No-Reference Image Quality Assessment and Application Based on Spatial Domain Coding

Chen Yong¹, Fang Hao¹, Liu Huanlin²

¹ Key Laboratory of Industrial Internet of Things & Network Control, Ministry of Education, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China.
² School of Telecommunications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China.

Corresponding author: Chen Yong (e-mail: chenyong@cqutp.edu.cn).

This research was funded by the Chongqing Education Committee Science of China for supporting the Foundation of program, No. KJ1400434.

ABSTRACT Aiming at the problem that no-reference image quality assessment algorithm is not accurate enough to predict the image quality at present, we proposed a method of no-reference image quality assessment based on spatial domain coding (SDC). We extracted the spatial structure features of the image on different bit planes. The extracted features quantify the structural information between pixels, which can more accurately reflect the distortion degree of the image. We used neural network to establish an image quality assessment model. The experimental results show that the proposed image assessment algorithm is more accurate than the existing mainstream image quality assessment algorithms and highly consistent with subjective perception of human eyes. Finally, the proposed algorithm is applied to the auto focus of the camera verifying the practicability and accuracy of the algorithm.

INDEX TERMS Local binary patterns, No-reference image quality assessment, Neural network

I. INTRODUCTION

Image is an important medium for the expression and transmission of information. However, different types and different degrees of distortion occur during the transmission and storage of digital image processing, which may have a serious impact on the integrity and intelligibility of the information in the image. Therefore, it’s very necessary to establish an effective image quality assessment (IQA) mechanism [1].

The existing IQA can be divided into the following three methods: full-reference (FR), reduced-reference (RR) and no-reference (NR) [2]. FR-IQA algorithm can obtain the image quality more accurately, but it needs to get undistorted original image as a reference in advance, and its scope of application is relatively narrow. RR type is proposed to solve the problem of IQA under wireless transmission conditions, this method needs to extract a small amount of original image features in the assessment of images, which limit the application. NR-IQA methods have the widest application prospect for it can predict the visual quality of any input image without original image. In the vast majority of practical applications, the original image of the distorted image is difficult to obtain. Therefore, the research of NR-IQA is very important.

Most of the existing NR-IQA methods are based on statistical natural features to establish the IQA model. Ref. [3] extracted features from frequency domain by wavelet transform and proposed a two-step framework combining prediction with classification. However, this method is not accurate enough for image quality assessment. Mittal et al. [4] used the generalized Gaussian distribution modeling to fit the histogram of mean subtracted contrast normalized (MSCN), then used model parameter as the feature to assess the quality of image. Yan et al. [5] utilized the singular value decomposition and asymmetric generalized Gaussian distribution (GGD) modeling to obtain natural redundancy statistics and used the natural redundancy statistics to capture the distortion degree. Song et al. [6] have found that using Gaussian model to fit histogram will cause information loss and reduce the accuracy of features, so they directly used the MSCN coefficient histogram as the feature and utilized Support Vector Machine (SVM) to establish the IQA model. Zhang et al. [7] used generalized local binary patterns (LBP) to extract image features and use SVM to evaluate image quality. Li et al. [8] combined LBP and MSCN, then proposed a NR-IQA method based on structure and brightness statistics. Ref. [9] utilized the local quantized pattern (LQP) to extract image features and also used SVM to establish the IQA model. Ref. [10]
applied LBP to the sub-bands of wavelet decomposed image separately, then use the histogram of LBP as the features to assess image quality. In fact, there are still some problems of the improvement of LBP in ref. [7-10]. Particularly when they extracted structure features of the images, the used threshold function causes information loss. Liu et al. [11] utilized a relative gradient magnitude feature which accounts for perceptual masking and utilized back-propagation neural network (BP-NN) to map the image features to image quality. They proved that BP-NN performs better than support vector regression (SVR) in image quality assessment. Yang et al. [12] studied the correlation between features to image quality. They proved that BP-NN performs which accounts for perceptual masking and utilized back-propagation neural network (BP-NN) to map the image features to image quality. They proved that BP-NN performs better than support vector regression (SVR) in image quality assessment.

But the structural information between pixels is very important for the description of image quality. As can be seen from Fig.1-(a) the local image is smoother, and the pixels in (b) are sharper, although the LBP algorithm gets the binary patterns are exactly the same. However, the original LBP algorithm is used to get texture information of the image and is not accurate enough to describe the relationship between pixels. Specifically, it will be judged as 1/0 regardless of whether the neighboring pixel is larger/smaller than the center pixel. But the structural information between pixels is very important for the description of image quality. As can be seen from Fig.1-(a) the local image is smoother, and the pixels in (b) are sharper, although the LBP algorithm gets the binary patterns are exactly the same.

**II. FEATURE EXTRACTION**

**A. LBP ALGORITHM**

Proposed by Ojala et al. [13], LBP has rotation invariance and can extract the structure information of the image. It has been widely used in face recognition, image retrieval, image segmentation and other fields. The LBP algorithm compares the gray value of each pixel with its 8 neighborhood pixels to get the structure information of the image pixels. The LBP code of one pixel is deduced as

$$LBP = \sum_{p=1}^{P} s(g_p - g_c)2^{p-1}$$

where \( g_c \) is the gray value of the central pixel in the 3×3 block of the image, the value of \( P \) is 8, \( g_p \) is the gray value of the 8 neighborhood pixels of the central pixel in the block, and \( s \) is the threshold function, which is defined as

$$s(g_p - g_c) = \begin{cases} 1 & g_p - g_c \geq 0 \\ 0 & g_p - g_c < 0 \end{cases} \quad (2)$$

After LBP encoding, the structural features of the image are extracted [14]. The features of the image can be obtained by gathering the LBP coefficients of each sub pixel. After that, the author proposed the following method for gray-scale and rotation invariant texture description

$$LBP^{nu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP) \leq 2 \\ P+1 & \text{otherwise} \end{cases} \quad (3)$$

where

$$U(LBP) = \| s(g_{p-1} - g_c) - s(g_0 - g_c) \|
\begin{align*}
&+ \sum_{p=1}^{P-1} [s(g_p - g_c) - s(g_{p-1} - g_c)]
\end{align*} \quad (4)$$

Superscript \( nu2 \) reflects the use of rotation invariant patterns that have \( U \) value of at most 2. At the same time, the dimension of LBP coefficient was reduced from 256 to 10.

**B. GLBP ALGORITHM**

From Eq. (2), it can be seen that each LBP coefficient is obtained by comparing the center pixel with the neighborhood pixels. However, the original LBP algorithm is used to get texture information of the image and is not accurate enough to describe the relationship between pixels. Specifically, it will be judged as 1/0 regardless of whether the neighboring pixel is larger/smaller than the center pixel. But the structural information between pixels is very important for the description of image quality. As can be seen from Fig.1-(a) the local image is smoother, and the pixels in (b) are sharper, although the LBP algorithm gets the binary patterns are exactly the same.

**FIGURE 1.** Example of local binary pattern to get the inter pixel structure information. (a) Local smooth structure image. (b) Local sharp structure image.

In general, the original LBP algorithm obtains the features of the image through two threshold functions. The method used the first threshold function (see Eq.(2)) to
extract image structure information, but this method needs to be improved if we apply it to IQA. The original LBP reduced the dimension of the features through another threshold function (see Eq. (3)) and reduced the feature dimension to 10, but this is also unadvisable. Although Ref. [7-10] also reduced the dimension of features, but they still used threshold function which caused information loss. How to reduce dimension without using threshold function and ensuring the accuracy of feature is the key to solve the problem, so we proposed the gauss local binary patterns (GLBP). We obtained eight different bit-planes by binary coding on the gray value of the image in spatial domain. The coding method is defined as

\[
B_i(x, y) = \text{dec2bin}(I(x, y), 8)
\]

(5)

where \( I(x, y) \) is the image to be encoded, \( i \) is the binary digit of the gray value, \( B_i(x, y) \) is the \( i \)-th binary plane obtained after binary coding. The binary weight of pixels in \( B_i \) is 0 or 1. The range of image gray value in the image databases is 0-255 and the gray value of each pixel is stored in 8-bit binary digits in the computer, so there will be 8-bit planes. The same pixel can get the same binary code, and the bigger the difference shows between binary codes, the bigger the difference exists in pixels. Therefore, the difference between pixels can be quantified by comparing the binary encoding of pixels. We made sliding windows for the bit-plane. The window size is \( 3 \times 3 \). The GLBP coefficient for each pixel in the bit-plane is deduced as

\[
B_{i,\text{GLBP}} = \sum_{p=1}^{P}(g_p - g_c)
\]

(6)

where \( g_c \) is the center pixel’s binary weight of the \( 3 \times 3 \) blocks, \( g_p \) is the binary weight of 8 neighborhood pixels of the central pixel in the block. In our method, we construct a Gauss-Laplace operator to implement operations in Eq. (6), so we call this method GLBP. The operator \( G \) is defined as

\[
G = \begin{bmatrix}
-1 & -1 & -1 \\
-1 & 8 & -1 \\
-1 & -1 & -1
\end{bmatrix}
\]

(7)

\[G = \begin{bmatrix}
-1 & -1 & -1 \\
-1 & 8 & -1 \\
-1 & -1 & -1
\end{bmatrix}\]

\[G = \begin{bmatrix}
-1 & -1 & -1 \\
-1 & 8 & -1 \\
-1 & -1 & -1
\end{bmatrix}\]

\[G = \begin{bmatrix}
-1 & -1 & -1 \\
-1 & 8 & -1 \\
-1 & -1 & -1
\end{bmatrix}\]

\[G = \begin{bmatrix}
-1 & -1 & -1 \\
-1 & 8 & -1 \\
-1 & -1 & -1
\end{bmatrix}\]
We take the convolution of bit-plane with the operator to implement the same operation as Eq. (6). This can be used to calculate the GLBP coefficient easier and faster. We merged each bit-plane $B_{GLBP}$ and got a 3-dimension matrix. The merge process is defined as

$$I_{GLBP}(x,y,z) = (B_{1,GLBP}, B_{2,GLBP}, B_{3,GLBP}, \ldots, B_{8,GLBP})$$  \hspace{5pt} (8)$$

where $I_{GLBP}(x,y,z)$ is a 3-dimension matrix with the height of 8, its long and wide are the same as the original image $I(x,y)$. After obtaining the GLBP coefficients of the image $I(x,y)$, Eq. (9) can be used to characterize the image statistical information

$$SH(k) = \frac{1}{8MN} \sum_{x=1}^{M} \sum_{y=1}^{N} \sum_{z=1}^{8} f(I_{GLBP}(x,y,z), k),$$  \hspace{5pt} (9)$$

where $M$ and $N$ are the length and width of $I_{GLBP}(x,y,z)$, $f(a,b)$ is the comparison function and $k$ is an integer between $\min(I_{GLBP}(x,y,z))$ and $\max(I_{GLBP}(x,y,z))$. By comparing the GLBP coefficients in $I_{GLBP}(x,y,z)$ with $k$, we obtained 17-dimensional feature vector from $SH(k)$.

It can be seen that the GLBP does not need threshold function, so that the complexity of the algorithm is reduced and this method can more accurately describe the structural information between neighborhood pixels. In addition, it can be obtained from Eq. (6) that

$$\begin{cases} 
\min(I_{GLBP}(x,y,z)) = -8 \\
\max(I_{GLBP}(x,y,z)) = 8 
\end{cases} \hspace{5pt} (10)$$

In fact, this is a method of dimensionality reduction at the expense of increasing the spatial dimension and the feature dimension of GLBP is not lower than that of LBP. However, we haven’t used threshold function in this dimension reduction process, and the accuracy of the feature is ensured. Compared with the classical LBP algorithm, the proposed GLBP also realized dimensionality reduction and obtained 17 dimensional features. Meanwhile, it also has lower complexity.

![Image](image-url)

**FIGURE 3.** The GLBP feature distribution of images with varying degrees of distortion: (a) is the feature distribution of FF distorted images, (b) is the feature distribution of JPEG distorted images, (c) is the feature distribution of JP2K distorted images, (d) is the feature distribution of Gblur distorted images, (e) is the feature distribution of WN distorted images.

We take the "statue" image in the LIVE image database [15] as an example. LIVE image database provides the Difference Mean Opinion Score (DMOS) of the images in the database. The range of DMOS value is 0-100. The smaller the DMOS value of the image, the better the subjective observation quality. This database contains 29 high-resolution RGB images and their distorted images, including 169 JPEG compression (JPEG) images, 175 images of JPEG2000 compression (JP2K), 145 images of Gauss blurring (Gblur), 145 images of white noise (WN) and 145 images of fast fading (FF).

Some different distortion type images of "statue" in the LIVE image database and its 17-dimension GLBP features distribution are shown in Fig. 2. It can be seen that the GLBP
probability histogram of the "statue" image shows a regular distribution. The middle part is high and the two ends are low. FF, JPEG, JP2K and Gblur distortion make this distribution more concentrated towards the middle, and the WN distortion makes this distribution divergent to both ends. On the one hand, the smaller the absolute value of the GLBP coefficient means that the image pixels contain less information. On the other hand, the first four distortions will lose image information, and WN distortion will produce noise that makes the image's microstructure more abundant.

In order to observe the relationship between the features extracted by the GLBP and the degree of image distortion, we did some further experiments. We found that with the increase of distortion, the 17-dimensional GLBP features of JPEG, JP2K, FF and Gblur distorted images gradually move closer to the middle. As the distortions intensify, the 17-dimensional GLBP features of WN distorted images gradually disperse to both ends (see Fig.3). It shows that the features are highly correlated with the image quality. Therefore, a mathematical model between feature and the degree of distortion can be established, so that the image quality can be predicted.

![Image Quality Assessment Model](image1.png)

**FIGURE 4.** Image quality assessment model. The input of the model is the image to be tested and the output is the image quality. C1 are bit-planes. C2 are the GLBP features. C3 is the hidden layer. C4 is the output layer.

### III. ASSESSMENT MODEL

Artificial neural network [16] simulates the structure and function of human brain through a large number of neurons. It belongs to the field of artificial intelligence and can be used in a wider scope. It can not only be applied to intelligent recognition, but also excellently performed in prediction or assessment [17].

We use BP-NN to establish a FR-IQA model between the features and the DMOS of the image (see Fig.4). C1 are bit-planes after binary encoding, C2 are the 17-dimensional GLBP features \(x_1, x_2, \ldots, x_{17}\) extracted from the bit-planes, and C3 is the hidden layer (In our experiment, the hidden layer is set to 3 layers), C4 is the output layer, and "+1" represents the bias node. \(a_i\) represents the activation value of the \(i\) unit in the layer \(l\). The activation value is deduced as

\[
a_i^{l+1} = f \left( \sum_{j=1}^{n} W_{ij} a_j^l + b_i^l \right)
\]

where \(W_{ij}\) represents the connection weight between the \(j\) unit in the \(l\) layer and the \(i\) unit of the \(l+1\) layer, \(b_i^l\) is the paranoid of the \(i\) unit in the \(l+1\) layer, \(n\) is the number of units in the \(l\) layer. When \(l=1\), \(a_i^1 = x_i\). C4 is the output layer. The quality of the image can be obtained by the output layer.

### IV. TEST AND ANALYSIS

In order to test the performance of the proposed algorithm for image quality assessment, the performance tests are conducted in LIVE, CSIQ [19], TID2013 [20] and The LIVE In the Wild Image Quality Challenge (CLIVE) [21, 22] image databases.

CSIQ image quality database contains 866 quality annotated images. There are 30 reference images in the image database and these reference images are distorted by 6 kinds of distortion. We tested 4 kinds of distortion types in the database, including JPEG, JP2K, Gblur and WN. The obtained DMOS values lie in the range 0-1, where a lower value indicates better visual quality.

TID2013 image database contains 3000 quality annotated images. These images contain 24 types of distortion and there are 25 reference images for each distortion type. Each reference image has 5 levels of distortion in each distortion type. The distortion type of TID2013 image database is the most, which contains most of the distortion types in real life. This makes the TID2013 a more challenging database. The obtained DMOS values lie in the range 0-9, where larger MOS indicate better visual quality.

The CLIVE image database contains 1169 images taken in real life conditions, as well as the various objects and scenes that are captured by different cameras in various luminance conditions. In this sense, the images in the database are a mixture of various distortions in the life. The DMOS value is within a range of 0-100, and higher value implies higher quality.

We tested the correlation between the predicted DMOS and subjective DMOS, the accuracy of prediction and the
complexity of the algorithm in the databases. The proposed algorithm is compared with some mainstream FR-IQA and NR-IQA methods.

A. CONSISTENCY EXPERIMENT

This section mainly tests the accuracy of the proposed IQA algorithm and the consistency with the subjective score of the human eye. We used the proposed algorithm to predict the quality of the image in the image databases. We randomly selected 80% of each type of distorted images as a training set, used the remaining 20% as the test set. We calculate performance indicators on all test sets and iterate 1000 times to take the average value as the final standard.

An acknowledged monotonic logistic function [7] is used to provide a nonlinear mapping between objective scores and subjective scores. Scatter plots of the predicted DMOS and objective DMOS in LIVE database are shown in Fig. 5. The closer to the center line, the better the prediction results are. It can be seen that the proposed assessment method has a good prediction effect for different types of distortion. In general, the proposed algorithm has good prediction ability for image quality.

![Image](http://www.ieee.org/publications_standards/publications/rights/index.html for more information.)

**FIGURE 5.** The predicted DMOS and subjective DMOS in LIVE database: (a) is the scatter plot of JP2K distorted images. (b) is the scatter plot of JPEG distorted images. (c) is the scatter plot of FF distorted images. (d) is the scatter plot of Gblur distorted images. (e) is the scatter plot of WN distorted images. (f) is the scatter plot of all the distorted images.

### TABLE I

<table>
<thead>
<tr>
<th>Methods</th>
<th>LCC</th>
<th>SROCC</th>
<th>KROCC</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>JP2K</td>
<td>0.968</td>
<td>0.962</td>
<td>0.838</td>
<td>6.402</td>
</tr>
<tr>
<td>JPEG</td>
<td>0.976</td>
<td>0.964</td>
<td>0.853</td>
<td>6.323</td>
</tr>
<tr>
<td>Gblur</td>
<td>0.975</td>
<td>0.963</td>
<td>0.852</td>
<td>6.923</td>
</tr>
<tr>
<td>FF</td>
<td>0.918</td>
<td>0.913</td>
<td>0.808</td>
<td>7.905</td>
</tr>
<tr>
<td>WN</td>
<td>0.983</td>
<td>0.974</td>
<td>0.866</td>
<td>6.182</td>
</tr>
<tr>
<td>All</td>
<td>0.962</td>
<td>0.956</td>
<td>0.832</td>
<td>7.248</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th>Methods</th>
<th>LCC</th>
<th>SROCC</th>
<th>KROCC</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>JP2K</td>
<td>0.918</td>
<td>0.911</td>
<td>0.692</td>
<td>0.112</td>
</tr>
<tr>
<td>JPEG</td>
<td>0.930</td>
<td>0.928</td>
<td>0.743</td>
<td>0.0981</td>
</tr>
<tr>
<td>Gblur</td>
<td>0.932</td>
<td>0.925</td>
<td>0.789</td>
<td>0.102</td>
</tr>
<tr>
<td>WN</td>
<td>0.963</td>
<td>0.956</td>
<td>0.820</td>
<td>0.0531</td>
</tr>
<tr>
<td>All</td>
<td>0.938</td>
<td>0.934</td>
<td>0.773</td>
<td>0.0945</td>
</tr>
</tbody>
</table>

B. CONTRASTIVE EXPERIMENT

In order to further test the performance of the proposed algorithm, some criteria are introduced. The widely used performance indicators are: Root mean squared error (RMSE), Pearson linear correlation coefficient (PLCC), Spearman’s rank-ordered correlation coefficient (SROCC) and Kendall rank order correlation coefficient (KROCC) [23]. The smaller value of RMSE means that the algorithm is more accurate for the image quality prediction. The larger PLCC value means that the algorithm is stronger for the correlation of image quality prediction, the value range of PLCC is 0 to 1. SROCC mainly measures the monotonicity of the algorithm. With the same range of 0 to 1, the larger SROCC value means the better the monotonicity.

The data in Table I-II are the performance indicators for the quality prediction of each distortion type image in the LIVE and CSIQ image databases. From the test results, we can see that the algorithm has good performance in LIVE and CSIQ databases. It performs best in the LIVE image database.
Additionally, because there are many distortion types in TID2013 image database, we use histogram to show the performance of our method (see Fig. 6). It can be see that our method still has good effect on most type of distorted images in TID2013 database.

![Figure 6](image)

**FIGURE 6.** The performance of our method performs on the performance of 24 distorted types of images in TID2013 database. The type of distortion: additive Gauss noise(WN), Differential additive noise in color components(DNC), space related noise(SRN), mask noise(MN), high frequency noise(HRN), pulse noise(PN), quantization noise(QN), Gblur, image denoising(ID), JPEG, JK2K, JPEG transmission error(JPEGTR), JP2K transmission error(JP2KTR), no-eccentricity type noise(NN), local block distortion with different strength(LB), mean shift (MS), contrast change(CC), color saturation change(CSC), multiplicative Gauss noise(MGN), comfortable noise(CN), compression of noisy images(CNI), Color quantization and fluctuation of images(CQF), Image chromatic aberration(ICA), sparse sampling and reconstruction(SSR).

The proposed method is compared with a number of NR-IQA and FR-IQA methods in LIVE, CSIQ and TID2013 databases. FR methods include PSNR, SSIM [24], IFC [25], VIF [26], NR methods include NIQE [27], BIQI [3], DIVINE [28], BLINDS2 [29], NRSE [5], NFERM [30], NR-GLBP [7]. In order to show the superiority of the BP-NN model, we also used SVR [31] set up an assessment regression model for comparison. The result of the experiment is shown in Table III. From the table, it can be seen that the performance of the proposed algorithm for quality prediction is the best in the NR methods in the LIVE and CSIQ databases. Although there are more distortion types in TID2013, our method still performs well. The performance of BP-NN in the evaluation of image quality is better than that of SVR. Compared with the FR Methods, our method is still competitive. The FR algorithm needs the original image as a reference, so the application prospect of the proposed algorithm is more extensive.

<table>
<thead>
<tr>
<th>Methods</th>
<th>LIVE</th>
<th>CSIQ</th>
<th>TID2013</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PLCC</td>
<td>SROCC</td>
<td>RMSE</td>
</tr>
<tr>
<td>PSNR</td>
<td>FR</td>
<td>0.884</td>
<td>0.873</td>
</tr>
<tr>
<td>SSIM[24]</td>
<td>FR</td>
<td>0.940</td>
<td>0.934</td>
</tr>
<tr>
<td>IFC[25]</td>
<td>FR</td>
<td>0.943</td>
<td>0.946</td>
</tr>
<tr>
<td>VIF[26]</td>
<td>FR</td>
<td>0.957</td>
<td>0.966</td>
</tr>
<tr>
<td>NIQE[27]</td>
<td>NR</td>
<td>0.916</td>
<td>0.913</td>
</tr>
<tr>
<td>BIQI[3]</td>
<td>NR</td>
<td>0.837</td>
<td>0.761</td>
</tr>
<tr>
<td>DIVINE[28]</td>
<td>NR</td>
<td>0.920</td>
<td>0.928</td>
</tr>
<tr>
<td>BLINDS2[29]</td>
<td>NR</td>
<td>0.931</td>
<td>0.933</td>
</tr>
<tr>
<td>NRSE[5]</td>
<td>NR</td>
<td>0.936</td>
<td>0.935</td>
</tr>
<tr>
<td>NFERM[30]</td>
<td>NR</td>
<td>0.942</td>
<td>0.938</td>
</tr>
<tr>
<td>NR-GLBP[7]</td>
<td>NR</td>
<td>0.943</td>
<td>0.938</td>
</tr>
<tr>
<td>SDC(SVR)</td>
<td>NR</td>
<td>0.956</td>
<td>0.953</td>
</tr>
<tr>
<td>SDC (BP-NN)</td>
<td>NR</td>
<td>0.962</td>
<td>0.956</td>
</tr>
</tbody>
</table>

C. GENERALIZATION EXPERIMENT

In order to verify the generalization of our method, we did some cross-database test to demonstrate the generalization ability of the algorithm. Specific tests are training in the LIVE database and testing in CSIQ and TID2013 databases. The cross experiment mainly tests the WN, JPEG, JP2K and Gblur distorted images in the image databases, because there
are only four kinds of distorted images of the same type in the three image databases. Moreover, because the DMOS value range of the three image databases are different, we have zoomed the scope of the DMOS value in the LIVE database to the same size as the test databases.

In addition, we also did some test in the CLIVE image database. The test data is shown in Table VI. CLIVE is a very challenging image database. Because this database is mixed with most of the distortion in the real life, quality assessment on CLIVE is much more difficult than that on other databases. Performances of all methods evaluated are much worse than in the single distorted image databases. The data in Table VI show that our method has also lost its due ability in this database. It can be concluded that our method has better generalization ability in the single distortion database, but it still can’t well evaluate for multi-mixed distorted images.

**D. COMPLEXITY EXPERIMENT**

Complexity is also a major factor affecting the performance of the algorithm. In many practical applications, it requires a relatively short period of time to complete the assessment of image quality. We used different methods to evaluate the quality of "statue" image in the LIVE image database and compared the time complexity. The computer used in the experiment is configured as: Intel Core i5-7400@3.0GHz, 4GB RAM, Win10 system, Matlab2017b software platform.

The time used of different algorithms is shown in Table VII. But running time only can reflect the complexity of algorithm unilaterally. Ref.[11] used \( O(\cdot) \) to more accurately describe time complexity of algorithm. In order to compare the complexity more accurately, we also increased the comparison of the time complexity ( \( O(\cdot) \) ) of each algorithm in Table VII. Compared with running time, time complexity is not affected by implementation methods and machine performance. In the Table, "Complexity" represents the time complexity of each algorithm for extracting features. \( K \) represents the number of pixels of the image and \( d \) is the window size (in pixels) used for feature computation. The complexity of our method for feature extraction is \( O(K+Kd) \) and our method uses about 0.174s to assessment the quality of "statue" image. It can be seen that our algorithm has relatively low complexity. Because our method extracts features in the spatial domain, without complex frequency domain transformation. In the experiment, the complexity of PSNR is the lowest, our method achieves the lowest complexity of the NR methods. This lays the foundation for the application of the algorithm in the next step.

**V. APPLICATION**

Blurriness is one of the most common factors that influence image quality. In our daily life, images are frequently distorted by blurriness which induced by defocusing of cameras, so adjusting the focus on the optimization of image quality is of great research significance. We set up an auto focusing experimental
platform (see Fig.7). In order to further illustrate the ability and practicality of the proposed method, we use MATLAB and C++ mixed programming to assess the quality of the pictures taken from camera and realized the auto focus of the camera.

We use MATLAB and C++ mixed programming to assess the quality of the pictures taken from camera and realized the auto focus of the camera.

The auto focus method used in the platform is shown in Fig. 8. According to the assessment model, the quality of the image collected by the camera is predicted. In order to make the motor rotate for the first time, we set up Quality(0)=100. If the image quality meets the requirements, the focus is finished. If not, the Quality(i) obtained by this calculation is compared with the last calculation quality. We let the motor continue to rotate when Quality(i) < Quality(i-1) and change the direction of rotation when Quality(i) > Quality(i-1). Once the image quality reaches the threshold T or the number of focal points reaches 12 times, the collected images are output. The maximum focus number is set to 12 times to prevent the quality scores from falling into the local optimal. we set T=20, because it can be determined that the quality of the image is "excellent" when the image quality score is less than 20.

![Focus test platform](image1)

**FIGURE 7.** Focus test platform, using PC to assessment the image from the camera and drive the motor to rotate according to the assessment results. (a) is the experiment platform (b) is the schematic diagram of the platform.

![Autofocus method diagram](image2)

**FIGURE 8.** Autofocus method, adjust the rotation direction of the motor by the image quality and output the image when image quality reaches the threshold.

![Images collected in the process of focusing](image3)

**FIGURE 9.** Images collected in the process of focusing and their quality scores: (a) is the Initial image and its predicted DMOS is 54.291 (b) is the image after the first focus and its predicted DMOS is 47.767 (c) is the image after the second focus and its predicted DMOS is 44.768 (d) is the image after the third focus and its predicted DMOS is 32.741 (f) is the image after the fourth focus and its predicted DMOS is 28.019 (g) is the output image and its predicted DMOS is 16.155.
Fig. 9 shows the images collected during the focus process and the predicted DMOS values. It can be seen that with the number of focus increased, the DMOS value of the image is getting smaller and smaller and the image is gradually clearer, so it can draw a conclusion that the proposed IQA methods is also very practical in practical applications.

**VI. CONCLUSION**

We proposed an efficient NR-IQA method based on spatial domain coding. The gray value of the image is encoded by binary encoding. The GLBP features are extracted from the coded binary bit-planes. The extracted features can more accurate to describe the degree of distortion of the image. Then the neural network is used to construct a mapping model between image features and image quality, afterwards the image quality is predicted by the model. The experimental in some single distorted image databases results show that the proposed method is superior to most mainstream IQA algorithms, which is in good agreement with human subjective perception and has lower time complexity. Finally, we set up an auto focus experimental platform and applied this algorithm to auto focus. Experiment shows that our method can also be well evaluated for the image captured by the camera.

Our method is effective in evaluating single distorted images, but we did not consider the impact on the image caused by multiple distortion. However, there are many multi-mixed distortion images in real life. Experiments in the CLIVE database show that the contrastive assessment method and our method are not ideal for multi-mixed distortion image evaluation. We will strengthen the research in this area to further improve the image assessment system.

**References**


CHEN Yong Professor at the School of Automation, Chongqing University of Posts and Telecommunication. He received his Ph.D. degree in mechanical engineering from Chongqing University in 2003. His research interest covers image processing, pattern recognition, and intelligent optimizing controls. E-mail: chenyong@cqupt.edu.cn.

FANG Hao Master student at the School of Automation, Chongqing University of Posts and Telecommunication. His research interest covers image processing, machine learning, and computer vision. E-mail: 1010507348@qq.com

LIU Huan-Lin Professor at the School of Telecommunication and Information Engineering, Chongqing University of Posts and Telecommunication. She received her Ph.D. degree in optical engineering from Chongqing University in 2008. Her research interest covers information acquiring and processing. E-mail: liuhl@cqupt.edu.cn.