Image Quality Assessment Based on Quaternion Singular Value Decomposition

QINGBING SANG1,2, YUNSHUO YANG1,2, LIXIONG LIU3, (Member, IEEE), XIAONING SONG1,2, (Member, IEEE), AND XIAOJUN WU1,2, (Member, IEEE)

1School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China
2Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computational Intelligence, Jiangnan University, Wuxi 214122, China
3School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China

Corresponding author: Qingbing Sang (sangqb@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61876072 and Grant 61672095, and in part by the Jiangsu Natural Science Foundation of Jiangsu Province of China under Grant BK20171142.

ABSTRACT We propose an image quality assessment metric based on quaternion singular value decomposition that represents a color image as a quaternion matrix, separates image noise information using singular value decomposition and extracts features from both the whole image and its noise information. In the proposed method, the color image and its local variance are represented by using quaternion and then performing singular value decomposition. Later, 75% of singular values are taken as image noise information. We extract a luminance comparison, contrast comparison, structure comparison, phase congruency and gradient magnitude from whole color images and extract the peak signal-to-noise ratio from image noise information as features. Finally, these features are used as the input to a kernel extreme learning machine to predict the quality of the tested images. Extensive experiments performed on four benchmark image quality assessment databases demonstrate that the proposed metric achieves high consistency with the subjective evaluations and outperforms state-of-the-art image quality assessment metrics.

INDEX TERMS Quaternion singular value decomposition, kernel extreme learning machine, image quality assessment.

I. INTRODUCTION

In this time of rapid information development, the visual quality of images is becoming increasingly important. Images may be distorted during acquisition, transmission, compression, restoration, and processing [1], [2]. Image quality assessment (IQA) [3] can usually be divided into subjective image quality assessment and objective image quality assessment. Subjective IQA metrics are expensive and time consuming. Objective IQA metrics research aims to design computational models that can automatically predict image quality. Therefore, research on objective IQA metrics is significant. According to the availability of original reference images, objective IQA metrics can be classified into full reference (FR), no-reference (NR) and reduced-reference (RR) metrics [4].

Here, we aim to develop an efficient FR IQA model. The most widely used FR IQA algorithms are the peak signal-to-noise ratio (PSNR) and mean squared error (MSE) [5], which calculate the distortion between the reference image pixels and corresponding distorted image pixels. However, they are criticized for their poor prediction of perceived image quality.

Many effective methods have been proposed to address the issue of FR IQA. Sheikh et al. introduced the information fidelity criterion (IFC) [6] and visual information fidelity (VIF) [2], which quantify the information shared between the distorted and the reference images. Wang et al. introduced the structural similarity index (SSIM) [4], which measures distorted image quality by luminance comparison, contrast comparison and structure comparison. The multi-scale extension of SSIM (MS-SSIM) [7] produces better results than its single-scale counterpart. Zhang et al. proposed a feature similarity index (FSIM) [8] that applies phase congruency (PC) and gradient magnitude (GM) in color image quality assessment. Recently, Reisenhofer et al. proposed a Haar wavelet-based perceptual similarity index (HaarPSI) [9] method that obtains local similarities using high-frequency
Haar wavelet coefficients. Although these methods delivered good performance, there still remains room for performance improvement in the field of FR IQA models.

In recent years, image quality assessment methods based on deep neural network are proposed. Zhang et al. [10] proposed a deep bilinear model for blind image quality assessment that works for both synthetically and authentically distorted images. He et al. [11] proposed a visual residual perception optimized network (VRPN) that separates the training of blind IQA into a distortion degree identification network and an image quality prediction network. Yan et al. [12] proposed a novel deep neural networks based multi-task learning approach. They introduced natural scene statistics features prediction as a task to our multi-task learning model and let this task to aid the quality prediction task. Bosse et al. [13] presented novel network architectures, incorporating an optional joint optimization of weighted average patch aggregation implementing a method for pooling local patch qualities to global image quality. Kim and Lee [14] proposed the full-reference image quality assessment metrics as intermediate training targets of the convolutional neural networks. The biggest defect of the deep learning-based IQA method is that a large number of training samples are needed, and we can just use a better-performing FR IQA method to make up for this defect. The specific approach is as follows: firstly, a large number of distorted images with different distortion levels and types are synthesized by the distortion algorithm on some reference images, and then the distorted images are labeled by the FR IQA algorithm, so that the synthesized samples can be used to pretrain the deep learning model. This is also a new application field of FR IQA algorithm.

Singular value decomposition (SVD) has been successfully used in the area of IQA. Azadeh and Ahmad [15] proposed a structural SVD-based image quality assessment (SSVD) that evaluates the image quality by reflecting on the left and right original singular vectors. Sang et al. [16] introduced an IQA metric based on a reciprocal singular value curve that obtains image quality indices by the area and curvature of image reciprocal singular value curves. We also deploy this tool to further explore a better method for gauging the perception of distortion.

Furthermore, we use an extreme learning machine in our proposed IQA method. Traditional neural networks must manually set a large number of network training parameters and easily fall into the local optimal solution. However, the extreme learning machine is a single-hidden layer feedforward neural network [17] that obtains a unique optimal solution through a set number of hidden layer nodes. Thus, it has great advantages in learning speed and generalization performance. Specifically, the kernel extreme learning machine is introduced into the kernel function, which can obtain the least square optimal solution. Here, we deploy it to implement quality prediction using various features.

The contributions of our proposed model are summarized as follows: 1) We use quaternion to represent the R, G, B three-channel and local variance of the color image, and conduct quaternion singular value decomposition to separate the noise part. 2) We extract different features from the whole image and noise part, and integrate them into our algorithm to make learning prediction by the kernel extreme learning machine. 3) We show that the proposed metric achieves high consistency with the subjective evaluations and outperforms state-of-the-art image quality assessment metrics.

The remainder of this paper is organized as follows. Section II introduces the theory of local variance and singular value decomposition of quaternion matrices. Section III presents the proposed algorithm. Section IV analyzes the experimental results. Section V summarizes the whole paper.

II. RELATED WORK

A. OVERVIEW OF LOCAL VARIANCE

The local variance in the image can describe the detailed information of an image. The study [18] found that local variance is more sensitive to blurred images than other distorted images. Therefore, the local variance is used to enhance the feature representation in our proposed algorithm.

Let \( \text{Var}(I_{x,y}) \) be the local variance in image \( I \), where \( I_{x,y} \) is an image block. \( \text{Var}(I_{x,y}) \) is computed within a local square window centered at the pixel \((x, y)\), which moves pixel-by-pixel over the entire image. We need to convert color images to YUV space (luminance channel Y, two chrominance channels U and V) and use the luminance layer Y to compute the local variance.

The local variance in an image is defined as

\[
\text{Var}(I_{x,y}) = \frac{\sum_{p=1}^{L} w_p (\eta_p - \bar{I}_{x,y})^2}{\sum_{p=1}^{L} w_p}
\]

where \( \bar{I}_{x,y} \) is the local mean

\[
\bar{I}_{x,y} = \frac{\sum_{p=1}^{L} w_p \eta_p}{\sum_{p=1}^{L} w_p}
\]

Considering that dividing the image into blocks damages the image structure correlation to some degree, we use a smoothing function to reduce such an influence. An \( 11 \times 11 \) circular-symmetric Gaussian weighting function \[4\]

\[
w = \{ w_i | i = 1, 2, \ldots, N \}
\]

is employed, with a standard deviation of 1.5 samples, normalized to the unit sum as follows:

\[
\sum_{i=1}^{N} w_i = 1
\]

where \( N \) is the total number of elements within a window.
B. OVERVIEW OF QUATERNION SINGULAR VALUE DECOMPOSITION

The concept of quaternions was first proposed by Hamilton in 1843 [19]. The quaternion \( q \) is defined as

\[
q = a + bi + cj + dk
\]

(4)

where \( a, b, c \) and \( d \) are real numbers, and \( i, j \) and \( k \) are imaginary number units and obey the rules as follows:

\[
i^2 = j^2 = k^2 = -1
\]

(5)

\[
j = -ji = k, \quad ki = -jk = j, \quad jk = -kj = i
\]

(6)

A color image can be decomposed into red (R), green (G), and blue (B) channels. Each pixel of a color image can be represented as a single pure quaternion-valued pixel [20].

\[
\begin{align*}
q(x, y) &= r(x, y)i + g(x, y)j + b(x, y)k
\end{align*}
\]

(7)

where \( r(x, y) \), \( g(x, y) \) and \( b(x, y) \) are the red, green and blue components, respectively, corresponding to the pixel at position \((x, y)\) in the color image.

According to (7), the real part of quaternion is 0. It can replace the local variance \( a = \text{Var}(x, y) \) to make the quaternion matrix capture more local details and structural information [21]. Thus, a quaternion of a color image can be defined as

\[
q(x, y) = \text{Var}(I_{x,y}) + r(x, y)i + g(x, y)j + b(x, y)k
\]

(8)

We only obtain luminance information from grayscale information by SVD. Since SVD can be applied to hypercomplex matrices [22], both luminance and chroma information can be obtained from color images by quaternion singular value decomposition (QSVD) [23], [24], i.e., more information of images can be obtained by QSVD. Let \( X_q \) the \( m \times n \) quaternion matrix with rank \( r \). Then, \( X_q \) can be decomposed uniquely, as follows:

\[
X_q = U_q \times S_q \times V_q^H
\]

(9)

where \( U_q \) is an \( m \times m \) quaternion unitary matrix, \( S_q \) is an \( m \times n \) diagonal matrix (i.e., the diagonal elements are singular values \( \sigma \)), and \( V_q \) is an \( n \times n \) quaternion unitary matrix. Let \( H \) conjugate transpose, and \( U_q, S_q \) and \( V_q \) are defined as

\[
U_q = [u_{1q} u_{2q} \cdots u_{m(q)}]
\]

(10)

\[
S_q = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)
\]

(11)

\[
V_q = [v_{1q} v_{2q} \cdots v_{n(q)}]
\]

(12)

where \( u_{i}(q)(i = 1, 2, \ldots, m) \) is the column vector of unitary matrix \( U_q \), \( v_{j}(q)(j = 1, 2, \ldots, n) \) is the column vector of unitary matrix \( V_q \), and \( \sigma_k(k = 1, 2, \ldots, r) \) is the diagonal element of diagonal matrix \( S_q \), \( r = \min(m, n) \). In addition, the singular values \( \sigma_k \) are in descending order, i.e., \( \sigma_1 > \sigma_2 > \cdots > \sigma_r \).

The quaternion singular value decomposition of color image \( X_q \) can also be denoted as the sum of the products of \( \sigma_i \) and \( u_{i}(q)v_{i}(q)^H \) as follows:

\[
X_q = \sum_{i=1}^{r} \sigma_i u_{i}(q)v_{i}(q)^H
\]

(13)

where \( u_{i}(q)v_{i}(q)^H \) is a color feature image of \( X_q \). The color image \( X_q \) can be treated as the linear combination of \( r \) color feature images. The \( i \)-th singular value \( \sigma_i \) denotes the luminance and chroma information of the \( i \)-th color feature image. Specifically, the singular value denotes the weights of the color feature image in the color image \( X_q \). The advantage of quaternion singular value decomposition is that it represents the color image as a whole and overcomes the shortcoming of traditional singular value decomposition.

III. OVERVIEW OF THE PROPOSED MODEL

The framework of the proposed model is shown in Figure 1. First, the reference image and distorted image are represented by using the quaternion. Specifically, we take the R, G, B components of the color image as the imaginary parts of the quaternion and the local variance as the real part of the quaternion. Second, we obtain image noise information by SVD. Third, we extract the luminance comparison, contrast comparison, structure comparison, PC and GM from the whole color image and extract the PSNR from image noise information. Finally, the six generated features are fed into a kernel extreme learning machine to predict the quality of the tested image.

A. IMAGE NOISE INFORMATION EXTRACTED FROM QSVD

Generally, the sum of the first 10% or even 1% of the singular values accounts for more than 99% of the total singular values, i.e., most of the energy of the image is concentrated on the color feature image corresponding to the larger singular values. Liu and Lin [25] proved that the latter 75% of singular values represent the data in which noise is the dominant factor. Thus, the noise information of the color image can be defined as

\[
N_o = \sum_{i=p+1}^{r} \sigma_i u_{i}(q)v_{i}(q)^H
\]

(14)

\[
N_d = \sum_{i=p+1}^{r} \sigma_i u_{i}(q)v_{i}(q)^H
\]

(15)

where \( r \) is the rank of a color image, \( p \in (0, r) \) is a threshold (here we set \( p = 0.25r \)). \( N_o \) and \( N_d \) are the noise information of the reference image and the distorted image, respectively.

B. FEATURES EXTRACTED FROM THE WHOLE IMAGE

Natural image signals are highly structured, i.e., their pixels have strong dependencies. It is generally known that each position of the image is considered to have different importance. Therefore, we extract features by referring to SSIM and FSIM.

SSIM combines luminance, contrast and structure information to assess image quality. We take the mean of the image as luminance information, the standard deviation of the image as contrast information, and the covariance of the image as structure information. The luminance comparison \( l(x, y) \), contrast comparison \( c(x, y) \) and structure comparison
where $\mu_x$ and $\mu_y$ are the mean of the distorted image block and the mean of the corresponding reference image block, respectively, $\sigma_x$ and $\sigma_y$ denote the standard deviation of the distorted image block and the standard deviation of the corresponding reference image block, and $\sigma_{xy}$ and $\sigma_x \sigma_y$ denote the covariance between the distorted image block and the corresponding reference image block. The constants $C_1$, $C_2$ and $C_3$ are included to avoid instability when the denominator is very close to zero. Specifically, we choose

$$C_1 = (K_1 L)^2$$

(19)

$$C_2 = (K_2 L)^2$$

(20)

$$C_3 = C_2 / 2$$

(21)

where $L$ is the dynamic range of the pixel values, $K_1 \ll 1$ and $K_2 \ll 1$ are small constants. We set $K_1 = 0.01$ and $K_2 = 0.03$.

SSIM has a deficiency that when pooling a single quality score from the local quality map, all pixel positions are considered to have the same importance. However, the edge areas of the image are definitely considered to be more important than the smooth areas [26]. Based on the above analysis, we also refer to FSIM in our proposed model. FSIM combines PC with GM to assess image quality. PC captures highly informative features and can be considered a measure of the significance of a local structure. GM also describes the contrast information of the image.

Rather than defining features directly at points with sharp changes in intensity, the PC model postulates that features are perceived at points where the Fourier components are maximal in phase. The method developed by Kovesi was employed in FSIM [27]. We can apply the 1D analysis over several orientations and then obtain the result using Gaussian as the spreading function. The PC of an image is defined as $G_y(x)$.

$$PC(x) = \frac{\sum_o |E_o(x) - T_o|}{\varepsilon + \sum_o \sum_n A_{no}(x)}$$

(22)

where $E_o(x)$ is the local energy, $T_o$ is a noise compensation factor, $A_{no}(x)$ is the local amplitude on scale $n$, $o$ is the orientation, and $\varepsilon$ is a small constant.

Gradient magnitude can be expressed by convolution masks. Three commonly used gradient operators are the Sobel operator [28], the Prewitt operator [28] and the Scharr operator [29]. Here, we adopt the Scharr operator. The GM of the image is defined as

$$G = \sqrt{G_x^2 + G_y^2}$$

(23)

where $G_x(x)$ and $G_y(x)$ are the partial derivatives of the image along the horizontal and vertical directions, respectively.

C. FEATURES EXTRACTED FROM IMAGE NOISE INFORMATION

Image noise is closely associated with image perceptual quality. Thus, we extract features from image noise information.
Specifically, we calculate the similarity between the reference and distorted image noise by using the PSNR. PSNR calculates the distortion between the reference image pixels and corresponding distorted image pixels. Although it is not consistent with the human visual system (HVS), PSNR is good at assessing noisy images. The PSNR of the $m \times n$ image is defined as

\[
PSNR = 10 \times \log_{10} \frac{L^2}{MSE} \tag{24}
\]

\[
MSE = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij} - y_{ij})^2 \tag{25}
\]

where $L$ is the dynamic range of the pixel values, $i$ and $j$ denote the row and column of the image, $x_{ij}$ and $y_{ij}$ denote the pixel values of the distorted image and reference image, respectively.

D. KERNEL EXTREME LEARNING MACHINE

After extracting features, we use an extreme learning machine (ELM) to map generated features to image quality. The ELM is a single-hidden layer feedforward neural network (SLFN) that can randomly initialize the weights and biases of the input and obtain a unique optimal solution only by setting the number of hidden layer nodes. The extreme learning machine is mathematically modeled as

\[
\sum_{i=1}^{N} \beta_i g(w_i x_j + b_i) = o_j, \quad j = 1, 2, \ldots, N \tag{26}
\]

where $w_i = [w_{i1}, w_{i2}, \cdots, w_{iN}]^T$ is the weight vector connecting the $i$th hidden node and the input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \cdots, \beta_{im}]^T$ is the weight vector connecting the $i$th hidden node and the output nodes, $b_i$ is the threshold of the $i$th hidden node. The structure of the extreme learning machine network is shown in Figure 2.

The ELM can be easily implemented and runs extremely fast. However, there is currently no effective method for accurately estimating the number of nodes of the hidden [30]. ELM is bedeviled by the curse of dimensionality, which runs in high-dimensional feature space [31].

To improve robustness and the capability of nonlinear approximation, Huang et al. [32] added the kernel function to an extreme learning machine. We call it the kernel extreme learning machine (KELM). The kernel matrix needs to satisfy the Mercer conditions and is defined as

\[
\Omega = HH^T : \Omega_{i,j} = h(x_i)h(x_j) = K(x_i, x_j) \tag{27}
\]

where $K(x_i, x_j)$ is the kernel function (we choose the Gaussian RBF kernel function), and $\Omega$ is the kernel matrix. The kernel extreme learning machine is mathematically modeled as

\[
f(x) = h(x)\beta = h(x)H^T(I + HH^T)^{-1}T
\]

\[
= \begin{bmatrix} K(x_1, x_1) \\ \vdots \\ K(x_N, x_N) \end{bmatrix}^T \begin{bmatrix} I \\ \vdots \\ \Omega \end{bmatrix}^{-1}T \tag{28}
\]

where $I$ is the unit diagonal matrix, and $C$ is the penalty coefficient. The parameter $\frac{1}{C}$ is added to the leading diagonal of diagonal matrix $HH^T$. According to (28), the weight of the output nodes is $\beta = (\frac{1}{C} + \Omega)^{-1}T$.

IV. EXPERIMENTAL RESULTS

A. IMAGE DATABASES

Four publicly available image databases are used to test the performance of the proposed method, including CSIQ [33], LIVE [34], TID2008 [35], and TID2013 [36]. The details of four image databases are summarized in Table 1.

| Database | Source images | Distorted images | Distortion types |
|----------|---------------|------------------|-----------------|
| CSIQ     | 30            | 866              | 6               |
| LIVE     | 29            | 982              | 5               |
| TID2008  | 25            | 1700             | 17              |
| TID2013  | 25            | 3000             | 24              |

TABLE 1. Information on the IQA databases.
TABLE 2. Performance (SROCC) of our proposed model with different parameters on the CSIQ database.

| Parameters | R=10   | R=500   | R=1000 | R=3000 |
|------------|--------|---------|--------|--------|
| K=1        | 0.9694 | 0.9714  | 0.9710 | 0.9703 |
| K=10       | 0.9570 | 0.968   | 0.9684 | 0.9688 |
| K=20       | 0.9522 | 0.9652  | 0.9658 | 0.9664 |
| K=30       | 0.9497 | 0.9627  | 0.9636 | 0.9645 |

TABLE 3. Performance (SROCC) of our proposed model with different parameters on the LIVE database.

| Parameters | R=10   | R=500   | R=786  | R=1000 | R=1500 |
|------------|--------|---------|--------|--------|--------|
| K=1        | 0.9707 | 0.9756  | 0.976  | 0.9762 | 0.9765 |
| K=10       | 0.9707 | 0.9756  | 0.976  | 0.9762 | 0.9765 |
| K=20       | 0.9676 | 0.9735  | 0.9739 | 0.9742 | 0.9747 |
| K=30       | 0.9665 | 0.9724  | 0.9728 | 0.973  | 0.9734 |

TABLE 4. Performance (SROCC) of our proposed model with different parameters on the TID2008 database.

| Parameters | R=10   | R=500   | R=1360 | R=1500 | R=2000 |
|------------|--------|---------|--------|--------|--------|
| K=1        | 0.9097 | 0.9204  | 0.9199 | 0.9198 | 0.9195 |
| K=10       | 0.8763 | 0.8995  | 0.9033 | 0.9036 | 0.9046 |
| K=20       | 0.8681 | 0.8914  | 0.896  | 0.8966 | 0.8977 |
| K=30       | 0.8637 | 0.8865  | 0.8916 | 0.8921 | 0.8933 |

TABLE 5. Performance (SROCC) of our proposed model with different parameters on the TID2013 database.

| Parameters | R=100  | R=2000  | R=2400 | R=3000 | R=3500 |
|------------|--------|---------|--------|--------|--------|
| K=1        | 0.9031 | 0.9103  | 0.9104 | 0.9106 | 0.9106 |
| K=10       | 0.8568 | 0.8805  | 0.8816 | 0.8830 | 0.8838 |
| K=20       | 0.8365 | 0.8679  | 0.8691 | 0.8706 | 0.8716 |
| K=30       | 0.8264 | 0.8596  | 0.8615 | 0.8630 | 0.8642 |

where $s_o$ is the objective outputs, $s_m$ is the mapped score, and $\beta_i, i = 1, 2, 3, 4, 5$ are the parameters to be fitted by minimizing the sum of squared errors between the mapped score $s_m$ and subjective scores. We use the 80% images in each database for training and the remaining 20% images for testing. The experiments are repeated one hundred times, and the average value of the evaluation results is taken as the final result.

C. PARAMETER SETTING

We set $Elm\_type = RBF\_kernel$ and $Kernel\_para = 1$. The parameter $Regularization\_coefficient$ is equal to the number of training sets. The numbers of training subsets in the CSIQ, LIVE, TID2008, and TID2013 databases are $866 \times 0.8 \approx 693, 982 \times 0.8 \approx 786, 1700 \times 0.8 \approx 1360,$ and $3000 \times 0.8 \approx 2400,$ respectively.

The SROCC of our proposed model with different parameters on each database is listed in Tables 2-5. The parameters $Kernel\_para$ and $Regularization\_coefficient$ are simplified to $K$ and $R$, respectively. It is obvious that the SROCC value is optimal when $K = 1$ and $R$ is equal to approximately the number of training subsets. For unity, the parameter $R$ is set as the number of training sets.

D. THE ADVANTAGE OF KELM

The kernel extreme learning machine has an advantage against other regression tools. We listed the SROCC, PLCC, KROCC and RMSE of our proposed model by support vector regression (SVR) [38], ELM, online sequential extreme learning machine (OSELM) [39], OSELM_VAR [39] and KELM on different image databases across a hundred trials in Tables 6-9. The quantity of initial training data used in the initial phase of OSELM and OSELM_VAR is 56. The size of the block of data learned by OSELM at each step is 1. The range from which the size of the data block is randomly generated at each iteration of the sequential learning phase in OSELM_VAR is $[10, 30]$. The results show that the KELM has the best performance improvement.

E. CONTRIBUTE OF LOCAL VARIANCE

To make the quaternion matrix capture more local details and structural information, the real part of quaternion is set with local variance of a color image in our proposed model. We also implemented a comparative experiment to study the contribution of local variance. The SROCC, PLCC, KROCC and RMSE values on different image databases are listed in Table 10. The top results are highlighted in boldface. It is obvious that the local variance contributes to the performance of our proposed model.

F. EACH FEATURE/FEATURE VECTOR CONTRIBUTION

We extracted luminance comparison, contrast comparison, structure comparison, PC and GM from a whole color image
and extracted PSNR from image noise information in our proposed model. The generated feature vector of the \( j \)th image is defined as

\[
x_j = [l_j, c_j, s_j, pc_j, g_j, n_j]^T
\]

(30)

where \( l_j, c_j, s_j, pc_j, g_j \) and \( n_j \) are luminance comparison, contrast comparison, structure comparison, PC, GM and PSNR, respectively. We randomly combined the extracted features. The combination of the luminance comparison, contrast comparison and structure comparison is called \( F_s \), and the combination of PC and GM is called \( F_f \). The PSNR is only called \( F_p \). The SROCC, PLCC, KROCC and RMSE for different feature combinations on different image databases across a hundred trials are listed in Tables 11-14. It is obvious that all six features delivered the best performance.

The GM and contrast comparison represent the contrast information of the image. We also list the SROCC, PLCC, KROCC and RMSE for different feature combinations on different image databases across a hundred trials are listed in Tables 11-14. In Table 16, the listed distortion types include white noise (WN), JPEG compression (JPEG), JPEG2000 compression (JP2K), additive pink Gaussian noise (APGN), Gaussian blur (GB), global contrast decrements (GCD), fast fading Rayleigh (FF), additive Gaussian noise (AGN), and additive white Gaussian noise, which is more intensive in color components than in the luminance component (ANC), spatially correlated noise (SCN), masked noise (MN), high-frequency noise (HFN), impulse noise (IN), quantization noise (QN), image denoising residual noise (DEN), JPEG transmission errors (JGTE), JPEG2000 transmission errors (J2TE), non eccentricity pattern noise (NEPN), local block wise distortions of different intensity (Block), intensity shift (IS), contrast change (CTC), change in color saturation (CCS), multiplicative Gaussian noise (MGN), comfort noise (CN), lossy compression of noisy image (LCNI), image

### G. PERFORMANCE ON EACH DISTORTION TYPE

The proposed method's ability to estimate image degradation caused by specific types of distortions is discussed in the experiment. To save space, only SROCC scores are listed. We compared our proposed model with ten IQA metrics.
TABLE 16. Performance (SROCC) of compared models on different image distortion types.

| DB   | type    | proposed | MS-SSIM | VIF | IWSSIM | SSIM  | NQM  | PSNR | IPC  | FSMR | HaarPSI | SVD  |
|------|---------|----------|---------|-----|--------|-------|------|------|------|------|---------|------|
| CSIQ | JPEG   | 0.9697   | 0.9375  | 0.933 | 0.8974 | 0.9384 | 0.9363 | 0.8431 | 0.9262 | 0.9694 | 0.9974  |
|      | JPEG2K | 0.9727   | 0.9634  | 0.9705| 0.9662 | 0.9564 | 0.9527 | 0.8881 | 0.8412 | 0.9702 | 0.9953  |
|      | APGN   | 0.9811   | 0.9683  | 0.9672 | 0.9683 | 0.9606 | 0.9631 | 0.9362 | 0.9252 | 0.9685 | 0.9792  |
|      | GB     | 0.9735   | 0.9711  | 0.9745 | 0.9782 | 0.9609 | 0.9584 | 0.9291 | 0.9527 | 0.9729 | 0.9743  |
|      | GCD    | 0.9559   | 0.9526  | 0.9545 | 0.9539 | 0.7922 | 0.9479 | 0.8621 | 0.4873 | 0.9420 | 0.9586  |
|      | ALL    | 0.9710   | 0.9130  | 0.9190 | 0.9210 | 0.8750 | 0.7400 | 0.8050 | 0.7670 | 0.9240 | 0.9550  |
| LIVE | JPEG   | 0.9818   | 0.9627  | 0.9696 | 0.9649 | 0.9614 | 0.9438 | 0.8543 | 0.9113 | 0.9717 | 0.9961  |
|      | JPEG2K| 0.9739   | 0.9581  | 0.9646 | 0.9808 | 0.9576 | 0.9644 | 0.9851 | 0.9848 | 0.9976 | 0.9877  |
|      | WN     | 0.9756   | 0.9398  | 0.9688 | 0.9405 | 0.9138 | 0.9597 | 0.8743 | 0.9331 | 0.9582 | 0.9617  |
|      | GB     | 0.9796   | 0.9561  | 0.9728 | 0.972 | 0.9517 | 0.8436 | 0.7882 | 0.9584 | 0.9708 | 0.9682  |
|      | FF     | 0.9654   | 0.9462  | 0.965 | 0.9442 | 0.9556 | 0.8112 | 0.8898 | 0.9629 | 0.9494 | 0.9596  |
|      | ALL    | 0.9790   | 0.9510  | 0.9630 | 0.9560 | 0.9470 | 0.9080 | 0.8750 | 0.9250 | 0.9630 | 0.9690  |

H. PERFORMANCE ON FOUR DATABASES

In this section, we compare the overall performance of the competing IQA metrics. These metrics include color quantization with dither (ICQD), chromatic aberrations (CHA) and sparse sampling and reconstruction (SSR).

From Table 16, it is obvious that the proposed method has the top results on different distortion types and is superior to the SSIM, FSIM and PSNR metrics.

**TABLE 17. Performance (SROCC, KROCC, PLCC, RMSE) on four IQA databases.**

| DB   | AGN     | 0.8941 | 0.8086 | 0.7879 | 0.7869 | 0.8107 | 0.7679 | 0.7347 | 0.8360 | 0.8566 |
|------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|      | ANC     | 0.8325 | 0.8054 | 0.8757 | 0.792 | 0.8029 | 0.749 | 0.8546 | 0.564 | 0.8527 |
|      | SCN     | 0.9008 | 0.8209 | 0.8698 | 0.7714 | 0.8141 | 0.772 | 0.7461 | 0.5843 | 0.8384 |
|      | MN      | 0.7523 | 0.8107 | 0.8683 | 0.8087 | 0.7795 | 0.612 | 0.6732 | 0.8019 | 0.7919 |
|      | HFN     | 0.9182 | 0.8694 | 0.9075 | 0.8662 | 0.8729 | 0.9015 | 0.8042 | 0.7318 | 0.9093 |
|      | IN      | 0.8072 | 0.6907 | 0.8327 | 0.6464 | 0.6732 | 0.7616 | 0.8596 | 0.5345 | 0.7452 |
|      | QQ      | 0.8686 | 0.8589 | 0.797 | 0.8177 | 0.8531 | 0.8209 | 0.7057 | 0.5857 | 0.8564 |
|      | GB      | 0.9545 | 0.9563 | 0.954 | 0.9636 | 0.9544 | 0.8846 | 0.774 | 0.8539 | 0.9472 |
|      | DBN     | 0.9449 | 0.9582 | 0.9161 | 0.9473 | 0.945 | 0.8522 | 0.9372 | 0.5963 | 0.9696 |
|      | JPEG   | 0.9268 | 0.9322 | 0.9168 | 0.9184 | 0.9252 | 0.9075 | 0.752 | 0.818 | 0.9279 |
|      | JPEG2K | 0.9791 | 0.970 | 0.9709 | 0.9738 | 0.9625 | 0.9532 | 0.6673 | 0.9437 | 0.9773 |
|      | JGTE    | 0.8911 | 0.8681 | 0.8585 | 0.8588 | 0.8687 | 0.7673 | 0.7328 | 0.7209 | 0.8708 |
|      | J2TE    | 0.8616 | 0.8606 | 0.8501 | 0.8203 | 0.8577 | 0.7262 | 0.7480 | 0.7901 | 0.8544 |
|      | NEPNN   | 0.8037 | 0.7377 | 0.7619 | 0.7724 | 0.7107 | 0.6800 | 0.5945 | 0.8418 | 0.7491 |
|      | Block   | 0.7651 | 0.7546 | 0.8324 | 0.7623 | 0.8462 | 0.2348 | 0.8513 | 0.677 | 0.8492 |
|      | IS      | 0.6831 | 0.7336 | 0.5059 | 0.7067 | 0.7231 | 0.5245 | 0.5576 | 0.425 | 0.6722 |
|      | CTC     | 0.8457 | 0.6381 | 0.8188 | 0.6301 | 0.5246 | 0.6191 | 0.4459 | 0.1713 | 0.6481 |
|      | ALL     | 0.9199 | 0.8530 | 0.7480 | 0.6550 | 0.7740 | 0.6240 | 0.5530 | 0.5670 | 0.8800 |

VOLUME 8, 2020
that the proposed metric achieves high consistency with the subjective evaluations and outperforms state-of-the-art image quality assessment metrics.

Figure 3 plotted the scatter distributions of subjective MOS versus the predicted scores of the proposed method on different databases. The curves shown in Figure 3 were obtained by a nonlinear fitting, according to (29). It is clear that the proposed method provides a satisfactory fitting effect and achieves very high consistency with human subjective evaluations.

I. CROSS-DATABASE VALIDATION

To test the feasibility and robustness of the proposed method, a cross-database validation is conducted. The results are listed in Table 18. Our metric outperforms all of the other three methods.

V. CONCLUSION

We proposed an image quality assessment metric based on the singular value decomposition of quaternion matrices. Since the correlation of red, green, and blue channels of color images, the characteristics of local variance, the acquisition of image noise information, and the feature combination are considered, the generated features can provide more comprehensive information from different angles. Furthermore, the selection of a kernel extreme learning machine further improves the performance of the proposed method. The experimental results show that our proposed method achieves very high consistency with human subjective evaluations.

TABLE 17. Performance of compared IQA metrics on different databases.

| Database | proposed | SSIM | FSIM | PSNR | SCQI | LLM | SPQI | SVD | SCMM |
|----------|----------|------|------|------|------|-----|------|------|------|
| CSIQ     | 0.9710   | 0.8750 | 0.9240 | 0.8050 | 0.9434 | 0.9050 | 0.9240 | 0.9344 | 0.8979 |
|          | 0.9725   | 0.8610 | 0.9120 | 0.8000 | 0.9268 | 0.9000 | 0.9344 | 0.9236 | 0.9520 |
|          | 0.8514   | 0.6900 | 0.7560 | 0.6080 | 0.7870 | 0.7238 | 0.7880 | 0.7234 | 0.8130 |
|          | 0.0619   | 0.1330 | 0.1070 | 0.1570 | 0.0986 | 0.1232 | 0.0934 | 0.1007 | 0.0810 |
| LIVE     | 0.9790   | 0.9470 | 0.9630 | 0.8750 | 0.9480 | 0.9608 | 0.9620 | 0.9600 | 0.9590 |
|          | 0.9840   | 0.9440 | 0.9590 | 0.8720 | 0.9373 | 0.9578 | 0.9597 | 0.9570 | 0.9560 |
|          | 0.8825   | 0.7960 | 0.8330 | 0.6860 | 0.8098 | 0.8230 | 0.8271 | 0.8203 | 0.8470 |
|          | 0.0994   | 8.9450 | 7.6780 | 13.3500 | 8.0590 | 7.6767 | 7.6288 | 7.9274 | 8.0660 |
| TID2008  | 0.9199   | 0.7740 | 0.8800 | 0.5530 | 0.9051 | 0.9077 | 0.9104 | 0.7375 | 0.9050 |
|          | 0.9194   | 0.7350 | 0.8730 | 0.5730 | 0.8899 | 0.8971 | 0.8927 | 0.7954 | 0.9010 |
|          | 0.7513   | 0.5760 | 0.6940 | 0.4020 | 0.7292 | 0.7368 | 0.7303 | 0.5628 | 0.7270 |
|          | 0.1392   | 0.8510 | 0.6520 | 1.0990 | 0.6120 | 0.5982 | 0.6046 | 0.8133 | 0.5830 |
| TID2013  | 0.9104   | 0.7410 | 0.8610 | 0.6390 | 0.9052 | 0.9037 | 0.9044 | 0.6757 | 0.8090 |
|          | 0.9202   | 0.7890 | 0.8580 | 0.6770 | 0.9071 | 0.9068 | 0.9091 | 0.7825 | 0.8660 |
|          | 0.7365   | 0.5580 | 0.6280 | 0.4690 | 0.7327 | 0.7209 | 0.7251 | 0.5146 | 0.6370 |
|          | 0.4904   | 0.7600 | 0.6340 | 0.9110 | 0.5219 | 0.5277 | 0.5165 | 0.7719 | 0.6200 |

TABLE 18. PLCC values of compared IQA models across three databases.

| Test_Database | Train_Database | CSIQ | LIVE | TID2008 |
|---------------|----------------|------|------|---------|
| proposed      | -              | 0.9377 | 0.8542 | -       |
| [47]          |                | 0.9106 | 0.8615 | -       |
| SVDR [48]     |                | 0.9056 | 0.7580 | -       |
| CD-MMF [32]   |                | 0.9109 | 0.8536 | -       |
| proposed      | -              | 0.9121 | -     | 0.7894 |
| [47]          |                | 0.9371 | -     | 0.8652 |
| SVDR          |                | 0.8581 | -     | 0.7428 |
| CD-MMF        |                | 0.9710 | -     | 0.8984 |
| proposed      | -              | 0.9318 | 0.9660 | -       |
| [47]          |                | 0.9274 | 0.9220 | -       |
| SVDR          |                | 0.8831 | 0.8862 | -       |
| CD-MMF        |                | 0.9127 | 0.9644 | -       |

REFERENCES

[1] Z. Wang and A. C. Bovik, Modern Image Quality Assessment. San Rafael, CA, USA: Morgan & Claypool, 2006. [Online]. Available: https://ieeexplore.ieee.org/document/6812930

[2] H. R. Sheikh and A. C. Bovik, “Image information and visual quality,” IEEE Trans. Image Process., vol. 15, no. 2, pp. 430–444, Feb. 2006.

[3] Y. Ding, D. Zhang, and S. Wang, “Full-reference image quality assessment using statistical local correlation,” Electron. Lett., vol. 50, no. 2, pp. 79–81, Jan. 2014.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.

[5] G. Zhu and X. Min. (2020). Perceptual Image Quality Assessment: A Survey. [Online]. Available: http://engine.scichina.com/publisher/scp/journal/SCIS/doi/10.1007/s11432-019-2757-1?slug=abstract
[6] H. R. Sheikh, A. C. Bovik, and G. de Veciana, “An information fidelity criterion for image quality assessment using natural scene statistics,” *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.

[7] Z. Wang, E. P. Simoncelli, and A. C. Bovik, “Multiscale structural similarity for image quality assessment,” in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, 2003, pp. 1398–1402.

[8] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “FSIM: A feature similarity index for image quality assessment,” *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[9] R. Reisenhofer, S. Bosse, G. Kutyniok, and T. Wiegand, “A Haar wavelet-based perceptual similarity index for image quality assessment,” *Signal Process., Image Commun.*, vol. 61, pp. 33–43, Feb. 2018.

[10] W. Zhang, K. Ma, J. Yan, D. Deng, and Z. Wang, “Blind image quality assessment using a deep bilinear convolutional neural network,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 1, pp. 36–47, Jan. 2020.

[11] L. He, Y. Zhong, W. Lu, and X. Gao, “A visual residual perception optimized network for blind image quality assessment,” *IEEE Access*, vol. 7, pp. 176087–176098, 2019.

[12] B. Yan, B. B. Bare, and W. Tan, “Naturalness-aware deep no-reference image quality assessment,” *IEEE Trans. Multimedia*, vol. 21, no. 10, pp. 4046–4057, Oct. 2019.

[13] S. Bosse, D. Maniry, K.-R. Müller, T. Wiegand, and W. Samek, “Deep neural networks for no-reference and full-reference image quality assessment,” *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 206–219, Jan. 2018.

[14] J. Kim and S. Lee, “Deep blind image quality assessment by employing FR-IAQ,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3180–3184.

[15] A. Mansouri and A. Mahmoudi-Aznaveh, “SSVD: Structural SVD-based image quality assessment,” *Signal Process., Image Commun.*, vol. 74, pp. 54–63, May 2019.

[16] Q. Sang, X. Wu, C. Li, and A. C. Bovik, “Blind image quality assessment using a reciprocal singular value curve,” *Signal Process., Image Commun.*, vol. 29, no. 10, pp. 1149–1157, Nov. 2014.

[17] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: Theory and applications,” *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.

[18] Y. Q. Wang, W. Y. Liu, and Y. Wang, “Image quality assessment based on local variance and structure similarity,” *J. Optoelectron. Laser*, vol. 19, no. 11, pp. 1546–1553, 2008.

[19] W. R. Hamilton, *Elements Quaternions*, 2nd ed. Harlow, U.K.: Longmans, 1899.

[20] S.-C. Pei, J.-H. Chang, and J.-J. Ding, “Quaternion matrix singular value decomposition and its applications for color image processing,” in *Proc. Int. Conf. Image Process.*, vol. 1, Sep. 2003, pp. 805–808.

[21] Y. Wang, W. Liu, and Y. Wang, “Color image quality assessment based on quaternion singular value decomposition,” in *Proc. Conge Image Signal Process.*, vol. 3, May 2008, pp. 433–439.

[22] F. Zhang, “Quaternions and matrices of quaternions,” *Linear Algebra Appl.*, vol. 251, no. 2, pp. 21–57, Jan. 1997.

[23] Q. G. Yang, M. Fan, and L. Wang, “Denoising color images by reduced quaternion matrix singular value decomposition,” *Multidimensional Syst. Signal Process.*, vol. 26, no. 1, pp. 307–320, Jan. 2015.

[24] N. Le Bihan and J. Mars, “Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing,” *Signal Process.*, vol. 84, no. 7, pp. 1177–1199, Jul. 2004.

[25] W. Liu and W. Lin, “Additive white Gaussian noise level estimation in SVD domain for images,” *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 872–883, Mar. 2013.

[26] B. Wang, Z. Wang, Y. Liao, and X. Lin, “HVS-based structural similarity for image quality assessment,” in *Proc. 9th Int. Conf. Signal Process.*, Oct. 2008, pp. 1194–1197.

[27] P. Kovesi, “Image features from phase congruency,” *J. Comput. Vis. Res.*, vol. 1, no. 3, pp. 1–26, 1999.

[28] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*. New York, NY, USA: McGraw-Hill, 1995.

[29] B. Jähne, H. Haubecker, and P. Geibler, *Handbook of Computer Vision and Applications*. Boston, MA, USA: Academic, 1999.

[30] N. Liu and H. Wang, “Ensemble based extreme learning machine,” *IEEE Signal Process. Lett.*, vol. 17, no. 8, pp. 754–757, Aug. 2010.

[31] X. Zhu, Z. Ni, M. Cheng, F. Jin, J. Li, and G. Weckman, “Selective ensemble based on extreme learning machine and improved discrete artifical fish swarm algorithm for hazard forecast,” *Appl. Intell.*, vol. 48, no. 7, pp. 1757–1775, Jul. 2018.

[32] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.

[33] E. C. Larson and D. M. Chandler, *Categorical Subjective Image Quality CSIQ Database*. Accessed: 2009. [Online]. Available: http://vision.okstate.edu/csiq

[34] H. R. Sheikh, Z. Wang, L. Cormack, and A. Bovik. *Live Image Quality Assessment Database Release 2*. Accessed: 2005. [Online]. Available: http://live.ece.utexas.edu/research/quality

[35] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, “Tid2008—A database for evaluation of full-reference visual quality assessment metrics,” *Adv. Modern Electron.* vol. 10, pp. 30–45, May 2009.

[36] N. Ponomarenko, L. Jin, O. Jeremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. J. Kuo, “Image database TID2013: Peculiarities, results and perspectives,” *Signal Process., Image Commun.*, vol. 30, pp. 57–77, Jan. 2015. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01109219

[37] A. Mittal, A. K. Moorthy, and A. C. Bovik, “No-reference image quality assessment in the spatial domain,” *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.

[38] C.-C. Chang, C.-W. Hsu, and C.-J. Lin, “The analysis of decomposition methods for support vector machines,” *IEEE Trans. Neural Netw.*, vol. 11, no. 4, pp. 1003–1008, Jul. 2000.

[39] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, “A fast and accurate online sequential learning algorithm for feedforward networks,” *IEEE Trans. Neural Netw.*, vol. 17, no. 6, pp. 1411–1423, Nov. 2006.

[40] Z. Wang and Q. Li, “Information content weighting for perceptual image quality assessment,” *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.

[41] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and R. Zhang, “Extreme learning machine and accurate online sequential learning algorithm for feedforward networks,” *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 1561–1572, Dec. 2008.

[42] G. Yang, D. Li, F. Lu, Y. Liao, and W. Wang, “RVSIM: A feature similarity method for full-reference image quality assessment,” *EURASIP J. Image Video Process.*, vol. 6, no. 1, p. 6, Dec. 2018.

[43] F. Zhang, B. Zhang, R. Zhang, and X. Zhang, “SPCM: Image quality assessment based on symmetry phase congruency,” *Appl. Soft Comput.*, vol. 87, Feb. 2020, Art. no. 105987. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1568494619307689

[44] S. Wang, D. Cui, B. Wang, B. Zhao, and J. Yang, “A perceptual image quality assessment metric using singular value decomposition,” *Circuits, Syst., Signal Process.*, vol. 34, no. 1, pp. 209–229, Jan. 2015.

[45] M. Narwaria and W. Lin, “SVD-based quality metric for image and video using machine learning,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 347–364, Apr. 2012.
YUNSHUO YANG received the bachelor’s degree in the Internet of Things from the Hebei University of Engineering, in 2017. She is currently pursuing the master’s degree with the School of Artificial Intelligence and Computer, Jiangnan University, Wuxi, China. Her research interests include image analysis and image quality assessment.

LI XIONG LIU (Member, IEEE) was a Visiting Researcher with the Laboratory for Image and Video Engineering, The University of Texas at Austin, USA, from 2009 to 2011. He is currently an Associate Professor with the School of Computer Science, Beijing Institute of Technology, Beijing, China. His research interests include image analysis and image quality assessment.

XIAONING SONG (Member, IEEE) received the B.S. degree in computer science from Southeast University, Nanjing, China, in 1997, the M.S. degree in computer science from the Jiangsu University of Science and Technology, Zhenjiang, China, in 2005, and the Ph.D. degree in pattern recognition and intelligent system from the Nanjing University of Science and Technology, Nanjing, in 2010. From 2014 to 2015, he was a Visiting Scholar with the Centre for Vision, Speech, and Signal Processing, University of Surrey, Guildford, U.K. He is currently a Full Professor with the School of Artificial Intelligence and Computer, Jiangnan University, Wuxi, China. His current research interests include pattern recognition, machine learning, and computer vision.

XIAOJUN WU (Member, IEEE) received the B.S. degree in mathematics from Nanjing Normal University, Nanjing, China, in 1991, and the M.S. and Ph.D. degrees in pattern recognition and intelligent system from the Nanjing University of Science and Technology, Nanjing, in 1996 and 2002, respectively. From 1996 to 2006, he was with the School of Electronics and Information, Jiangsu University of Science and Technology, where he was involved in teaching and promoted to a Professor. From 1999 to 2000, he was a Fellow of the International Institute for Software Technology, United Nations University. From 2003 to 2004, he was a Visiting Researcher with the Centre for Vision, Speech, and Signal Processing, University of Surrey, U.K. Since 2006, he has been with the School of Information Engineering, Jiangnan University, where he is currently a Professor of computer science and technology. He has published over 200 articles. His current research interests include pattern recognition, computer vision, and computational intelligence. He received the Most Outstanding Postgraduate Award from the Nanjing University of Science and Technology.