Frontiers of Information Technology & Electronic Engineering
www.zju.edu.cn/jzus; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Image quality assessment method based on nonlinear feature extraction in kernel space*

Yong DING[†‡], Nan LI[†], Yang ZHAO, Kai HUANG

(*Institute of VLSI Design, Zhejiang University, Hangzhou 310027, China*)

[†]E-mail: dingy@vlsi.zju.edu.cn; linan@vlsi.zju.edu.cn

Received Dec. 7, 2015; Revision accepted July 4, 2016; Crosschecked Sept. 26, 2016

Abstract: To match human perception, extracting perceptual features effectively plays an important role in image quality assessment. In contrast to most existing methods that use linear transformations or models to represent images, we employ a complex mathematical expression of high dimensionality to reveal the statistical characteristics of the images. Furthermore, by introducing kernel methods to transform the linear problem into a nonlinear one, a full-reference image quality assessment method is proposed based on high-dimensional nonlinear feature extraction. Experiments on the LIVE, TID2008, and CSIQ databases demonstrate that nonlinear features offer competitive performance for image inherent quality representation and the proposed method achieves a promising performance that is consistent with human subjective evaluation.

Key words: Image quality assessment, Full-reference method, Feature extraction, Kernel space, Support vector regression

http://dx.doi.org/10.1631/FITEE.1500439 **CLC number:** TP753

1 Introduction

Image quality assessment (IQA) plays an important role in image processing systems, which can be used as feedback for evaluation, optimization, and monitoring. Since subjective quality metrics are costly, time-consuming, and impractical, they cannot be integrated within real-world systems (Zhang *et al.*, 2015). This triggers the need to develop reliable objective quality assessment to replicate human perception by using computational models of visual appearance and potential mathematical expressions of images. It is desirable to predict the perceived visual quality as human subjective perception.

In terms of methodology, early works focused on error visibility, which means treating an image as the sum of the original reference and error signals. Methods such as mean squared error (MSE) and peak signal-to-noise ratio (PSNR) have been applied in IQA because they are simple to calculate and have clear physical meanings (Wang *et al.*, 2003; 2004). Over the last decades, various effective IQA metrics have been proposed. Since the characteristics and quality of features determine the performance of the IQA method, the key difficulty is to model the similarities and differences in a set of computational tractable features which can represent the crucial content and information closely related to the inherent quality of images. A structure similarity index measure (SSIM) (Wang *et al.*, 2004) and other decomposition methods have been developed based on the simple fact that natural images are highly structured, and these methods simulate the basic function of the human visual system well

‡ Corresponding author

* Project supported by the National High-Tech R&D Program (863) of China (No. 2015AA016704c), the National Science & Technology Support Program of China (No. 2013BAH03B01), and the Zhejiang Provincial Natural Science Foundation of China (No. LY14F020028)

ORCID: Yong DING, http://orcid.org/0000-0002-5226-7511

(Sheikh *et al.*, 2005; Zhang *et al.*, 2014). Currently, researchers are attempting to reveal the statistical characteristics within the natural scene, which is less subjective and drives more applications of statistical tools and methods (Zhang and Chandler, 2013; Ding and Dai, 2014). It is hypothesized that distortion disrupts the normal statistical property of an original image regardless of the type of distortion. Based on this principle and model, statistical analysis is superior to that in previous methods.

In the context of feature extraction, independent component analysis which decomposes image data by a linear transformation turns out to be a powerful tool (Ding and Dai, 2014). However, in general, such methods fail to cope with nonlinear problems because of unique and complicated mathematical expressions in coded digital image data. To overcome the drawback of independent component analysis, Bach and Jordan (2003) proposed kernel independent component analysis (KICA) based on minimizing mutual information on the entire function space of nonlinearities using the kernel method. The kernel method or kernel trick has been widely used in learning and optimization algorithms as a nonlinear similarity measure. Some applications based on KICA have been proposed in research areas such as face recognition, blind source separation, and image watermark (Yang *et al.*, 2005; Li *et al.*, 2007), but none of the KICA-based works address image quality assessment. In this study, an IQA method based on feature extraction in nonlinear kernel space is proposed. After analyzing the nonlinear feature representation in three color channels of distorted and original images, image quality evaluation is given by a pooling strategy.

According to the availability of an original image, IQA can be classified into full-, reduced-, and no-reference methods (Zhang *et al.*, 2014). Considering the complexity and difficulty of exploiting the commonality and difference among digital images based on nonlinear features, we focus only on the full-reference method. The basic idea of full-reference IQA is to find the visual or statistical similarities among the original images (images without distortion for reference) and the difference between the reference images and distorted images, and then to quantify and synthesize the difference into a distortion index by a certain pooling strategy. Reduced- and no-reference methods measure the image quality with only partial information or no information of reference images being available, respectively (Li and Wang, 2009; Rehman and Wang, 2012; Ma *et al.*, 2013; Wu *et al.*, 2015; 2016).

The framework of the proposed method is illustrated in Fig. 1. In summary, there are three processing stages. In the color space conversation stage, the input reference and distorted images are separated into RGB channels. In the feature extraction stage, nonlinear features for each channel are extracted after dimension reduction. Finally, with a dedicated support vector machine based pooling strategy, the aggregation evaluation is given by synthesizing the features in each channel.

2 Related work

Based on the hypothesis that visual perception is highly adapted for extracting structural information from a scene, most existing IQA methods are based on structure information or decomposition (Wang *et al.*, 2002). For example, SSIM is a representative method which combines luminance, contrast, and structure to simulate a human visual system. It is widely used in image processing systems because of its good performance and simplicity and has been extended by many researchers (Rao

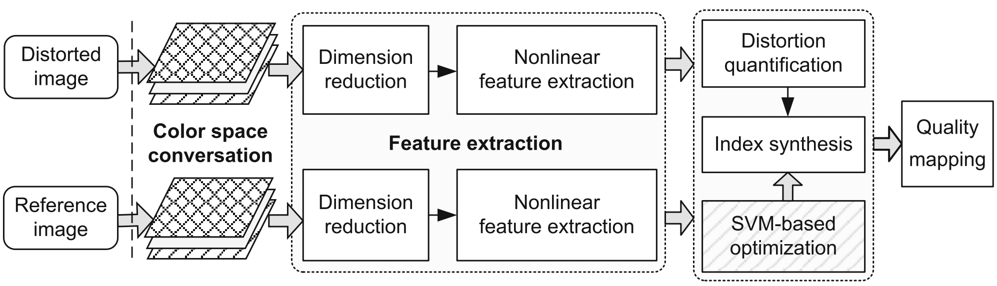

Fig. 1 Framework of the proposed method

and Reddy, 2009; Wang and Li, 2011). Yang and Kaveh (2010) proposed a full-reference IQA method measuring the change in the angle between the principal singular vectors from original and distorted image blocks. The change is used to quantize the loss of structural content after distortion. Multi-scale geometric analysis emulated by a contourlet transform was applied in Liu and Yang (2009). These transforms have advantages in multi-scale or multi-direction aspects in analyzing images. In general, these methods using linear features combined with a linear or nonlinear aggregation method to prove image quality find it difficult to analyze nonlinear distortion types of noise in a separated way while ignoring most of the complex image information.

Recently, natural image statistics (NIS) has been introduced into IQA based on quantization of statistical differences affected by a distortion (Zhang *et al.*, 2011; Liu *et al.*, 2012). Hypothesizing that distortion would affect the statistics characteristics of features, NIS provides a promising way to assess image quality that is different from structure-based methods. Most decomposition methods like independent component analysis (ICA) decompose images or image patches into orthogonal components. Then statistical characteristics are summarized from these components and known as features (Chang *et al.*, 2015). However, the distortion types are hard to separate using these methods and this still limits the development of a metric.

As discussed above, feature extraction is a critical step for image quality quantification. There should be some restrictions on the extracted features. For example, features should be as independent as possible from each other; otherwise, it is difficult to deal with features carrying redundancy information in a pooling strategy. The limited number of features is another reason for independency as well as consideration of computation cost. It turns out that the obtained components extracted by many feature extraction methods (Mittal *et al.*, 2012) relying on linear expressions by different kinds of decomposition and transformation in the spatial and frequency domains, are not really independent. For example, an ICA model finds the most independent components that are possible by a linear transformation, but a linear transformation has so few parameters that the estimated components are often quite far from being independent (Hyvärinen *et al.*, 2009). Moreover, linear methods cannot separate noise and error information thoroughly. In practice, the expressions of digital images should be nonlinear in a high-dimensional space representing the implicit content within an image, which turns out to be very similar to complex cells. As we all know, coefficients of high-order terms of a polynomial equal zero, which changes a problem from nonlinear to linear. In this particular nonlinear space, the complex mix of image content and error information may be easy to separate and measure. Applying the same method to the original and distorted images, we can compare the features in pairs to tell the difference. Thus, measuring nonlinear features to evaluate image quality is more comprehensive, efficient, and promising.

Although the state-of-the-art IQA methods perform well in terms of computation efficiency or concordance with human vision, in this study, we focus on exploring the nonlinear features to evaluate image quality in a way that is different from conventional linear decomposition or transformation.

3 Nonlinear feature extraction and image distortion quantification

3.1 Dimension reduction by principal component analysis

Dimension reduction is developed along with larger datasets and a larger number of variables with more observations. High-dimensional datasets provide many opportunities and mathematical challenges. Less but important information clears the restriction of some computationally expensive methods. In general, to keep important information as much as possible, the original data can be processed in a covariance matrix based on second-order statistics or be projected into a low-dimensional space. Traditional methods for dimension reduction include principal component analysis (PCA) (Jolliffe, 2002), random projections, multi-dimensional scaling, etc.

In this study, dimension reduction is used to specify the preprocessing of the image for further feature extraction. We are facing two major problems in nonlinear feature extraction. For one thing, the cardinality of the set of features of an image is very large and unknown, so the result will be more accurate with a larger number of samples. Considering the computing issues and algorithm complexity of KICA, data should be dimensionally reduced efficiently in advance. The other reason is that, when dealing with a small window size, it is possible with a small probability that two or more image patches contain the same content. However, the problem is that KICA offers only a solution for full rank data. The data must be sufficiently preprocessed.

Although PCA is not a successful model in terms of modeling a visual system, it provides the basis for the subsequent nonlinear feature extraction by KICA. Therefore, we employ PCA as the preprocessing tool for reducing the dimension of the input images where the maximum amount of the variance is preserved. For the implementation of PCA, readers are referred to Hyvärinen *et al.* (2009) and Abdi and Williams (2010). As an example, the data after preprocessing by PCA is given in Fig. 2.

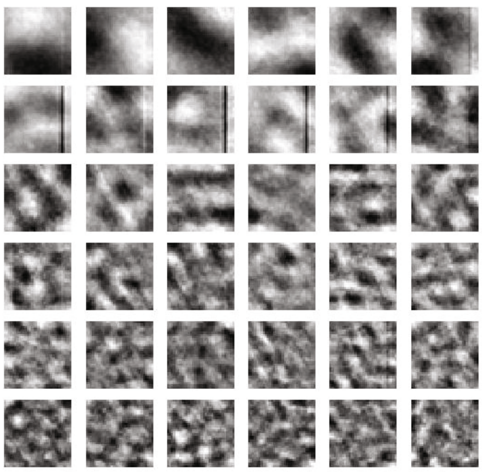

Fig. 2 Data derived by principal component analysis

3.2 Nonlinear feature extraction in kernel space

In general, independent component analysis aims to find the statistically independent factors or factors as independent as possible, which constitute the observed variables through linear combination. The statistical independency ensures that each component can be modeled, processed, and compared separately. On the condition that observation $\boldsymbol{x} = [x_1, x_2, \ldots, x_d]^{\mathrm{T}}$ from an image can be regarded as a product of a mixing matrix of content components and distortion components, $\boldsymbol{A} \in \mathbb{R}^{d\times m}$ represents weights, and component set $\boldsymbol{s} = [s_1, s_2, \ldots, s_m]^{\mathrm{T}}$, it is possible to estimate the components ($\hat{\boldsymbol{s}}$) and the demixing matrix $\boldsymbol{W}$ from the observation with no prior knowledge:

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} \Rightarrow \hat{\boldsymbol{s}} = \boldsymbol{W}\boldsymbol{x}. \tag{1}$$

Here, the demixing matrix is used only to separate the component matrix which contains weights of components to obtain original image patches. The extracted components, which we call features, represent the original image. However, linear decomposition is better at dimension reduction but not data separation. As discussed above, we need to find a method to transform the linear demixing or decomposition problem into a nonlinear one. In such a situation, the kernel method is considered to be a good solution and is introduced here. Actually, the kernel method owes its name to kernel functions with the capability of turning any linear model into nonlinear ones (Schölkopf and Smola, 1998). In other words, if a point set in a low dimension is hard to separate using a plane, after transformation into a high dimension, separation would be possible. Thus, data in a low-dimensional space may be better classified in a high dimension, and the description of data in a high dimension is better.

The nonlinear transformation is processed by mapping the data of image patches into an implicit reproducing kernel Hilbert space (RKHS):

$$\phi : \boldsymbol{x} \in \mathbb{R}^n \to \phi(\boldsymbol{x}) \in \boldsymbol{F}, \tag{2}$$

where $\boldsymbol{x}$ stands for the input image data, $\boldsymbol{F}$ is an RKHS on $\mathbb{R}$, $\phi(\boldsymbol{x}) = K(\cdot, \boldsymbol{x})$ is the feature map, and $K(\cdot, \boldsymbol{x})$ is a kernel function in $\boldsymbol{F}$ for each $\boldsymbol{x}$. To obtain statistical independency, the $\boldsymbol{F}$-correlation between two features within the same image data is denoted as

$$\begin{aligned}\rho_{\boldsymbol{F}} &= \max_{f_1, f_2 \in \boldsymbol{F}} \operatorname{corr}(f_1(\boldsymbol{x}_1), f_2(\boldsymbol{x}_2)) \\ &= \max_{f_1, f_2 \in \boldsymbol{F}} \frac{\operatorname{cov}(f_1(\boldsymbol{x}_1), f_2(\boldsymbol{x}_2))}{\sqrt{\operatorname{var}(f_1(\boldsymbol{x}_1)) \cdot \operatorname{var}(f_2(\boldsymbol{x}_2))}},\end{aligned} \tag{3}$$

where $f_1(\boldsymbol{x}_1)$ and $f_2(\boldsymbol{x}_2)$ are any two features from the RKHS of the input image.

The contrast function takes the following form:

$$I_{\rho_{\boldsymbol{F}}} = -\frac{1}{2}\log(1 - \rho_{\boldsymbol{F}}), \tag{4}$$

which is always non-negative and equals zero if and only if the variables are independent.

Based on the definitions above, KICA attempts to separate individual nonlinear components by estimating and minimizing the $\boldsymbol{F}$-correlation of the transformed image patch:

$$\hat{\rho}_{\boldsymbol{F}} = \max_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^m} \frac{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{K}_i\boldsymbol{K}_j\boldsymbol{\beta}}{\sqrt{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{K}_i^2\boldsymbol{\alpha}}\sqrt{\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{K}_j^2\boldsymbol{\beta}}}, \tag{5}$$

$$C(\boldsymbol{W}) = \hat{I}_{\rho_{\boldsymbol{F}}}(\boldsymbol{K}_1, \boldsymbol{K}_2, \ldots, \boldsymbol{K}_m), \tag{6}$$

where $\boldsymbol{K}_i$ ($i \in \mathbb{N}_+$) is the Gram matrix for each image patch.

In this way, nonlinear component extraction is turned into calculating the demixing matrix $\boldsymbol{W}$ in Eq. (6). By computing the inner products between the images of all pairs of data in the feature space instead of the coordinates of the data in the original space, the computation difficulty and cost are decreased greatly. A sample of features extracted by nonlinear decomposition KICA is given in Fig. 3.

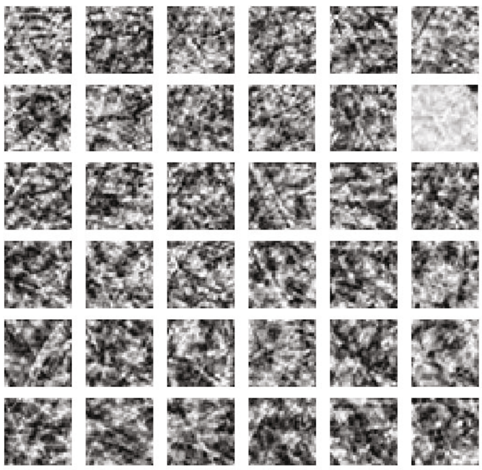

Fig. 3 Features derived by kernel independent component analysis

3.3 Distortion quantification and quality mapping

In general, with the features extracted from the reference and testing images, the distortion index of the testing image can be quantified by measuring discrepancies of features from that of the reference image. Because of the robustness of KICA to probability density functions, deviations of distributions among the kernel components are not obvious. As an example, we plot the distribution of a random sample of kernel independent components in Fig. 4. Three distributions are shown: one is fitted from the features of a reference image ('monarch.bmp' in the LIVE database), one is from a distorted image with serious fast-fading distortion ('img136.bmp' in the LIVE database) but the same content as the first reference image, and the third one is from another original image with different content ('ocean.bmp' in the LIVE database). Obviously, although there is a tremendous difference of feature meaning and distortion degree, the feature distributions exhibit almost the same probability trends. Therefore, the discrepancies and deviations between the nonlinear components of each patch extracted from a distorted image and the corresponding components from the reference image are hardly measured by conventional methods, such as Kullback-Leibler distance and Euclidean distance. In this study, we use the correlation coefficient to quantify such deviations:

$$\mathrm{corr}(C_i(x,y)) = \frac{\sum_{j=0}^{m}(x_j-\bar{x})(y_j-\bar{y})}{\sqrt{\sum_{j=1}^{m}(x_j-\bar{x})^2 \cdot \sum_{j=1}^{m}(y_j-\bar{y})^2}}, \tag{7}$$

where x and y represent every estimated component from the reference and distorted images, respectively, m is the number of image patches, and $\bar{x}$ and $\bar{y}$ are the mean values. In the set of features derived from the demixing result of content patches, the implicit meanings are distinguishable consisting of distortion, redundancy, and image content mathematically. If distortion is slight, the features extracted from the original and distorted images are more relevant. Taking an extreme example, if there is no distortion, the

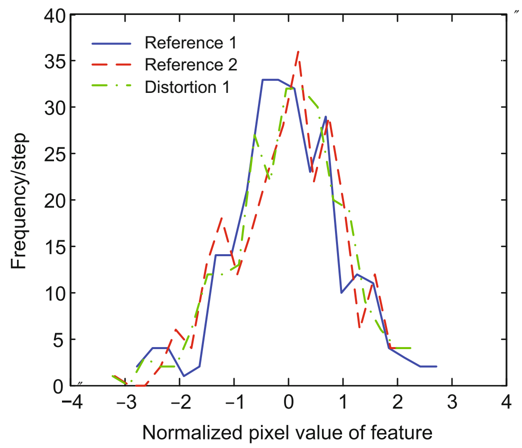

Fig. 4 Distribution of sample kernel independent components

correlation of the two matrices is 1.

Then such deviations of all the components in the RGB color space are synthesized into a distortion index with a pooling strategy. Instead of the conventional pooling strategy that uses the Minkowski equation which carries with it the tedious business of weight coefficient optimization, in this study, support vector regression (SVR) is employed to learn the optimal weights of each extracted component (feature) by supervised training. Thus, the distortion index of all the components in each channel is combined by an SVR fuser, i.e.,

$$D_{\mathrm{r,g,b}} = \sum_{i=0}^{N} (w_i |\mathrm{corr}(C_i)|), \tag{8}$$

where w_i is the weight for the ith component which is obtained from SVR training and N is the total number of components in each color channel. We choose the radial basis function (RBF) as the kernel function in SVR. The SVR training and predicting platform is LIBSVM (Chang and Lin, 2011).

Based on the fact that human vision is more sensitive to green channel than red and blue channels, the final index aggregated from features of each channel is synthesized and evaluated as

$$D = 0.299 D_{\mathrm{r}} + 0.587 D_{\mathrm{g}} + 0.114 D_{\mathrm{b}}. \tag{9}$$

Finally, a nonlinear mapping function with five parameters is employed to map the distortion index to an objective score (Video Quality Experts Group, 2003; Wang and Li, 2011):

$$Q(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1+\exp[\beta_2(x-\beta_3)]} \right) + \beta_4 x + \beta_5, \tag{10}$$

where x is the image distortion index obtained from Eq. (9) and β_1–β_5 are the parameters for optimization.

4 Experimental results and discussion

Experimental results and comparisons on three public databases, including LIVE (http://live.ece.utexas.edu/research/quality), TID2008 (http://www.ponomarenko.info/tid2008.htm), and CSIQ (Larson and Chandler, 2010), are presented in this section. In each database, there are hundreds of color images contaminated by a variety of distortion types as well as their corresponding reference images. For each distorted image, the databases provide a subjective evaluation score, e.g., mean opinion score (MOS) or difference mean opinion score (DMOS), which is obtained by psychometric tests. The main characteristics of the three databases are summarized in Table 1.

Three criteria are employed for quantitative performance evaluation of IQA methods, i.e., Pearson linear correlation coefficient (PLCC), Spearman rank-order correlation coefficient (SRCC), and root mean square error (RMSE). PLCC is an indicator of prediction accuracy. SRCC operates on the rank of the data points and ignores the relative distance between data points. It is used to evaluate prediction monotonicity. RMSE is used to evaluate the prediction consistency with subjective scores. The larger values of PLCC and SRCC as well as the smaller value of RMSE indicate a better performance of the IQA method.

For performance evaluation, the proposed method is compared with some representative methods, including ASVD (Yang and Kaveh, 2010), MSDD (Liu and Yang, 2009), VIF (Sheikh *et al.*, 2005), MS-SSIM (Wang *et al.*, 2003), GSM (Liu *et al.*, 2012), SSIM (Wang *et al.*, 2004), QDFS (Zhang *et al.*, 2014), and FSIM (Zhang *et al.*, 2011). PSNR is selected as the basic IQA because of its simplicity and superior capability in predicting the quality of images with additive noise (Liu *et al.*, 2013).

4.1 Performance evaluation on individual distortion types

To examine the IQA performance on each distortion type, a thorough performance evaluation is conducted on the LIVE database. The LIVE database contains 174 fast-fading images, 174 Gauss

Table 1 Databases for performance evaluation

Database name	Number of reference images	Number of distorted images	Number of distortion types	Image size (pixel)
LIVE	29	779	5	Various
TID2008	25	1700	17	512×384
CSIQ	30	866	6	512×512

blurred (Gblur) images, 227 JPEG2000 (JP2K) images, 233 JPEG images, and 174 white noise images. In the experiments, the Gauss kernel is chosen for population and precision (Genton, 2001).

Detailed performance comparison with the competing methods on different image distortion types in the LIVE database is listed in Table 2, where SRCC is used as the evaluation measure and the best result of each distortion type has been highlighted in boldface. Using other measures such as PLCC and RMSE, similar conclusions can be drawn. The proposed method achieves a comparatively larger SRCC on each distortion type, validating it as a powerful method for image quality assessment. Although JPEG and JPEG2000 produced significant differences in edges and structures, the proposed method achieves results equivalent to algorithms based on texture analysis and visual characteristics. Since white noise is an additive noise such that PSNR outperforms other methods directly calculating the difference, the performance of the proposed method indicates that nonlinear features can also be used to measure linear distortion. In addition, the proposed method performs best among these methods on fast-fading which is produced during transmission channel fading, showing that nonlinear features have advantages in measuring distortions with local differences over distortions with global differences. In general, the competitive experimental results demonstrate that the proposed method is a stable and powerful model in accordance with the perceptual quality.

Table 2 Spearman rank-order correlation coefficient (SRCC) comparison on individual distortion types in the LIVE database

Method	SRCC				
	JP2K	JPEG	WN	Gblur	FF
PSNR	0.8954	0.8809	**0.9854**	0.7823	0.8907
ASVD	0.9146	0.9112	0.9425	0.8226	0.9048
MSDD	0.8991	0.8828	0.9461	0.9480	0.9226
SSIM	0.9614	0.9764	0.9694	0.9517	0.9556
MS-SSIM	0.9654	0.9793	0.9731	0.9584	0.9321
VIF	0.9683	**0.9842**	0.9845	**0.9722**	0.9652
FSIM	0.9717	0.9834	0.9652	0.9708	0.9499
GSM	**0.9759**	0.9392	0.8577	0.9589	0.8925
QDFS	0.9603	0.9517	0.9656	0.9527	0.9415
Proposed	0.9639	0.9766	0.9724	0.9633	**0.9658**

JP2K: JPEG2000 images; WN: white noise images; Gblur: Gauss blurred images; FF: fast-fading images. The best result of each distortion type is in boldface

In addition, the scatter diagrams of the proposed method on the LIVE database are exhibited in Fig. 5. Fig. 5a presents the scatter plots of the objective scores for the entire LIVE database of the proposed method versus subjective DMOS, and Figs. 5b–5f show the subjective ratings of perception versus predicted values for each type of distortion. If the predicted score reflects the DMOS faithfully, scatter plots should be close to the fitted curve. Intuitively, with most points close to the fitted logistic curve in the scatter plots as shown in Fig. 5, the proposed method provides good prediction of DMOS and stability on different distortion types.

4.2 Performance comparison on different databases

To validate the performance and robustness of IQA schemes on different databases, the comparison on all types of distortions in the three databases is listed in Table 3, where the two best results for each criterion on each database have been highlighted in boldface. From Table 3, we can see that the proposed

Table 3 Performance comparison on different databases

Database	Criterion	PSNR	SSIM	MS-SSIM	MSDD	FSIM	VIF	GSM	Proposed
LIVE	PLCC	0.8723	0.9449	0.9409	0.8900	**0.9597**	**0.9598**	0.9437	0.9476
	SRCC	0.8756	0.9479	0.9513	0.8901	**0.9634**	**0.9632**	0.9554	0.9543
	RMSE	13.3597	8.9454	9.2593	**7.3413**	7.6780	7.6670	9.0376	**7.6502**
TID2008	PLCC	0.5726	0.7710	0.8451	–	**0.8738**	0.8090	0.8462	**0.8467**
	SRCC	0.5794	0.7749	0.8542	–	**0.8805**	0.7496	**0.8554**	0.8488
	RMSE	1.1003	0.8546	0.7173	–	**0.6525**	0.7888	**0.7151**	0.7210
CSIQ	PLCC	0.7998	0.8612	0.8990	–	0.9120	**0.9227**	0.8979	**0.9143**
	SRCC	0.8005	0.8756	0.9133	–	**0.9242**	**0.9195**	0.9126	0.9124
	RMSE	0.1576	0.1334	0.1150	–	0.1077	**0.0980**	0.1156	**0.1068**

PLCC: Pearson linear correlation coefficient; SRCC: Spearman rank-order correlation coefficient; RMSE: root mean square error. Two best results for each criterion on each database are in boldface

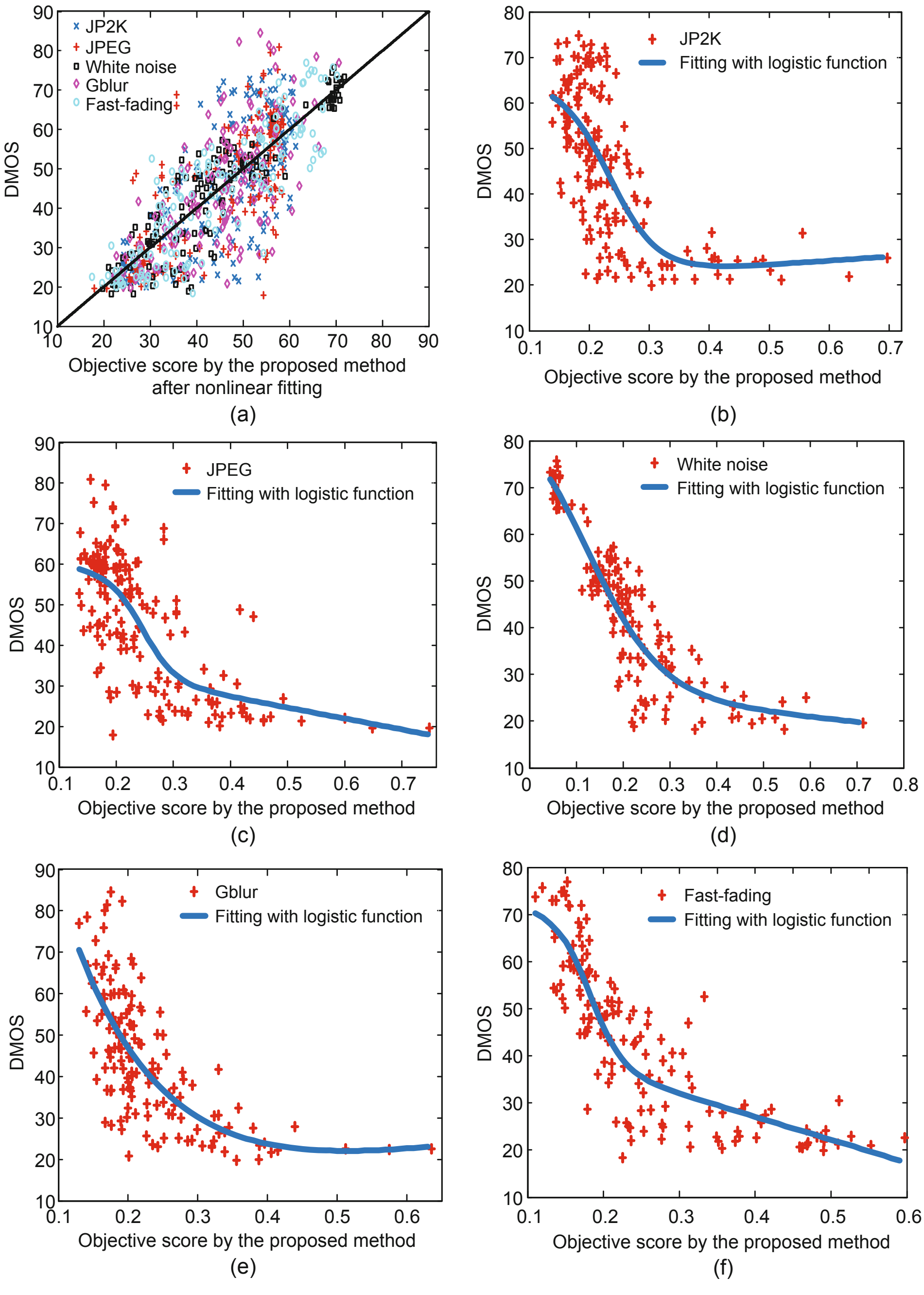

Fig. 5 Scatter diagrams of the proposed method on different distortion types in the LIVE database: (a) for the whole database; (b) for JPEG2000 distortion; (c) for JPEG distortion; (d) for white noise distortion; (e) for Gauss blurred distortion; (f) for fast-fading distortion

method, FSIM, and VIF give the best performance on almost all the three databases.

The robustness of the proposed method on different databases is also demonstrated in Table 3. Clearly, VIF is not good for TID2008 although it is the best for LIVE. Likewise, GSM is the best for TID2008 but is relatively poor for the LIVE and CSIQ databases. The proposed scheme and FSIM give more consistent and stable performances across all the three databases in comparison with the other schemes. The experimental results demonstrate the robustness of our method across different databases.

5 Conclusions

In this paper, a novel image quality assessment method has been proposed as a successful attempt to extract nonlinear features of digital images and apply them in the framework of IQA. In comparison with existing image quality assessment approaches, the results show that nonlinear features have equivalent and even better ability to evaluate the image quality in accord with the human visual system without considering human visual properties in IQA development. The accuracy and robustness across different distortion types and databases demonstrate the proposed method to be a general-purpose and stable approach. Future works include improving the computation efficiency and studying the characteristics of features further for no-reference IQA applications.

References

Abdi, H., Williams, L.J., 2010. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.*, **2**(4):433-459. http://dx.doi.org/10.1002/wics.101

Bach, F.R., Jordan, M.I., 2003. Kernel independent component analysis. Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, p.IV-876-9. http://dx.doi.org/10.1109/ICASSP.2003.1202783

Chang, C.C., Lin, C.J., 2011. LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, **2**(3):Article 27. http://dx.doi.org/10.1145/1961189.1961199

Chang, H.W., Zhang, Q.W., Wu, Q.G., *et al.*, 2015. Perceptual image quality assessment by independent feature detector. *Neurocomputing*, **151**:1142-1152. http://dx.doi.org/10.1016/j.neucom.2014.04.081

Ding, Y., Dai, H., 2014. Image quality assessment scheme with topographic independent components analysis for sparse feature extraction. *Electron. Lett.*, **50**(7):509-510. http://dx.doi.org/10.1049/el.2013.4298

Genton, M.G., 2001. Classes of kernels for machine learning: a statistics perspective. *J. Mach. Learn. Res.*, **2**:299-312.

Hyvärinen, A., Hurri, J., Hoyer, P.O., 2009. Natural Image Statistics: a Probabilistic Approach to Early Computational Vision. Springer-Verlag London, UK.

Jolliffe, I., 2002. Principal Component Analysis. Springer-Verlag New York, USA.

Larson, E.C., Chandler, D.M., 2010. Most apparent distortion: full-reference image quality assessment and the role of strategy. *J. Electron. Imag.*, **19**(1):011006. http://dx.doi.org/10.1117/1.3267105

Li, Q., Wang, Z., 2009. Reduced-reference image quality assessment using divisive normalization-based image representation. *IEEE J. Sel. Topics Signal Process.*, **3**(2):202-211. http://dx.doi.org/10.1109/JSTSP.2009.2014497

Li, Y.C., Wu, K.H., Ma, Y.L., *et al.*, 2007. Image digital watermarking technique based on kernel independent component analysis. Proc. 11th Int. Conf. on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, p.467-474. http://dx.doi.org/10.1007/978-3-540-72530-5_56

Liu, A., Lin, W., Narwaria, M., 2012. Image quality assessment based on gradient similarity. *IEEE Trans. Image Process.*, **21**(4):1500-1512. http://dx.doi.org/10.1109/TIP.2011.2175935

Liu, M., Yang, X., 2009. Image quality assessment using contourlet transform. *Opt. Eng.*, **48**(10):107201. http://dx.doi.org/10.1117/1.3241996

Liu, T.J., Lin, W., Kuo, C.C.J., 2013. Image quality assessment using multi-method fusion. *IEEE Trans. Image Process.*, **22**(5):1793-1807. http://dx.doi.org/10.1109/TIP.2012.2236343

Ma, L., Li, S., Ngan, K.N., 2013. Reduced-reference image quality assessment in reorganized DCT domain. *Signal Process. Image Commun.*, **28**(8):884-902. http://dx.doi.org/10.1016/j.image.2012.08.001

Mittal, A., Moorthy, A.K., Bovik, A.C., 2012. No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.*, **21**(12):4695-4708. http://dx.doi.org/10.1109/TIP.2012.2214050

Rao, D.V., Reddy, L.P., 2009. Contrast weighted perceptual structural similarity index for image quality assessment. Proc. Annual IEEE India Conf., p.1-4. http://dx.doi.org/10.1109/INDCON.2009.5409432

Rehman, A., Wang, Z., 2012. Reduced-reference image quality assessment by structural similarity estimation. *IEEE Trans. Image Process.*, **21**(8):3378-3389. http://dx.doi.org/10.1109/TIP.2012.2197011

Schölkopf, B., Smola, A.J., 1998. Learning with Kernels. MIT Press.

Sheikh, H.R., Bovik, A.C., de Veciana, G., 2005. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Process.*, **14**(12):2117-2128. http://dx.doi.org/10.1109/TIP.2005.859389

Video Quality Experts Group, 2003. Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II (fr_tv2). Available from http://www.vqeg.org.

Wang, Z., Li, Q., 2011. Information content weighting for perceptual image quality assessment. *IEEE Trans. Image Process.*, **20**(5):1185-1198. http://dx.doi.org/10.1109/TIP.2010.2092435

Wang, Z., Bovik, A.C., Lu, L.G., 2002. Why is image quality assessment so difficult? Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, p.1-4. http://dx.doi.org/10.1109/ICASSP.2002.5745362

Wang, Z., Simoncelli, E.P., Bovik, A.C., 2003. Multiscale structural similarity for image quality assessment. Proc. 37th Asilomar Conf. on Signals, Systems and Computers, p.1398-1402. http://dx.doi.org/10.1109/ACSSC.2003.1292216

Wang, Z., Bovik, A.C., Sheikh, H.R., *et al.*, 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.*, **13**(4):600-612. http://dx.doi.org/10.1109/TIP.2003.819861

Wu, Q., Li, H., Meng, F., *et al.*, 2015. No reference image quality assessment metric via multi-domain structural information and piecewise regression. *J. Vis. Commun. Image Represent.*, **32**:205-216. http://dx.doi.org/10.1016/j.jvcir.2015.08.009

Wu, Q., Li, H., Meng, F., *et al.*, 2016. Blind image quality assessment based on multichannel feature fusion and label transfer. *IEEE Trans. Circ. Syst. Video Technol.*, **26**(3):425-440. http://dx.doi.org/10.1109/TCSVT.2015.2412773

Yang, C.A., Kaveh, M., 2010. Image quality assessment using singular vectors. Proc. SPIE, Article 752910. http://dx.doi.org/10.1117/12.839796

Yang, J., Gao, X., Zhang, D., *et al.*, 2005. Kernel ICA: an alternative formulation and its application to face recognition. *Patt. Recog.*, **38**(10):1784-1787. http://dx.doi.org/10.1016/j.patcog.2005.01.023

Zhang, H., Ding, Y., Huang, K., *et al.*, 2014. Image quality assessment by quantifying discrepancies of multifractal spectrums. *IEICE Trans. Inform. Syst.*, **97**(9):2453-2460. http://dx.doi.org/10.1587/transinf.2014EDP7036

Zhang, L., Zhang, L., Mou, X., *et al.*, 2011. FSIM: a feature similarity index for image quality assessment. *IEEE Trans. Image Process.*, **20**(8):2378-2386. http://dx.doi.org/10.1109/TIP.2011.2109730

Zhang, M., Muramatsu, C., Zhou, X., *et al.*, 2015. Blind image quality assessment using the joint statistics of generalized local binary pattern. *IEEE Signal Process. Lett.*, **22**(2):207-210. http://dx.doi.org/10.1109/LSP.2014.2326399

Zhang, Y., Chandler, D.M., 2013. No-reference image quality assessment based on log-derivative statistics of natural scenes. *J. Electron. Imag.*, **22**(4):043025. http://dx.doi.org/10.1117/1.JEI.22.4.043025