features and image quality, a bagging-based SVR module is trained in this paper. The ratios of features before and after degradation are the input of this module, while the output is a NRCDM score that represents the quality of the test image. The performances of the proposed NRCDM metric and the existing classical and state-of-the-art IQA algorithms on the SIQD database are compared in this paper. The experimental results show the poor effect on sonar images of the IQA algorithms that are designed for NSIs and the superiority of the reference-free NRCDM metric over the available quality evaluation models.

REFERENCES

[1] Wang, Z., Bovik, A., Sheikh, H., & Simoncelli, E., "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[2] Chandler, D., & Hemami, S., "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2284-2298, 2007.

[3] Liu, A., Lin, W., Narwaria, M., "Image quality assessment based on gradient similarity," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 21, no. 4, pp. 1500-1512, Apr. 2012.

[4] Gao, F., Tao, D., Gao, X., & Li, X., "Learning to Rank for Blind Image Quality Assessment," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 10, pp. 2275-2290, 2015.

[5] Hou, W., Gao, X., Tao, D., & Li, X., "Blind Image Quality Assessment via Deep Learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 6, pp. 1275-1286, 2015.

[6] Liang, L., Wang, S., Chen, J., Ma, S., Zhao, D., & Gao, W., "No-reference perceptual image quality metric using gradient profiles for JPEG2000," *SIGNAL PROCESSING-IMAGE COMMUNICATION*, vol. 25, no. 7, pp. 502-516, 2010.

[7] Kang, L., Ye, P., & Doermann, D., "Convolutional neural networks for no-reference image quality assessment," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1733-1740. 2014.

[8] Xue, W., Zhang, L., & Mou, X., "Learning without human scores for blind image quality assessment," *Computer Vision and Pattern Recognition (CVPR)*, pp. 995-1002, 2013.

[9] Ma, K., Liu, W., Zhang, K., Duanmu, Z., Wang, Z., & Zuo, W.,"End-to-End Blind Image Quality Assessment Using Deep Neural Networks," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1202-1213, 2018.

[10] Ma, K., Liu, W., Liu, T., Wang, Z., & Tao, D., "dipIQ: Blind image quality assessment by learning-to-rank discriminable image pairs," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 3951-3964, 2017.

[11] Saad, M. A., Bovik, A. C., & Charrier, C., "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339-3352, 2012.

[12] K. Gu, V. Jakhetiya, J.-F. Qiao, X. Li, W. Lin, and D. Thalmann, "Model-based referenceless quality metric of 3D synthesized images using local image description," *IEEE Trans. Image Process.*, pp. 394-405, vol. 27, no. 1, Jan. 2018.

[13] Zhang, J., Chang, S., Zhang, L., Su, Y., & Fu, X., "Sparse Wavelet Transform for Underwater Acoustic Image Compressed Sensing," *OCEANS'18 MTS/IEEE Kobe*, 2018.

[14] Sung, M., Joe, H., Kim, J., & Yu, S., "Convolutional Neural Network based Resolution Enhancement of Underwater Sonar Image Without Losing Working Range of Sonar Sensors," *OCEANS'18 MTS/IEEE Kobe*, 2018.

[15] Chen, W., Yuan, F., Cheng, E., & Lin, W., "Subjective and objective quality evaluation of sonar images for underwater acoustic transmission," *International Conference on Image Processing*, 2017.

[16] Rouse, D., Hemami, S., Ppion, R., & Patrick, L., "Estimating the usefulness of distorted natural images using an image contour degradation measure," *Journal of the Optical Society of America. A, Optics, Image Science, and Visions*, vol. 28, no.2, pp. 157-188, 2011.

[17] G. Loffler, "Perception of contours and shapes: low and intermediate stage mechanisms," *Vis. Res*, vol. 48, pp. 2106-2127, 2008.

[18] S. O. Dumoulin, S. C. Dakin, & R. F. Hess, "Sparsely distributed contours dominate extra-striate responses to complex scenes," *NeuroImage*, vol. 42, pp. 890-901, 2008.

[19] Kabrisky, M., Tallman, O., Day, C., & Radoy, C., "A theory of pattern perception based on human physiology," *Ergonomics*, vol. 13, no. 1, pp. 129-147, 1970.

[20] Carl, J., & Hall, C., "The Application of Filtered Transforms to the General Classification Problem," *IEEE Transactions on Computers*, vol. C-21, no. 7, 1972.

[21] Hall, C., Hall, E., "A Nonlinear Model for the Spatial Characteristics of the Human Visual System," *Systems Man & Cybernetics IEEE Transactions on*, vol. 7, no. 3, pp. 161-170, 1977.

[22] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Learning a blind quality evaluation engine of screen content images," *Neurocomputing*, vol. 196, pp. 140-149, Jul. 2016.

[23] Hurley, N., & Rickard, S., "Comparing measures of sparsity," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4723-4741, 2009.

[24] Ye, Z., & Mohamadian, H., "Multilevel Wavelet Transform Based Sparsity Reduction for Compressive Sensing," *9th International Conference on Intelligent Environments*, vol. 17, pp. 4-11. July, 2013.

[25] Zhang, M,, Chen, T., Shi, X., & Cao, P., "Image Arbitrary-Ratio Down- and Up-Sampling Scheme Exploiting DCT Low Frequency Components and Sparsity in High Frequency Components," *IEICE Transactions on Information & Systems*, vol. E99D, no. 2, pp. 475-487, Feb. 2016.

[26] Dan, K., "Singularly Valuable Decomposition: The SVD of a Matrix," *College Mathematics Journal*, vol. 27, no. 1, pp. 2-23, 1996.

[27] K. Gu, J. Zhou, J. Qiao, G. Zhai, W. Lin, and A. C. Bovik, "No-reference quality assessment of screen content pictures," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 4005-4018, Aug. 2017.

[28] Tao, D., Tang, X., Li, X., & Wu, X., "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 28, no. 7, pp. 1088-1099, Jul. 2006.

[29] Ma, K., Zhao, T., Zeng, K., & Wang, Z., "Objective Quality Assessment for Color-to-Gray Image Conversion," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4673-4685, 2015.

[30] Rouse, D., Pepion, R., Hemami, S., & Patrick, L., "Image utility assessment and a relationship with image quality assessment," *Proceedings of SPIE - The International Society for Optical Engineering*, 2009.

[31] Xue, W., Zhang, L., Mou, X., & Bovik, A., "Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index," *IEEE Transactions on Image Processing*, vol. 23, no.2, pp. 684-695, 2014.

[32] Moorthy, A., Bovik, A., "Blind Image Quality Assessment: From Natural Scene Statistics to Perceptual Quality," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 20, no. 12, pp. 3350-3364, 2011.

[33] Sheikh,H., Bovik, A., & Cormack, L., "No-reference quality assessment using natural scene statistics: JPEG2000[J," *IEEE Transactions on Image Processing*, vol.14, no. 11, pp. 1918-1927, 2005.

[34] Liu, T., Liu, K., "No-Reference Image Quality Assessment by Wide-Perceptual-Domain Scorer Ensemble Method," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1138-1151, 2018.

[35] He, K., Sun, J., & Tang, X., "Guided Image Filtering," *European Conference on Computer Vision, ECCV 2010*, pp. 1397-1409, 2010.

[36] Chen, W., Yuan, F., & Cheng, E., "Adaptive Underwater Image Compression with High Robust Based on Compressed Sensing," *IEEE International Conference on Signal Processing, Communications and Computing*, 2016.

[37] Said, A., & Pearlman, W., "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on circuits and systems for video technology.*, vol. 6, no. 3, pp. 243-250, Jun. 1996.

[38] Tomasi, B., Toni, L., Casari, P., Preisig, J., & Zorzi, M., "A study on the SPIHT image coding technique for underwater acoustic communications," *Proceedings of Acm International Workshop on Underwater Networks Wuwnet*, 2011.

[39] Yin, Y., Zhou, F., Qiao, G., & Liu, S., "Orthogonal multicarrier M-ary cycle shift keying spread spectrum underwater acoustic communication," *Acta Physica Sinica.*, vol. 62, no. 22, pp. 224302-224302, 2013.

[40] Mittal, A., Moorthy, A. K., & Bovik, A. C., "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, 2012.

[41] Xu, L., Yan, Q., Xia, Y., & Jia, J., "Structure extraction from texture via relative total variation," *Acm Transactions on Graphics*, vol. 31, no. 6, pp. 1-10, 2012.

[42] Breiman, L., "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123-140, 1996.

[43] Krasula, L., Fliegel, K., Callet, P., & Milos, K., "On the accuracy of objective image and video quality models: New methodology for performance evaluation," *Eighth International Conference on Quality of Multimedia Experience. IEEE*, 2016.

[44] K. Gu, D. Tao, J.-F. Qiao, and W. Lin, "Learning a no-reference quality assessment model of enhanced images with big data," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 29, no. 4, pp. 1301-1313, Apr. 2018.

[45] K. Gu, J. Qiao, X. Min, G. Yue, W. Lin, and D. Thalmann, "Evaluating quality of screen content images via structural variation analysis," *IEEE Trans. Visual. Comput. Graphics*, vol. 24, no. 10, pp. 2689-2701, Oct. 2018.

[46] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Transactions on Broadcasting*, vol. 60, no. 3, pp. 555-567, Sep. 2014.

[47] Zhang. L., Zhang, L., & Bovik, A., "A feature-enriched completely blind image quality evaluator," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579-2591, Aug. 2015.

[48] Liu, L., Dong, H., Huang, H., & Bovik, A., "No-reference image quality assessment in curvelet domain," *Signal Processing Image Communication*, vol. 29, no. 4, pp. 494-505, 2014.

[49] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50-63, Jan. 2015.

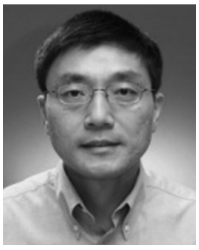
[50] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218-3231, Oct. 2015.

[51] X. Min, K. Gu, G. Zhai, J. Liu, X. Yang, and C. W. Chen, "Blind quality assessment based on pseudo-reference image," *IEEE Trans. Multimedia*, vol. 20, no. 8, pp. 2049-2062, Aug. 2018.

[52] Xu, J., Ye, P., Li, Q., Du, H., Liu, Y., & Doermann, D., "Blind Image Quality Assessment based on High Order Statistics Aggregation," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4444-4457, 2016.
[53] M. Liu, K. Gu, G. Zhai, P. Le Callet, and W. Zhang, "Perceptual reduced-reference visual quality assessment for contrast alteration," *IEEE Transactions on Broadcast.*, vol. 63, no. 1, pp. 71-81, Mar. 2017.

Weiling Chen received the B.S. and PhD degrees in communication engineering from Xiamen University, Xiamen, China, in 2013 and 2018. She is currently a lecturer in Fuzhou University, Fuzhou, China. From Sep. 2016 to Dec. 2016, she was visiting at the School of Computer Science and Engineering, Nanyang Technological University, Singapore. And she is the reviewer for ICIP 2016, ICIP 2017, IEEE ACCESS, DSP, T-CSVT and T-IP. Her current research interests include image quality assessment, image compression, and underwater acoustic communication.

Ke Gu received the B.S. and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015, respectively. He is currently a Professor with the Beijing University of Technology, Beijing, China. His research interests include environmental perception, image processing, quality assessment, and machine learning. He received the Best Paper Award from the IEEE Transactions on Multimedia (T-MM), the Best Student Paper Award at the IEEE International Conference on Multimedia and Expo (ICME) in 2016, and the Excellent Ph.D. Thesis Award from the Chinese Institute of Electronics (CIE) in 2016. He was the Leading Special Session Organizer in the VCIP 2016 and the ICIP 2017, and serves as a Guest Editor for the Digital Signal Processing (DSP). He is currently an Area Editor for Signal Processing Image Communication (SPIC), and an Associate Editor for the IEEE Access and the IET Image Processing (IET-IP). He is a Reviewer for 20 top SCI journals.



Weisi Lin (F'16) received the Ph.D. degree from Kings College London. He is currently an Associate Professor with the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include image processing, visual quality evaluation, and perception-inspired signal modeling, with more than 340 refereed papers published in international journals and conferences. He has been on the Editorial Board of the IEEE Transactions on Image Processing (T-IP), the IEEE Transactions on Multimedia (T-MM, 2011-2013), the IEEE Signal Processing Letters (SPL), and the Journal of Visual Communication and Image Representation (JVCI). He has been elected as an APSIPA Distinguished Lecturer (2012/13). He served as a Technical-Program Chair for Pacific-Rim Conference on Multimedia 2012, the IEEE International Conference on Multimedia and Expo 2013, and the International Workshop on Quality of Multimedia Experience 2014. He is a fellow of Institution of Engineering Technology, an Honorary Fellow of the Singapore Institute of Engineering Technologists, and a Fellow of IEEE.

Zhifang Xia received the B.S. degree in measuring and control instrument from Anhui University, Hefei, China in 2008 and received the Master degree in control science and engineering from Tsinghua University, Beijing, China in 2012. She is currently an engineer and a registered consultant (investment) with State Information Center, Beijing, China. Her interests include image processing, quality assessment, machine learning and e-government. She won the second prize of National excellent engineering consultation award in 2016.

Patrick Le Callet received the M.Sc. and Ph.D. degrees in image processing from École Polytechnique de l'Université de Nantes, Nantes, France, in 1993 and 2001, respectively. He was an Assistant Professor from 1997 to 1999 and a full-time Lecturer from 1999 to 2003 with the Department of Electrical Engineering, Technical Institute, University of Nantes (IUT), Nantes, France. Since 2003, he has been a Teacher with the Electrical Engineering and Computer Science Department, Ecole Polytechnique de l'Université de Nantes. He is currently a Full Professor with the Computer Science Department, Ecole Polytechnique de l'Université de Nantes. Since 2006, he has been the Head of the Image and Video Communication Lab, CNRS IRCCyN, Nantes, France. He is also a Co-Chair within the Video Quality Expert Group, for the "JointEffort Group" and "3DTV" activities. He has authored or co-authored more than 200 publications and communications, and holds 13 international patents. His research interests include the application of human vision modeling in image and video processing, 3D image and video quality assessment, watermarking techniques, and visual attention modeling and applications. Prof. Le Callet is currently serving as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the Springer EURASIP Journal on Image and Video Processing, and SPIE Electronic Imaging.

En Cheng received the B.S., M.S. and Ph.D. degrees in electronic engineering from Xiamen University, Xiamen, China in 1985, 1988 and 2006, respectively. Now he is a professor and supervisor for Ph.D. candidates in Xiamen University. And he is the standing deputy director of the Key Laboratory of Underwater Acoustic Communication and Marine Information Technology (Xiamen University), Ministry of Education, Xiamen, China. Since 2002, he has been the senior member of the Chinese Institute of Electronics. And from 2003 to 2012, he was the department head of the department of communication engineering, Xiamen University, China. He published more than 50 papers in international journals and conferences. Prof. Cheng's main research interests include underwater acoustic communication and networks.