A blind contour-aware quality model for sonar images

Weiling Chen¹, Rongfu Lin¹, Rongxin Zhang², Yi Zhu³

¹ Fujian Key Lab for Intelligent Processing and Wireless Transmission of Media Information, Fuzhou University, Fuzhou, China
² School of Electronic Science and Engineering (National Model Microelectronics College), Xiamen University, Xiamen, China
³ Key Laboratory of Underwater Acoustic Communication and Marine Information Technology, Ministry of Education, Xiamen University, Xiamen, China

Correspondence
Rongxin Zhang, School of Electronic Science and Engineering (National Model Microelectronics College), Xiamen University, Xiamen, China.
Email: zhrx@xmu.edu.cn

Abstract
Since the penetration depth of visible light is limited in muddy and dark deep marine environment, sonar imaging, which is independent of natural light, has been attracting much attention in underwater detections. However, the visual quality of sonar image is inevitably damaged during the acquisition and the transmission in the sophisticated and dynamic underwater environment. The quality decrease can further bring negative impacts on oceanic information analysis. To measure the quality decrease of sonar images, this paper proposes a contour-aware model for sonar image quality assessment (CaMSIQA). The CaMSIQA metric is established upon the conclusion that the quality of sonar images can be defined as a measure of their utilities in real application scenarios. We exploit a contour information statistic (CIS) model that is critical to object recognition of sonar images here. Quantifying the changes of the CIS model between the unimpaired and test sonar images makes it possible to perceive the quality decrease of sonar images. Extensive experimental results have demonstrated the effectiveness of the proposed method compared with current mainstream image quality assessment (IQA) methods without full reference information.

1 INTRODUCTION

As the emergence of ocean explorations in recent years, underwater detection has got increasing attention. The object imaging in muddy and dark deep marine environments is generally achieved via sonar devices since they require no natural light for imaging. A complete sonar system would take a series of processing steps, that is, image acquisition, compression and transmission, each of which often leads to distortions in consequence of the limitations of bandwidth and the deterioration in the underwater acoustic channels. The ability to estimate the impacts of these impairments on sonar image quality is a significant problem in underwater detection tasks. Therefore, reliably exploiting sonar image quality assessment (SIQA) methods for monitoring the wide field of image processing procedures is of high practical relevance.

Due to the insufficient capabilities of current image processing, human intervention is still required in underwater detection. Definitely, numerous objective image quality assessment (IQA) approaches designed from human-oriented perspectives have been proposed for optical images. Presently, IQA algorithms could be commonly categorized into three types. The first type is a statistic-based approach. These algorithms outline the image intrinsic characteristics by utilizing a statistic model from different aspects. The representative work of this type includes natural scene statistics (NSS) model [1, 2], information and structure models [3, 4], and so on. The second type predicts perceptual quality by modelling the characteristics of the human visual system (HVS) from psychological vision science. These approaches quantify the visual attributes in image understanding and appreciation by making assumption of the HVS’s behaviour [5–8]. The third type trains a deep neural network to predict image quality based on a corpus of high-level features [9–11].

Though a large set of IQA methods has been developed for optical images, their performance indices are far less than the ideal results in SIQA. The reasons are threefold: (1) Sonar images have different statistical characteristics from optical images. To be specific, the optical images display rich colour variations, complex texture content and thick lines, while sonar images are greyscale with low mean value and bimodal histogram. For more intuitive comparison, an example of a sonar image compared with an optical image is shown in Figure 1. Moreover, images obtained by distinct sonar devices have different characteristics. Therefore, the SIQA is equipment specific...
and those statistic-based models cannot be applied. (2) Popular deep learning–based IQA methods may cause serious overfitting in sonar images. The expensive and time-consuming acquisition of sonar images limits the size of sonar image database. (3) Most optical IQA algorithms are general purpose and visual perception oriented whereas the SIQA should be task oriented since the evaluation criteria of sonar image quality is application specific. We tend to assess the utility quality of sonar images under specific applications which cannot be substituted by perceptual quality [12, 13].

We define the image utility quality as a practicability indicator of the test image in specific applications. Some device- and operation-based parameters have been extracted for synthetic aperture sonar (SAS) image quality assessment. The SAS image that performs well at adapting the route of a sonar-equipped autonomous underwater vehicle is considered to have sufficient quality. In this case, the measurements of sonar-platform motion, environmental characteristics and the degree of navigation errors have been employed to represent the utility quality of SAS images [14, 15]. In [16], some low-level features are correlated with the utility quality of forward looking sonar (FLS) images. The extracted features are the indicators of obstacle recognition performance of the test image. In these works, the utility quality of a sonar image with low precision and resolution is predicted based upon the performance of relevant underwater equipment in specific applications. Nevertheless, an increasing number of sonar images require human decision in underwater detection due to the appearance of high-precision sonar imaging technologies in complex detection tasks. For instance, the advanced acoustic lens and side-scan sonar devices are commonly used in underwater detection in recent years because of their sufficient precision and resolution [17]. In human-involved underwater detection, sonar images with clear and visually recognizable object are considered as high quality. According to this principle, we have proposed a series of SIQA methods in [18–20]. These methods provide predicted utility quality of acoustic lens and side-scan sonar images in respect of visual characteristics of object detection.

Despite the aforementioned efforts, SIQA is still a largely underexplored field compared with optical IQA. The important and challenging issues remained in SIQA can be divided into three categories. First, in order to adapt to the requirements of complex detection, the utility quality should be evaluated in consideration of HVS. However, existing SIQA approaches for low-resolution images only take service indicators into account. Second, due to the limited bandwidth of underwater acoustic channels, the transmission burden should be as less as possible. Based upon the accessibility of the reference information, IQA algorithms can be classified into three groups. The full-reference (FR) [21, 22], no-reference (NR) [23, 24] and reduced-reference (RR) [25, 26] methods predict quality of test image by comparing relevant features with the complete, no and partial reference information, respectively. The state-of-the-art FR and RR SIQA methods [19, 20] need at least 1300 bits to represent the reference information, which is a hefty burden on underwater acoustic channels. Third, the training process of the SIQA methods should be lightweight and efficient to the greatest extent on account of the lack of sonar image materials and the real-time demands in many underwater scenarios. The state-of-the-art NR SIQA metric [18] trains and aggregates multiple base learners for overfitting avoidance, the computational complexity of which is many times higher than that utilizing only one learner.

To address this problem, we propose an NR contour-aware model for sonar image quality assessment (CaMSIQA) for acoustic lens and side-scan sonar devices. The major contributions of this paper are summarized in the following three aspects:

1) Proposing an SIQA method that excludes any reference in the pipeline. This is significantly practical in underwater acoustic transmission with limited channel bandwidth.
2) Proposing a contour information statistics (CIS) model to depict the statistics of the main morphology of objects in a sonar image. The CIS model is motivated by the requirements of common application scenarios of sonar images, that is, underwater detection.
3) Quantifying the differences between the parameters of CIS model of distorted and unimpaired sonar images by metric learning. It is a good trade-off between accuracy and computational complexity in contrast to conventional learning machines.

The overall structure of this paper is divided into four parts, including this introduction. The details about the CIS model as well as the metric learning approach are presented in Section 2. Section 3 provides the quantitative and qualitative performance comparisons between the CaMSIQA method and the existing state-of-the-art RR and NR IQA methods. In Section 4, this study is concluded.

2 | METHODOLOGY

In this section, the details of the CaMSIQA method are expounded. The method is proposed based on a corpus of contour-aware features using the conclusion that image contour is significant for underwater detection [18]. Then we find that the statistical regularities of exploited contour-aware features can be observed by $\beta$ fit. Therefore, the CIS model is
derived from the $\beta$ fit of the contour-aware features. The benchmark of the CIS model is established from a collect of undistorted sonar images so that the CaMSIQA can work without unimpaired reference. We also estimate CIS model parameters from the test sonar image. The distance between parameters of these two models are utilized as the indicator of utility quality of the test image. To make distance measurement more accurate, we exploit a supervised metric learning algorithm consisting of dimensionality reduction and Euclidean distance measurement.

### 2.1 Contour-aware features

The importance of structural information of the HVS for semantic objects understanding has been verified in [13]. Although the structure-based IQA has been well researched, existing structure-relevant works perform unsatisfactorily on sonar images. The reasons can be concluded as the indiscrimination of the global and local structures. Specifically, the global structure dominates the object description, while local structure mostly affects the comfort of an image. Thus, instead of taking account of all components of image structures in the SIQA, extracting global structure-based features is a preferable alternative. We extract contour-aware features, which is one of the primary representations of global structure, in accordance with feature extraction framework proposed in [18].

To collect contour information from multi-aspects, the sonar image is firstly transformed into frequency and spatial domain. For frequency domain, the discrete cosine transformation (DCT) and the Cohen-Daubechies-Feauveau 9/7 (C-D-F 9/7) wavelet transform are employed for fixed- and multi-scale analysis. The coefficient matrices of the DCT and the C-D-F 9/7 wavelet transform are denoted as $C_D$ and $C_W$. Meanwhile, the singular value decomposition (SVD) is utilized to depict contour information in spatial domain. We denote $C_J$ as the diagonal matrix of the SVD transform. According to [22] and [27], the elements of $C_D$, $C_W$ and $C_J$ with large values are concise representatives of image contour. Thus, we define the sparsity of these coefficients matrices as the indicator of image contour. Since Gini index and Hoyer measure have varying levels of sensitivity to sparsity, that is, the sensitivity of these two sparsity measures to contour distortion is different [18]. Both the Gini index and Hoyer measure, which outperform other sparsity measures [28], are utilized to more comprehensively describe the contour information. The Gini index and Hoyer measure are defined in Equations (1) and (2):

$$H(n) = \frac{\sqrt{M} - (\sum_{i=1}^{M} |n_i|)}{\sqrt{\sum_{i=1}^{M} n_i^2}} \sqrt{M - 1}, \quad (1)$$

$$G(n) = 1 - 2 \sum_{k=1}^{M} \frac{n_{(k)}}{\|n\|_1} \left( M - k + \frac{1}{2} \right) \frac{1}{M}, \quad (2)$$

where $n = [n_1, n_2, ..., n_M]$ is a vector flattened from a certain coefficient matrix, and $n_{(k)}$ is the $k$-th smallest element of $n$. Ultimately, the image contour is expressed by the Gini index and Hoyer measure of $C_D$, $C_W$ and $C_J$.

### 2.2 CIS model

In this part, we try to capture the inherent properties of extracted features for a universal and concise expression. For optical images, some statistical models have been proposed to capture the statistical behaviour of image features [29, 30]. For example, given a set of natural scene images (NSIs), the statistics of their mean subtracted contrast normalized (MSCN) coefficients can be characterized by an asymmetric generalized gaussian distribution (AGGD) model. Moreover, it has been proved to be held in [31] that the MSCN coefficients have characteristic statistical properties changed by the presence of distortion. Then the NSS model, which captures the aforementioned statistical properties, is established based upon this hypothesis. Since sonar images are visually and practically different from conventional NSIs, the NSS model is inferior for sonar images. Nevertheless, inspired by this pipeline, we also discover the statistical behaviour of the extracted contour-aware features from sonar images.

In order to visualize contour-aware features, Figure 2 plots the histogram of contour-aware features for a set of undistorted sonar images, with some examples shown in Figure 3. From Figure 2, it can be found that the statistics of undistorted sonar images can be captured using $\beta$ distribution. Furthermore, using mean opinion score (MOS), a cluster of $\beta$ function fitted to the histograms with different quality decreases are plotted in Figure 4. From Figure 4, we can find that though quality decrease modifies the statistics of contour-aware features characteristically, the statistical behaviour of distorted contour-aware features can still be well approximated by $\beta$ distribution. It is clear that the peak and the tail behaviours of $\beta$ functions are distinguishable relevant to different quality decreases. From this
β distributions of contour-aware features extracted from sonar images with different quality decreases

observation, the β model is given by:

\[
β(x; a, b) = \frac{x^{a-1}(1-x)^{b-1}}{B(a, b)},
\]

where \( x \) denotes the contour-aware features. The ‘shape’ of β function is controlled by \( a \) and \( b \) in Equation (1), which can be estimated by maximizing a modified likelihood that integrates the zeros or ones by considering them as if they were values that have been left-censored at the square root of the smallest sample or right-censored at \( 1 - \varepsilon \) (\( \varepsilon \) is a small content), respectively [32].

\section{2.3 Metric learning}

To measure the distance between the different CIS parameter tuples, a supervised metric learning algorithm is employed for the CaMSIQA method, which is characterized by marginal Fisher analysis (MFA) [33] and Euclidean distance. The MFA is a dimensionality reduction algorithm which overcomes the limitation of linear discriminant analysis algorithm that the samples of each class should obey a Gaussian distribution. According to the procedure of MFA, we divided the training sonar images into \( \eta \) groups according to the MOS values. Images in the same group were assigned as the same index. The MFA can be stated as four steps.

In the first step, we obtain a transformation matrix \( W_{PCA} \) by projecting the training data into the principal component analysis (PCA) subspace. Then the intraclass compactness and interclass separability graph are constructed by:

\[
\tilde{S}_{c} = 2\omega^T X (D_{c} - W_{c}) X^T \omega,
\]

where \( \omega \) is the projection matrix and \( X \) is the vertex set of a graph whose Laplacian matrix is defined as \( D_{c} \) and \( W_{c} \).

\[
W_{c} = \begin{cases} 1 & \text{if and only if } x_i \text{ and } x_j \text{ is the } k_{c} \text{th nearest between their neighbours,} \\ 0 & \text{otherwise} \end{cases}, \quad \forall i,j.
\]

The interclass separability can be characterized as:

\[
\tilde{S}_{s} = 2\omega^T X (D_{s} - W_{s}) X^T \omega,
\]

where \( W_{s} = \begin{cases} 1 & \text{if and only if } x_i \text{ or } x_j \text{ is the } k_{s} \text{th nearest across different classes and } D_{s} \text{ is defined as} \\ 0 & \text{otherwise} \end{cases}, \quad \forall i,j.
\]

The mean \( \mu \) and the variance \( \delta \) of the β distribution are also useful, which are defined as Equations (6) and (7):

\[
\mu = \frac{a}{a + b},
\]

\[
\delta = \frac{ab}{(a + b + 1)(a + b)^2}.
\]
In the third step, the objective function is given as follows:

$$\omega^* = \arg \min_{\omega} \omega^T X (D - Wc)^T X^T \omega$$

Finally, the projection direction becomes:

$$\omega = W_{\text{PCA}} \omega^*.$$  \hspace{1cm} (13)

Thanks to the projection matrix $\omega$, the parameter vector obtained from the CIS model can be projected into a new space and the distance between the parameter tuples of different CIS models can be measured using the Euclidean distance. The quality of the test sonar image is expressed as the distance between different CIS models:

$$D(\mathbf{p}_r, \mathbf{p}_d) = \sqrt{\sum (p_{ri} - p_{di})^2},$$

where $\mathbf{p}_r$ and $\mathbf{p}_d$ indicate the dimension-reduced parameter tuples extracted form the CIS model for undistorted and test sonar images, respectively.

### 3 EXPERIMENTAL RESULTS

#### 3.1 Sonar image quality database

In [19], Sonar Image Quality Database (SIQD) was constructed for evaluating SIQA indices with adequate ground-truth human subjective data. Thus, all experiments here were conducted on the SIQD database. We tabulate the important information of the SIQD database in Table 1. The SIQD database includes 40 reference images captured by advanced acoustic lens sonar and side-scan sonar in underwater scenarios, with the resolution of $320 \times 320$. The objects of the selected sonar images include most common scenes for underwater detection, such as underwater creatures, swimmers, seabed and shipwrecks.

Totally, there are 800 distorted images in the SIQD database. The distortion of sonar images is caused by the lossy compression, that is, Compression based on Gradient-Based Recovery (ComGBR) coding [34] and Set Partitioning In Hierarchical Trees (SPIHT) coding [35], and manmade bit errors within the bitstreams. The bit error ratios were set in accordance to the latest achievements in underwater acoustic communications [36, 37]. Each type of distortions has four to five distortion levels. Figure 6 shows examples of four types of distortions. As illustrated in Figure 6, the compression coding causes blurs, while the bit errors introduce noise and eccentricity into the sonar image, respectively. All of the aforementioned types of distortions take typical peculiarities of the HVS into account, and are common in practical underwater application scenarios. In subjective tests, all images were divided into 20 groups, each of which was tested by 25 viewers with expertise in underwater target recognition. The Single Stimulus with Multiple Repetition (SSMR) method with a five-category discrete scale were employed to collect subjective scores. The collected scores were firstly converted to Z-scores. Then we rejected scores from the unreliable viewers according to the procedure specified in ITU-R BT 500-11 [38]. After viewer rejection, Z-scores were linearly rescaled to lie in the range of $[0, 100]$. Finally, the mean of the rescaled Z-scores was computed as the MOS value of each test image. In this way, all images in the SIQD database are finally labelled by their MOS values.

#### 3.2 Evaluation protocols

To reveal the effectiveness of proposed approach, comprehensive experiments were conducted on four popular IQA benchmarks. Selected IQA benchmarks can be classified into three groups according to their physical meanings. The first group includes the Spearman rank-order correlation coefficient (SROCC) and the Kendall rank-order correlation coefficient (KROCC) for prediction monotonicity measurement defined as...
Equations (15) and (16):

\[
SROCC = 1 - \frac{6 \sum_{i=1}^{M} d_i^2}{M(M^2 - 1)}, \quad (15)
\]

\[
KROCC = \frac{2(M - M_d)}{M(M - 1)}, \quad (16)
\]

where \(d\) indicates the discrepancy between the rank of MOS value and predicted objective quality for the \(i\)th sonar image. \(M\), \(M_d\), and \(M_r\) represent the number of test sonar images, the concordant and the discordant sonar image pairs in the testing set, respectively. These indices operate only on the ranks of the sample points without considering the relative distance between sample points. The second group of IQA benchmarks is relevant to the prediction accuracy. The prediction accuracy is indicated by the Pearson linear correlation coefficient (PLCC) between MOS values and the mapped objective scores \(Q_m\), that is,

\[
PLCC = \frac{\sum_{i=1}^{M} (Q_m - \bar{Q}_m)(s_i - \bar{s})}{\sqrt{\sum_{i=1}^{M} (Q_m - \bar{Q}_m)^2 \cdot \sum_{i=1}^{M} (s_i - \bar{s})^2}}, \quad (17)
\]

where \(Q_m\) and \(\bar{Q}_m\) are the mapped objective score for the \(i\)th sonar image and the mean of all \(Q_m\), \(s_i\) and \(\bar{s}\) represent the MOS value of the \(i\)th sonar image and the mean of all \(s_i\), respectively. The mapped objective score can be obtained as:

\[
Q_m = \theta_1 \left( \frac{1}{2} - \frac{1}{1 + \exp(\theta_2(Q_p - \theta_3))} \right) + \theta_4 Q_p + \theta_5, \quad (18)
\]

where \(Q_p\) and \(Q_m\) are the predicted objective score from the corresponding IQA index and the mapped objective scores after the non-linear regression, respectively. \(\theta_i, i = 1, 2, \ldots, 5\) are parameters to be fitted during the curve fitting process. The IQA index in the third group is the root mean square error (RMSE) as depicted in Equation (19), which quantifies the differences between \(Q_m\) and MOS values for prediction consistency measurement.

\[
RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (Q_m - s_i)^2}, \quad (19)
\]

Based on the abovementioned four benchmarks, we compared the performance of our method with the state-of-the-art NR and RR IQA methods proposed for NSIs and screen content images (SCIs). Besides, SIQA methods for side-scan and acoustic lens sonar images were also compared with the CaMSIQA method. We implemented them through the source codes provided by the authors and adopted their default settings if feasible. The 11 selected NR IQA methods are BLInd Image Integrity Notator using DCT Statistics II (BLINDS II) [2], Accelerated Screen Image Quality Evaluator (ASIQE) [23], Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [31], No-reference Free Energy based Robust Metric (NFERM) [39], AR-based Image Sharpness Metric (ARISM) [40], Blind Quality Measure for Screen content images (BQMS) [41], Integrated Local Natural Image Quality Evaluator (ILNIQE) [42], Blind Pseudo Reference Image based quality measure (BPR) [43], blind IQA method based on High Order Statistics Aggregation (HOSA) [44], Six-Step BLInd Metric (SISBLIM) [45] and NF IQA in Curvelet domain (CurveletQA) [46]. Three widely cited RR IQA methods have been compared with the proposed method, they are Reduced-reference Image Quality Metric for Contrast Change (RIQMC) [25], Reduced-reference Wavelet-domain Quality Measure of Screen content images (RWQMS) [26] and Quality Metric Contrast (QMC) [47]. The SIQA methods chosen for comparison are No-Reference Contour Degradation Measurement (NRCDM) [18], FR Sonar Image Quality Predictor (SIQP) [19] and Partial-reference Sonar Image Quality Predictor (PSIQP) method [20].

Based on the previously mentioned IQA methods, we compared the performances of selected IQA methods on the SIQD database, wherein 4/5 is used for training and the rest is used for testing. There are 40 distorted sonar images in each group, every 20 of whom are of the same content. To exclude the performance bias, random choice of training and testing sets was
repeated for 1000 times. The average performance across 1000 iterations is finally reported for comparisons.

### 3.3 Quantitative performance comparison

In this study, we utilize metric learning to measure the distance between the CIS model parameter tuples of the test and undistorted sonar images. Based on the metric learning, the model parameters are firstly projected into a new subspace for dimensionality reduction, wherein the transform matrix is optimized by the training samples. We show that the proposed method driven by metric learning significantly outperforms the conventional distance metrics. The results are tabulated in Table 2, wherein the better performances are highlighted in boldface. When metric learning is excluded from the framework, the Euclidean distance is directly applied between parameters of different models. From Table 2, it is obvious that applying metric learning leads to a performance gain, about 20% for the SROCC, 24% for the KROCC, 46% for the PLCC and 21% for the RMSE. The significant performance gain achieved by metric learning demonstrates the superiority due to its good separability for samples with different characteristics.

Table 2: Performance comparisons with and without metric learning

| Metric Learning | SROCC   | KROCC   | PLCC    | RMSE    |
|-----------------|---------|---------|---------|---------|
| Without         | 0.593   | 0.412   | 0.500   | 11.551  |
| With            | 0.714   | 0.512   | 0.731   | 9.120   |

Due to the pipeline of the CaMSIQA method, it works in an NR approach. We now compare the performance of the proposed method with that of other NR IQA methods on the testing set, where the images and the distortion types are completely separated from the training set. The results are reported in terms of the selected performance metrics in Table 3, where the best and the next-best performances are boldfaced and underlined, respectively. From Table 3, we can conclude that CaMSIQA performs better than all of the selected state-of-the-art NR IQA algorithms. It should be emphasized that only CaMSIQA method can attain performance greater than 0.7 for the SROCC and the PLCC, which are the most common performance metrics. With respect to the prediction consistency, the proposed method also achieves the best performance. Compared with the second-ranking method in Table 3, the performance gain of the CaMSIQA method is approximately 11% for the SROCC, 8% for the KROCC, 9% for the PLCC and 2% for the RMSE.

Moreover, Table 4 tabulates the performance metrics of the CaMSIQA method as well as the selected widely cited RR IQA methods. We can see from Table 4 that the proposed method performs the best regarding to the prediction monotonicity among the SIQD database. As for the prediction accuracy and consistency, the RWQMS performs better than all the other competitors. The RWQMS method is an RR SIQA method, which needs 24 features as reference information. In practice, at least 480 bits are needed for accurate representation of the reference information. As for the proposed method, no reference is required, which is a very important improvement for underwater acoustic transmissions due to the limited bandwidth of underwater acoustic channels. Comprehensively considering Tables 3 and 4, the CaMSIQA beats all competing IQA methods. The performance gain achieved is due to the fact that our method mainly concerns the global structures of images which are directly related to the utilities of sonar images. In contrast, most common used IQA methods for NSIs or SCIs were designed either based on statistics or on visual characteristics.

We have also conducted the $F$-test between the proposed work and the compared IQA methods. We define the critical threshold $F_c$ according to the value of residuals of the compared IQA methods and confidence level which is assumed as 95%. Then the ratio of the prediction residual variances of objective qualities to MOS values is expressed as $F_t$, where the objective qualities are regressed using the five-parameter logistic nonlinear regression function. When $F_t > F_c$, two IQA methods participating in the comparison can be regarded as significantly distinct. Table 5 tabulates the results of the significance comparison. In Table 5, “+1” represents that the CaMSIQA method is significantly better than the compared method, “0” shows that the performances of the CaMSIQA method and the compared method are statistically undistinguishable, while “−1” indicates that the CaMSIQA method is statistically worse than the compared method. It can be concluded that the CaMSIQA method...

Table 3: Performance comparisons of NR IQA algorithms on SIQD database

| IQA       | SROCC   | KROCC   | PLCC    | RMSE    |
|-----------|---------|---------|---------|---------|
| BRISQUE   | 0.641   | 0.470   | 0.663   | 9.318   |
| BLINDS II | 0.389   | 0.281   | 0.514   | 10.702  |
| NFVLRD    | 0.401   | 0.294   | 0.538   | 10.515  |
| ARISM     | 0.127   | 0.089   | 0.556   | 10.451  |
| BQMS      | 0.486   | 0.363   | 0.651   | 9.642   |
| IL-NIQE   | 0.523   | 0.386   | 0.629   | 9.716   |
| ASIQE     | 0.379   | 0.286   | 0.669   | 9.440   |
| BPI       | 0.386   | 0.280   | 0.528   | 10.751  |
| HOSA      | 0.572   | 0.429   | 0.591   | 9.774   |
| SISBLIM   | 0.422   | 0.317   | 0.646   | 9.648   |
| CurvObsQA | 0.225   | 0.151   | 0.280   | 12.900  |
| CaMSIQA   | 0.714   | 0.512   | 0.731   | 9.120   |

Table 4: Performance comparisons of RR IQA algorithms on SIQD database

| IQA       | SROCC   | KROCC   | PLCC    | RMSE    |
|-----------|---------|---------|---------|---------|
| RWQMS     | 0.643   | 0.480   | 0.772   | 7.937   |
| QMC       | 0.564   | 0.428   | 0.733   | 8.130   |
| RIQMC     | 0.379   | 0.284   | 0.716   | 8.683   |
| CaMSIQA   | 0.714   | 0.512   | 0.731   | 9.120   |
To compare the performance between SIQA methods, we exhibit performance metrics indicating the prediction monotonicity (the SROCC and the KROCC) and the prediction accuracy (the PLCC) in a radar chart as shown in Figure 7. There are three axes in the radar chart for the SROCC, the PLCC and the KROCC. We mark the scale of three performance metrics in the centring axis. It can be clearly observed that the SIQP metric outperforms other SIQA methods with regard to the prediction monotonicity, and comparable with the PSIQP metric concerning prediction accuracy among selected test images. Since NR methods lose some prior knowledge of the content, they are universally inferior to the reference-based SIQA methods. However, due to the efficient distance measurement employed here, this study achieves comparable performance with the NRCDM metric using lower computational complexity. The processing time of methods fall under the last category for one sonar image exceeds 1 s. The last category includes five methods. It is clearly that the computational cost of the CaMSIQA method is in the leading position compared with most of the selected competing NR IQA methods. Only the BRISQUE metric and the SISBLIM metric take less processing time than the CaMSIQA method. As for these IQA methods with lower computational time, the proposed method can obtain at least 11%, 8.9%, 10% and 2% performance gain in terms of the SROCC, the KROCC, the PLCC and the RMSE, respectively.

### 3.4 Qualitative performance comparison

To visually compare the performance of the proposed method and the competing IQA approaches, we illustrated the differences between the objective quality predicted by each selected method and the MOS value with an example of the distorted sonar image in Figure 8. Figure 8(a) shows the reference sonar image, which is captured by a acoustic lens sonar for a nurse shark during Fabien Cousteau’s Mission 31 ground breaking underwater study aboard the Aquarius. We exhibited the average run-time of each method in Table 7. This computational cost tests were conducted on the software platform of MATLAB R2019b on a computer with 3.60 GHz CPU processor and 64.00G RAM. Selected methods in Table 7 can be classified into four categories according to their computational costs. Methods in the first category take less than $10^{-3}$ st to process one sonar image. The BRISQUE and SISBLIM make up this category. The second category includes method with average run-time less than $10^{-1}$, and our CaMSIQA falls under this category. Four methods can be classified into the third category since they all take less than 1 s for quality prediction per image. The last category includes five methods. Similar performance comparison between the existing SIQA methods was tabulated in Table 6.
distorted version of this image, whose MOS value is 30.45, in Figure 8(b). The most typical distortion existed in this sonar image is blur and non-eccentricity distortion with regard to SPIHT compression, which can affect the identification of rock and shark. The quality evaluated by the CaMSIQA method is very close to the MOS value provided by the SIQD database. Furthermore, we also tabulate the objective qualities predicted by selected methods in Figure 9, which have been mapped to the range [1,100]. The horizontal axis of Figure 9 indicates the NR IQA methods, which are indexed according to Table 8. The vertical axis is the predicted quality. The MOS value of this sonar image is indicated by red line. It is obvious that the proposed method is able to explore the impacts of those distortions on the utility quality and quantize them to the quality score. By contrast, the selected NR IQA methods all overestimate the quality for this image sample. This is because these IQA methods are designed for NSIs or SCIs without considering the visual characteristics based on corresponding utility for sonar images.

4 | CONCLUSION

We have proposed a contour-based approach for NR SIQA using a statistics model (i.e. the CIS model). Based on this model, the contour information is expressed as a set of parameters of Beta fit. The quality of the test sonar image is quantified by measuring the distances between the parameter tuples of the CIS models of the test and the undistorted sonar images in a proper subspace. We have validated that our method correlates highly with human decision of utility quality for sonar images by four common-used performance metrics. Besides, the proposed CaMSIQA method is also well behaved in comparison with the state-of-the-art NR IQA methods, and is even highly comparative against the performance with the widely cited RR IQA methods. Furthermore, how to extend the CaMSIQA method to an unsupervised approach without performance loss is an interesting and challenging problem yet to be explored.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (61901119,62001404).

REFERENCES

1. Qin, M., et al.: Hybrid NSS features for no-reference image quality assessment. IET Image Proc. 11(6), 443–449 (2017)
2. Saad, M.A., et al.: Blind image quality assessment: A natural scene statistics approach in the DCT domain. IEEE Trans. Image Process. 21(8), 3339–3352 (2012)
3. Yang, X., Wang, T., Ji, G.: No-reference image quality assessment via structural information fluctuation. IET Image Proc. 14(2), 384–396 (2020)
4. Liu, Z. et al.: Structured entropy of primitive: Big data-based stereoscopic image quality assessment. IET Image Proc. 11(10), 854–860
5. Tang, Z., et al.: Full-reference image quality assessment by combining features in spatial and frequency domains. IEEE Trans. Broadcast. 65(1), 138–151 (2019)
6. Liu, Z. et al.: Structured entropy of primitive: Big data-based stereoscopic image quality assessment. IET Image Proc. 11(10), 854–860
7. Rahul, K., Tiwari, A.K.: FQI: Feature-based reduced-reference image quality assessment method for screen content images. IET Image Proc. 13(7), 1170–1180 (2019)
8. Gu, K., et al.: Learning a no-reference quality assessment model of enhanced images with big data. IEEE Trans. Neural Networks Learn. Syst. 29(4), 1301–1313 (2018)
9. Bosse, S., et al.: No-reference image quality metric based on multiple deep belief networks. IET Image Proc. 13(6), 1321–1327 (2019)
10. Chen, W., et al.: Subjective and objective quality evaluation of sonar images for underwater acoustic transmission. 2017 IEEE International Conference on Image Processing, pp. 176–180. Beijing, China (2017)
1. Rouse, D.M., et al.: Estimating the usefulness of distorted natural images using an image contour degradation measure. J. Opt. Soc. Am. A. 28(2), 157–188 (2011)

2. Williams, D.P.: Image-quality prediction of synthetic aperture sonar imagery. 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 2114–2117. Dallas, TX, USA (2010)

3. Debes, C., et al.: Quality assessment of synthetic aperture sonar images, pp. 1–4. OCEANS 2009-EUROPE, Bremen (2009)

4. Kalwa, J., Madsen, A.: Sonar image quality assessment for an autonomous underwater vehicle. In: Proceedings World Automation Congress, pp. 33–38. Seville, Spain (2004)

5. Zhu, M., et al.: Sonar system of Jiaolong human-occupied vehicle. Chin. Sci. Bull. 59(35), 3462–3470 (2014)

6. Chen, W., et al.: Reference-free quality assessment of sonar images via contour degradation measurement. IEEE Trans. Image Process. 28(11), 5336–5351 (2019)

7. Chen, W., et al.: Statistical and structural information backed full-reference quality measure of compressed sonar images. IEEE Trans. Circuits Syst. Video Technol. 30(2), 334–348 (2020)

8. Chen, W., et al.: Partial-reference sonar image quality assessment for underwater transmission. IEEE Trans. Aerosp. Electron. Syst. 56(6), 2776–2787 (2018)

9. Gu, K., et al.: Saliency-guided quality assessment of screen content images. IEEE Trans. Multimedia 18(6), 1098–1110 (2016)

10. Gu, K., et al.: A fast reliable image quality predictor by fusing micro- and macro-structures. IEEE Trans. Ind. Electron. 64(5), 3903–3912 (2017)

11. Gu, K., et al.: No-reference quality assessment of screen content pictures. IEEE Trans. Image Process. 26(8), 4005–4018 (2017)

12. Gu, K., et al.: Model-based referenceless quality metric of 3D synthesized images using local image description. IEEE Trans. Image Process. 27(1), 394–405 (2018)

13. Gu, K., et al.: The analysis of image contrast: From quality assessment to automatic enhancement. IEEE Trans. Cybern. 46(1), 284–297 (2016)

14. Wang, S., et al.: Subjective and objective quality assessment of compressed screen content images. IEEE J. Emerging Sel. Top. Circuits Syst. 6(4), 532–543 (2016)

15. Dan, K.: Singularly valuable decomposition: The SVD of a matrix. Coll. Math. J. 27(1), 2–23 (1996)

16. Hurley, N., Rickard, S.: Comparing measures of sparsity. IEEE Trans. Inf. Theory 55(10), 4723–4741 (2009)

17. Chen, Z., et al.: Blind stereoscopic video quality assessment: from depth perception to overall experience. IEEE Trans. Image Process. 27(2), 721–734 (2018)

18. Yan, B., et al.: Naturalness-aware deep no-reference image quality assessment. IEEE Trans. Multimedia 21(10), 2603–2615 (2019)

19. Mittal, A., Moorby, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process. 21(12), 4695–4708 (2012)

20. Gu, K., et al.: No-reference image sharpness assessment in autoregressive parameter space. IEEE Trans. Image Process. 24(10), 3218–3231 (2015)

21. Gu, K., et al.: Learning a blind quality evaluation engine of screen content images. Neurocomputing 196(C), 140–149 (2016)

22. Zhang, L., Zhang, L., Bovik, A.C.: A feature-enriched completely blind image quality evaluator. IEEE Trans. Image Process. 24(8), 2579–2591 (2015)

23. Min, X., et al.: Blind quality assessment based on pseudo reference image. IEEE Trans. Multimedia 20(8), 2049–2062 (2018)

24. Xu, J., et al.: Blind image quality assessment based on high order statistics aggregation. IEEE Trans. Image Process. 25(9), 4444–4457 (2016)

25. Gu, K., et al.: Hybrid no-reference quality metric for singly and multiply distorted images. IEEE Trans. Broadcast. 60(3), 555–567 (2014)

26. Liu, L., et al.: No-reference image quality assessment in curvelet domain. Signal Process. Image Commun. 29(4), 494–505 (2014)

27. Gu, K., et al.: Automatic contrast enhancement technology with saliency preservation. IEEE Trans. Circuits Syst. Video Technol. 25(9), 1480–1494 (2014)

How to cite this article: Chen W, Lin R, Zhang R, Zhu Y. A blind contour-aware quality model for sonar images. IET Image Process. 2021;1–10. https://doi.org/10.1049/ipt.12202