J. Vis. Commun. Image R. 32 (2015) 205–216

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

No reference image quality assessment metric via multi-domain structural information and piecewise regression ☆

Qingbo Wu [a,*], Hongliang Li [a], Fanman Meng [a], King Ngi Ngan [a,b], Shuyuan Zhu [a]

[a] School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu, China
[b] Department of Electronic Engineering, The Chinese University of Hong Kong, ShaTin, Hong Kong

ARTICLE INFO

Article history:
Received 17 October 2014
Accepted 15 August 2015
Available online 28 August 2015

Keywords:
No reference image quality assessment
Human vision system
Image representation
Quality aware feature
Multi-domain structural information
Gradient of wavelet domain
Piecewise regression
HEVC

ABSTRACT

The general purpose no reference image quality assessment (NR-IQA) is a challenging task, which faces two hurdles: (1) it is difficult to develop one quality aware feature which works well across different types of distortion and (2) it is hard to learn a regression model to approximate a complex distribution for all training samples in the feature space. In this paper, we propose a novel NR-IQA method that addresses these problems by introducing the multi-domain structural information and piecewise regression. The main motivation of our method is based on two points. Firstly, we develop a new local image representation which extracts the structural image information from both the spatial-frequency and spatial domains. This multi-domain description could better capture human vision property. By combining our local features with a complementary global feature, the discriminative power of each single feature could be further improved. Secondly, we develop an efficient piecewise regression method to capture the local distribution of the feature space. Instead of minimizing the fitting error for all training samples, we train the specific prediction model for each query image by adaptive online learning, which focuses on approximating the distribution of the current test image's *k*-nearest neighbor (KNN). Experimental results on three benchmark IQA databases (i.e., LIVE II, TID2008 and CSIQ) show that the proposed method outperforms many representative NR-IQA algorithms.

1. Introduction

No reference image quality assessment (NR-IQA) aims at designing a computational model to automatically predict the perceptual quality of a test image without its undistorted reference version [1–3]. In many practical applications, the reference images are unaccessible, which makes the NR-IQA algorithm more desirable in comparison with the full-reference (FR)-IQA [4–7] and the reduced-reference (RR)-IQA [8–12] metrics. For example, many denoising algorithms require the manual parameters to obtain a good result. Since there is no reference image, the NR-IQA is very desirable for parameter tuning, such as [13]. Similarly, the NR-IQA based image enhancement applications have also been discussed in recent works [14,15]. To date, many NR-IQA methods have been developed, which are usually composed of the quality aware feature extraction and perceptual quality regression modules.

The quality aware feature corresponds to an efficient image representation, which could capture the variation of the perceptual image quality caused by the distortion. Many efficient features have been proposed in existing algorithms. Seo et al. introduce the visual saliency into the perceptual quality metric in [16]. Ghanem et al. [17] utilize the inter/intra-segment interactions to measure the image quality variation. In [18,19], Moorthy et al. employ the parametric statistical model to describe the natural scene statistics (NSS) of an image in the wavelet domain. Meanwhile, some nonparametric wavelet statistics are also introduced in [20]. In [21,22], the NSS information from the multi-scale DCT domain is utilized. In [23], Tang et al. propose the α-stable model in the wavelet domain to describe the image, and Sang et al. [24] introduce the singular value curve to measure the image degradation caused by blur. After extracting the image features, the regression module is used to map these quality aware features to a subjective quality score. Many different regression schemes have been discussed in existing methods, such as, the general regression neural network (GRNN) [25], the support vector regression (SVR) [18,19] and the multiple kernel learning (MKL) [20].

Although the aforementioned methods have achieved good result in capturing the human perception, some important local

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author. Fax: +86 028 61830064.
E-mail addresses: wqb.uestc@gmail.com (Q. Wu), hlli@uestc.edu.cn (H. Li), fmmeng@uestc.edu.cn (F. Meng), knngan@ee.cuhk.edu.hk (K.N. Ngan), eezsy@uestc.edu.cn (S. Zhu).

http://dx.doi.org/10.1016/j.jvcir.2015.08.009
1047-3203/

properties in both the image representation and perceptual quality regression are still underutilized. Firstly, in the feature extraction module, the NSS is usually extracted from some specific domains [18–22]. When the transform coefficients are quantized into different bins, only the global frequency distribution is reserved and the local spatial neighboring relationship is lost. In [19,20], the neighboring coefficients' joint distribution were introduced to address this issue. However, the similar problem still exists in the coefficients' quantization process. Secondly, in the regression module, the existing methods usually learn a map function by off-line training, which focuses on minimizing the average fitting error over all training samples. For some local training data, its performance may not be very well [26].

Based on these analysis, we develop our method from two aspects: (1) For the image representation, we introduce a new local image representation, which is then combined with a widely-used global image feature. As discussed in [27,28], the visual cortex integrates both the spatial-frequency and spatial information. Most existing methods don't consider this characteristic and only extract the image features from single domain, such as, wavelet domain [18,19], DCT domain [21,22] or spatial domain [29]. In contrast, we describe the local image structures from both the spatial-frequency and spatial domains. Particularly, the spatial-frequency information is derived from a novel orientation statistics of the gradients in the local patch of each wavelet subband. The spatial information is captured with the classic local binary pattern (LBP) [30]. Since the human vision captures both the local and global information from the natural image [31], we further introduce the global distribution of the wavelet coefficients to compensate our local features. (2) For the perceptual quality regression, we design an efficient local regression method to further improve the prediction accuracy. Inspired by our previous work [32], the piecewise regression criterion [33] is employed in our method. Unlike previous single-phase regression [33] which uses all training data to learn a regressor, we try to build the appropriate segmented data or training sample subset for each test image. Then, the specific regression parameters can be learned from these training sample subset by online training. Experimental results on three IQA databases show that the proposed method is highly efficient and robust.

The rest of this paper is organized as follows. Sections 2 describes the quality-aware features in our method and Section 3 presents the proposed piecewise quality regression model. The experimental results are shown in Section 4. Finally, we draw the conclusion in Section 5.

2. Quality aware features

For the image representation, the spatial domain, transform domain [34,35], saliency [36,37] and segmentation [38–42] information have been widely used. Among these analysis tools, the wavelet transform is particularly popular due to its abilities in multiresolution analysis and spatial-frequency information representation [35]. For clarity, the spatial-frequency structure of four scales wavelet transform is illustrated in Fig. 1. The label in the top-right corner of each sub-image denotes its scale and direction information. The HL*x*, LH*x* and HH*x* represent the horizontal, vertical and diagonal details for the scaled image under the *x*th level. The block label in the top-left corner denotes the subbands with the same frequency from the fine scale to the coarse one. In this section, we describe the quality aware local and global features in details.

2.1. Local multi-domain structural information

In order to represent the local spatial-frequency structure in the wavelet domain, we partition each subband into non-overlapped cell units and count the distribution of the gradient in some quantized orientations in the form of histogram of oriented gradient (HOG) [43]. Here, the wavelet coefficient's gradient is first calculated by convolving the 1-D derivatives ($[-1; 0; 1]$) in the horizontal and vertical directions, where g_x and g_y denote the horizontal and vertical gradient in each location of a cell. Then, the gradient orientation G_o and the gradient magnitude G_m can be calculated as,

$$\begin{aligned} G_o &= \arctan(g_y/g_x) \\ G_m &= \sqrt{g_y^2 + g_x^2} \end{aligned} \tag{1}$$

As shown in Fig. 2, the gradient orientation histogram statistic is applied in each cell, where the histogram bins indicate the quantized gradient orientation ranging from -180° to 180°. The gradient magnitude voting scheme is employed in counting the histogram. In our method, we count the wavelet domain gradient

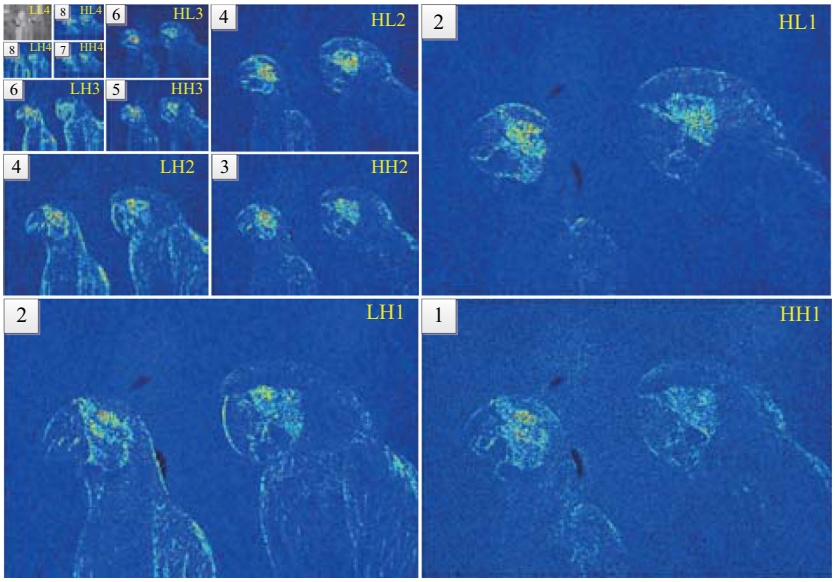

Fig. 1. The wavelet subband and frequency band locations.

orientation histogram across all of directions and scales. Let $h_{i,j}$ denote the normalized histogram for the wavelet subband at the ith direction and jth scale, where $i = 1, 2, 3$ denote the directions in HL, LH and HH. Then, we can obtain the cascaded multiscale histogram in each direction, i.e.,

$$\begin{aligned} h_{HL} &= [h_{1,1}, h_{1,2}, \ldots, h_{1,L}] \\ h_{LH} &= [h_{2,1}, h_{2,2}, \ldots, h_{2,L}] \\ h_{HH} &= [h_{3,1}, h_{3,2}, \ldots, h_{3,L}] \end{aligned} \tag{2}$$

Since the distortion presents in all color channels [32], we extract the wavelet domain gradient orientation histogram in the Y, Cb and Cr channels of a color image. Then, the multi-channel histogram feature H can be represented as

$$\begin{aligned} H^{Y} &= \left[h^{Y}_{HL}, h^{Y}_{LH}, h^{Y}_{HH}\right] \\ H^{Cb} &= \left[h^{Cb}_{HL}, h^{Cb}_{LH}, h^{Cb}_{HH}\right] \\ H^{Cr} &= \left[h^{Cr}_{HL}, h^{Cr}_{LH}, h^{Cr}_{HH}\right] \\ H &= [H^{Y}, H^{Cb}, H^{Cr}] \end{aligned} \tag{3}$$

where H^{Y}, H^{Cb} and H^{Cr} are the histograms in each color channel.

Fig. 3 shows some samples of the $h_{i,j}$ for the *manfishing* image and its associated distortion versions across all distortion types in the LIVE II database. In these examples, we partition each wavelet subband into 3×3 cells and quantize the gradient orientation into 9 bins in each cell. In this way, we can obtain 81 bins in computing the $h_{i,j}$. The difference mean opinion score (DMOS) of each image is labeled in the legend, where 0 corresponds to the undistorted image and a larger value represents a severer distortion.

As shown in Fig. 3, the $h_{i,j}$ present significant differences under different distortion degrees across all of the five distortion types. Since the JPEG2000 compression will eliminate the small wavelet coefficients and preserve the large ones, the gradient magnitude gap will grow in the wavelet domain, i.e., a large gradient will become larger and vice versa. As shown in Fig. 3(a), when a bin of $h_{i,j}$ shows high value in the reference image, its value would become higher for the JEPG2000 images. As the distortion level increasing, this gap will grow as well. Since the JPEG compression also suppress the high frequency information, it induces similar gradient variation in the wavelet domain, where the large value bins will become larger and the small value bins will further reduce as shown in Fig. 3(b). The White Noise introduces the high frequency interference to overwhelm the directional local structure in an image, which results in more homogeneous gradient distribution across different directions in the wavelet domain. As shown in Fig. 3(c), it is seen that the $h_{i,j}$ of the undistorted image presents peaks on some specific gradient orientation bins. However, the noisy images' $h_{i,j}$ is relatively uniform across all bins. For the images with higher DMOS, the variations in their $h_{i,j}$ are more gentle. Since the Blur operation also eliminates the original image structure, it also makes the wavelet domain gradient distribution more homogeneous. As shown in Fig. 3(d), we can find that the $h_{i,j}$ of the distorted images show smaller change across different bins than the natural image. For the Fast Fading, besides the wavelet domain quantization, there are also some random channel error, such as, the mismatch of the wavelet subbands. Thus, in Fig. 3(e), we can find that the distorted images' $h_{i,j}$ grows higher than the natural images in some bins just like Fig. 3(a). Meanwhile, we can also find some $h_{i,j}$ peak mismatches. That is, in some bins, the $h_{i,j}$ of the natural image could reach the peaks, and the distorted images can't reach.

Besides the spatial-frequency local structure in the scaled wavelet subband, we also employ the classic LBP descriptor [30] to capture the local structure distribution in the original scale spatial domain. It is noted that the LBP feature only preserves the relative intensities between the center pixel and its neighboring samples, which is too sensitive for some invisible small LBP difference in the chroma channels (i.e., Cb and Cr). Thus, we only extract the LBP feature in the gray image space like [32].

2.2. Global frequency distribution

As discussed in pervious works [18,19], the non-Gaussian property of the wavelet coefficients distribution is an important global clue in measuring the naturalness of an image. The parameterized models like generalized Gaussian distributions (GGD) [44] and

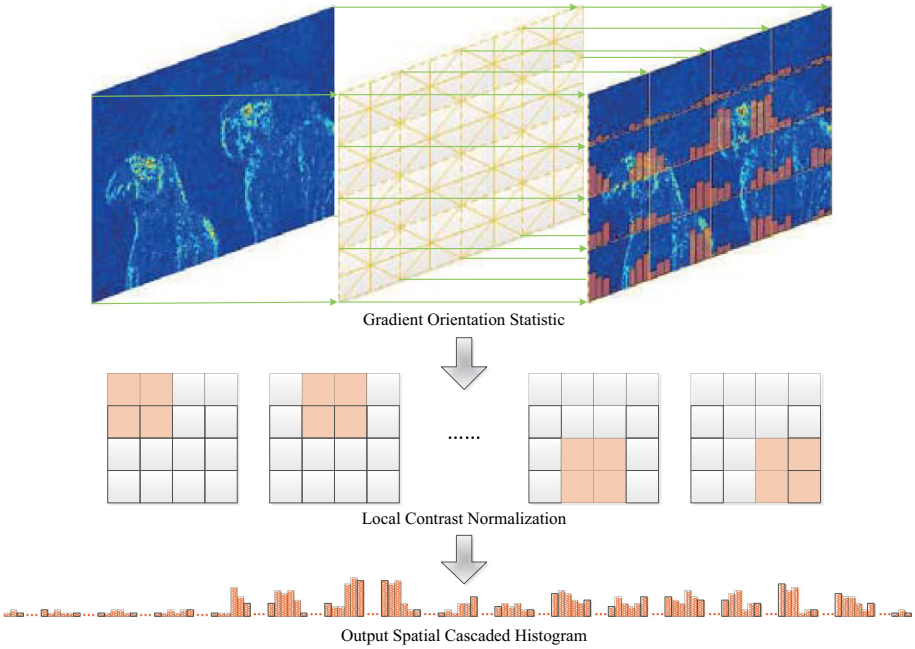

Fig. 2. The framework of wavelet domain gradient orientation statistics.

Gaussian scale mixture (GSM) [45] are often utilized to represent this non-Gaussianity.

Since the fitting error is inevitable in modeling the wavelet coefficients distribution, we directly use the margin distribution of the wavelet coefficients as the global frequency feature. Here, we only focus on the coefficients' (e.g., high and low) frequency distribution. Thus, the margin distribution is counted for the coefficients in the same frequency band as labeled in Fig. 1. Let C_k denote the coefficient set in the kth frequency band. The probability distribution p_k in this band can be represented as

$$p_k = norm_{l1}(hist(C_k)) \tag{4}$$

where $hist(\cdot)$ denotes the histogram statistic operator and $norm_{l1}(\cdot)$ represents the $l1$ normalization operator.

Similar with the local spatial-frequency structure statistic in the previous section, we also count the global frequency distribution P in all of the color channels, i.e.,

$$\begin{aligned} P^Y &= [p_1^Y, p_2^Y, \cdots, p_{2L}^Y] \\ P^{Cb} &= [p_1^{Cb}, p_2^{Cb}, \cdots, p_{2L}^{Cb}] \\ P^{Cr} &= [p_1^{Cr}, p_2^{Cr}, \cdots, p_{2L}^{Cr}] \\ P &= \left[P^Y, P^{Cb}, P^{Cr}\right] \end{aligned} \tag{5}$$

where P^Y, P^{Cb} and P^{Cr} are the wavelet coefficients probability distribution in each color channel.

Fig. 4 shows some samples of the p_k distribution for the *manfishing* image and its associated distortion versions under five distortion types in the LIVE II database. The DMOS of each image is also labeled in the legend. It can be seen that the p_k shows different distributions for the images assigned with different perceptual qualities. For the JPEG2000, JPEG, Blur and Fast Fading distortions, an important statistical expression is the high frequency loss, i.e., the more wavelet coefficients would gather to zero as the distortion degree becoming larger. In contrast, the wavelet coefficients would prefer distributing to the nonzero values due to the additive random noise for the white noise.

3. Piecewise regression based prediction model

In existing NR-IQA methods, there are two typical schemes for training the quality prediction model, i.e., one-step and two-step schemes. The one-step [21,22] scheme evaluates different distortion types equally and train single quality prediction function for all test images. In contrast, the two-step [18,19] scheme considers the differences across various distortion types and evaluates the

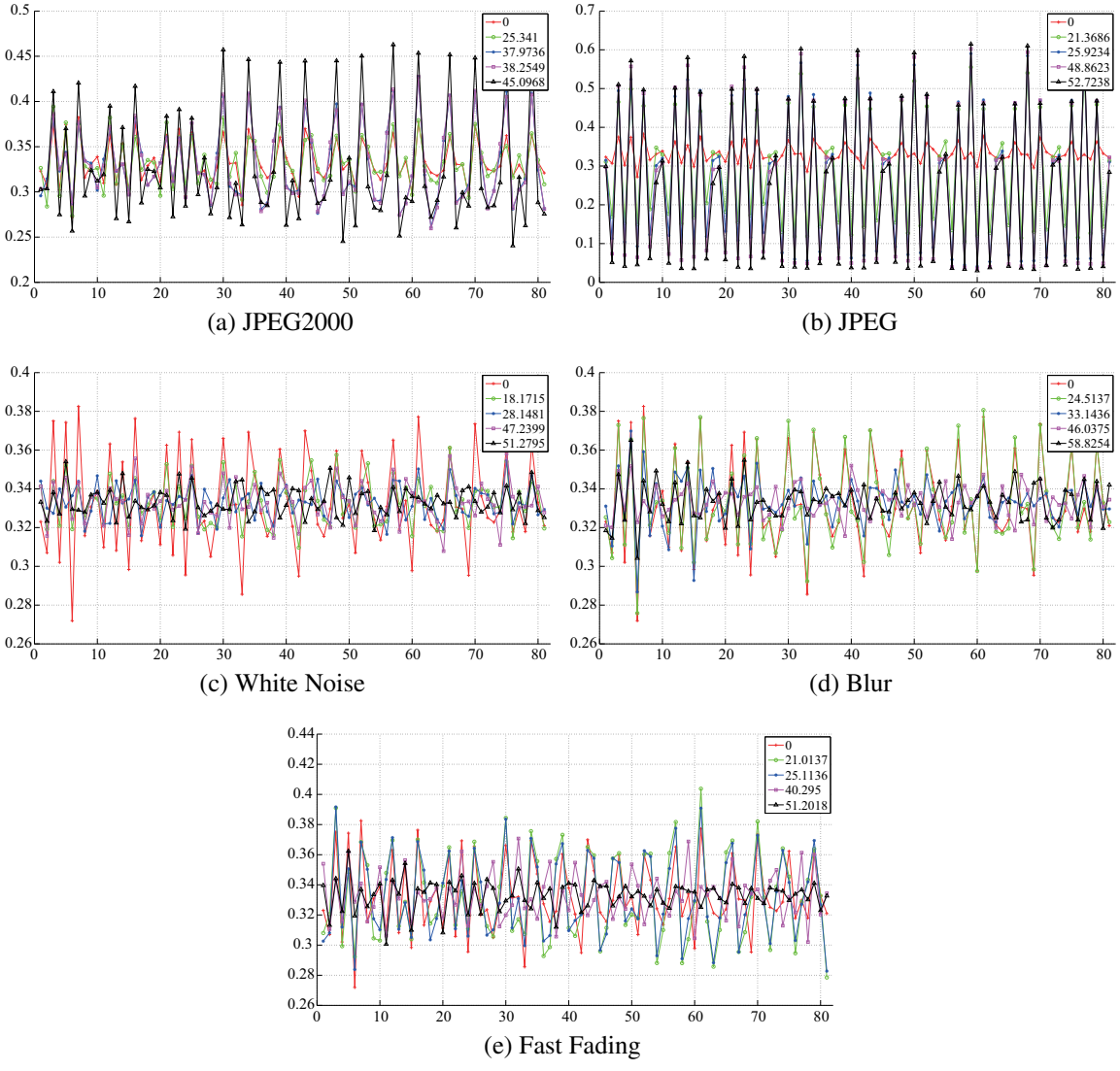

Fig. 3. The normalized histogram of oriented gradient in the wavelet subband under different distortion types. x-axis denotes the gradient orientations across all cells and y-axis denotes the occurrence frequency of each bin. Legend: the DMOS assigned for the images with different qualities.

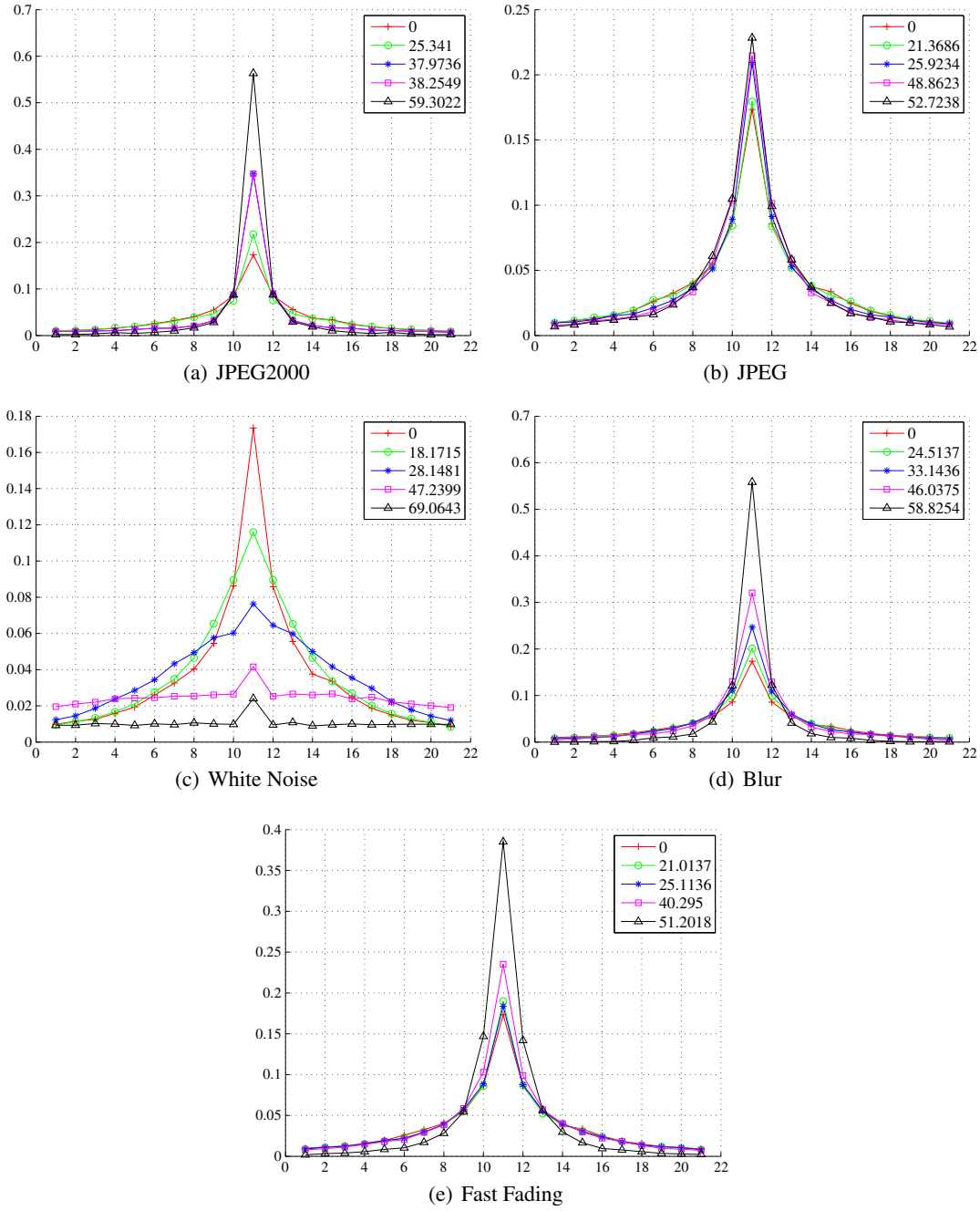

Fig. 4. The marginal distribution of wavelet coefficients in a frequency band under different distortion types. Legend: the DMOS assigned for the images with different qualities.

image quality with two operations, i.e., distortion identification and distortion-specific quality evaluation.

In this paper, we follow the two-step framework to design our prediction model, whose diagram is shown in Fig. 5. Firstly, a pre-trained classifier is utilized to identify the distortion type of an input image. Secondly, the subjective score is generated by a distortion-specific prediction function whose training data share the same distortion type with the test image.

It is noted that the conventional two-step scheme [18,19] uses all training samples to learn the distortion-specific regression model, which neglects the diversity of the samples under the same distortion type. To address this issue, we introduce the piecewise regression [33] which designs different prediction functions on different samples subset within one class of distortion type, i.e.,

$$f_c(\boldsymbol{x};\ \Theta_c, \boldsymbol{\alpha}_c) = \begin{cases} f_c(x_1;\ \theta_c(1)), & x_1 \in \alpha_c(1); \\ f_c(x_2;\ \theta_c(2)), & x_2 \in \alpha_c(2); \\ \vdots & \vdots \\ f_c(x_n;\ \theta_c(n)), & x_n \in \alpha_c(n). \end{cases} \tag{6}$$

where $f_c(\boldsymbol{x}, \Theta, \boldsymbol{\alpha}_c)$ denotes the parameterized quality prediction function for the cth class of distortion, $\boldsymbol{x}$ denotes all test images' features, Θ denotes the parameter set for all of samples subset $\boldsymbol{\alpha}$ in current distortion type and $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}, \Theta_c = \{\theta_c(1), \theta_c(2), \ldots, \theta_c(n)\}, \boldsymbol{\alpha}_c = \{\alpha_c(1), \alpha_c(2), \ldots, \alpha_c(n)\}$.

In this piecewise quality prediction model, the key task is to determine the appropriate samples subset partition, which is similar with the breakpoints estimation in [46,47]. An ideal samples

subset partition should satisfy two requirements: (1) completeness, i.e., the number of samples should be large enough for parameters regression and (2) homogeneity, i.e., the perception quality distribution in the feature space should be consistent and homogenous. Here, the clustering is an intuitive idea to implement the sample partition work. While, there are many difficulties to satisfy the previous partition requirements by directly using a clustering method. First, the clustering results are not robust for some representative methods like *k*-means [48] and *k*-medoid [49]. Second, it is hard to determine appropriate number of the clusters by hand or automatically [48]. Thirdly, the completeness is not a natural property in the clustering method, which may produce some incomplete subsets and reduce the accuracy of the learned model parameters.

Fortunately, many local learning algorithms provide us the clues of constituting different sample subsets. In [50–52], it is reported that the approaches of local learning a particular data based on its neighbors often outperform those methods using all training data. Inspired by these works, we develop a simple and efficient KNN based subset partition method. Fig. 6 shows an example of piecewise regression for the JPEG2000 compressed images, where all test images are labeled by *Test image-1* to *Test image-n* and their distortion types are identified as the JPEG2000. The candidate sample set is consist of all annotated JPEG2000 images whose DMOS are available. In the quality prediction process, we first search for the KNN of each test image from the candidate sample set. The selected KNN images could construct the training sample subsets. For example, the KNN of *Test image-1* construct its training subset $\alpha_c(1)$. Then, the unique model parameters $\theta_c(1)$ for *Test image-1* can be learned from $\alpha_c(1)$ by online training. The same process would be implemented for all the other test images. Since the model parameters are designed for each test image, it can adequately capture the local specificity of the perception quality distribution in the feature space.

In our experiment, the KNN of a test image are the K samples which produce the smallest chi-square distances [53] with respect to this test image in the candidate sample set. Let x_i denote the feature of the ith test image and $\hat{x}_j$ denote the feature of the jth candidate image. Then, the definition of chi-square distance $D(x_i, \hat{x}_j)$ can be given by

$$D(x_i, \hat{x}_j) = \prod_{m=1}^{3} d(x_i^m, \hat{x}_j^m)$$
$$d(x_i^m, \hat{x}_j^m) = \sum_{u=1}^{U_m} \frac{\left(x_i^m(u) - \hat{x}_j^m(u)\right)^2}{x_i^m(u) + \hat{x}_j^m(u)} \tag{7}$$

where x_i^m denotes the mth feature of x_i and $m = \{1, 2, 3\}$ correspond to the H, LBP and P features in Section 2, respectively. $x_i^m(u)$ denotes the uth element in the U_m dimension feature vector x_i^m.

Meanwhile, the support vector regression (SVR) [54] is employed to implement the online training in Fig. 6. In this way,

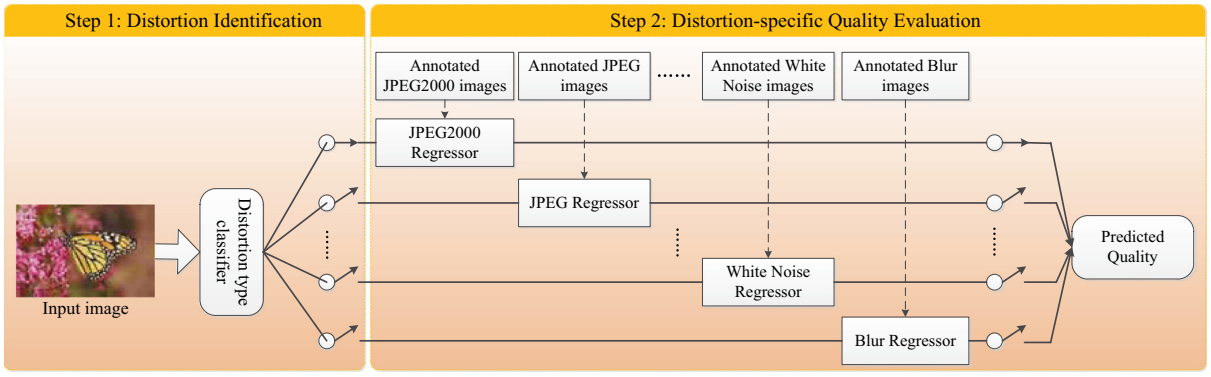

Fig. 5. The diagram of two-step NR-IQA scheme.

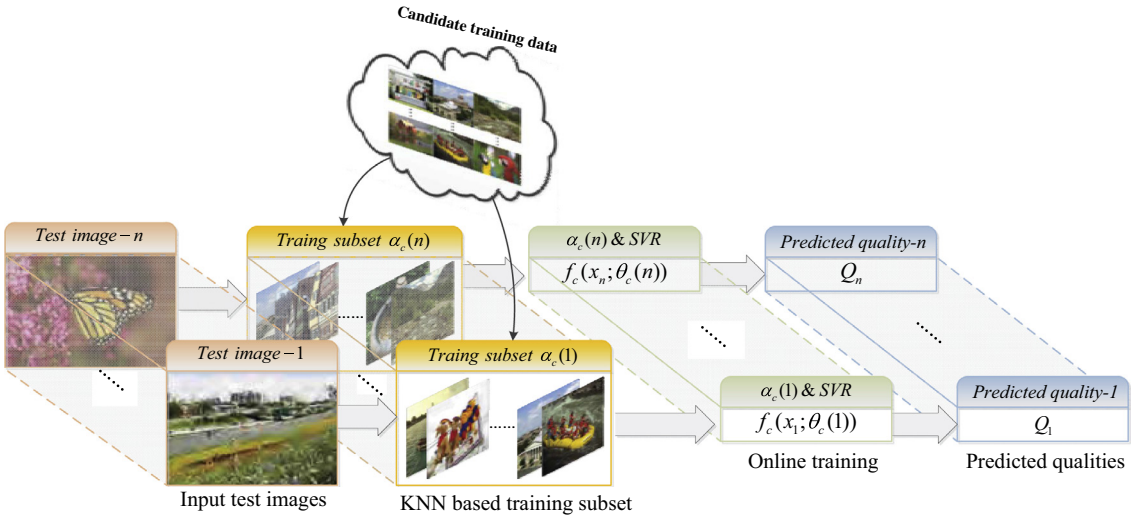

Fig. 6. The framework of the piecewise regression.

the SVR model parameters construct $\theta_c(i)$. As discussed in [54], the prediction function $f_c(x_i;\ \theta_c(i))$ in kernel methods can be represented as

$$f_c(x_i;\ \theta_c(i)) = \sum_{k=1}^{K} a_k \mathbb{K}(x_i, \hat{x}_k) + b \tag{8}$$

where a_k and b denote the SVR model parameters learned from the training sample subset and $\theta_c(i) = \{a_k, b | 1 \leqslant i \leqslant K\}$. $\hat{x}_k$ represents the kth sample in the training sample subset, $\mathbb{K}$ is the kernel function which is used to measure the similarity of two samples in the high dimension feature space. To capture different local distribution properties, both the polynomial kernel (P-K) and radial basis function kernel (RBF-K) are used here, i.e.,

$$\begin{aligned} \mathbb{K}_{P-K}(x_i, \hat{x}_k) &= (x_i^T \hat{x}_k + c)^d \\ \mathbb{K}_{RBF-K}(x_i, \hat{x}_k) &= exp\left(-\frac{\|x_i - \hat{x}_k\|_2^2}{2\sigma^2}\right) \end{aligned} \tag{9}$$

where c, d and σ are the custom parameters for the two kernels. The kernel selection and custom parameter setting are determined by cross validation.

4. Performance evaluation

4.1. Protocol

In our proposed method, the detailed experimental setting is as follows: (1) For the wavelet domain gradient orientation statistics, the spatial cell partition is fixed to 3×3 and the block normalization unit is set to 2×2 cells. In each cell, the gradient orientation is quantized to 9 bins; (2) For the LBP descriptor, the radius and neighbor parameters are set to 2 and 16, respectively; (3) For the global frequency distribution, the wavelet coefficients are quantized to 100 bins in each frequency band; and (4) The LIBSVM toolbox [55] is used for the distortion type classification and subjective quality regression.

In our method, there are two elements introduced to improve the performance of the BIQA, i.e., multi-domain structural information and piecewise regression. To evaluate the contribution of each part, we implement the proposed method with the following four combinations: global frequency distribution feature + single-phase regression (referred to as GF-SR); global frequency distribution feature + piecewise regression (referred to as GF-PR); multi-domain structural and global frequency features + single-phase regression (referred to as MSGF-SR); multi-domain structural and global frequency features + piecewise regression (referred to as MSGF-PR).

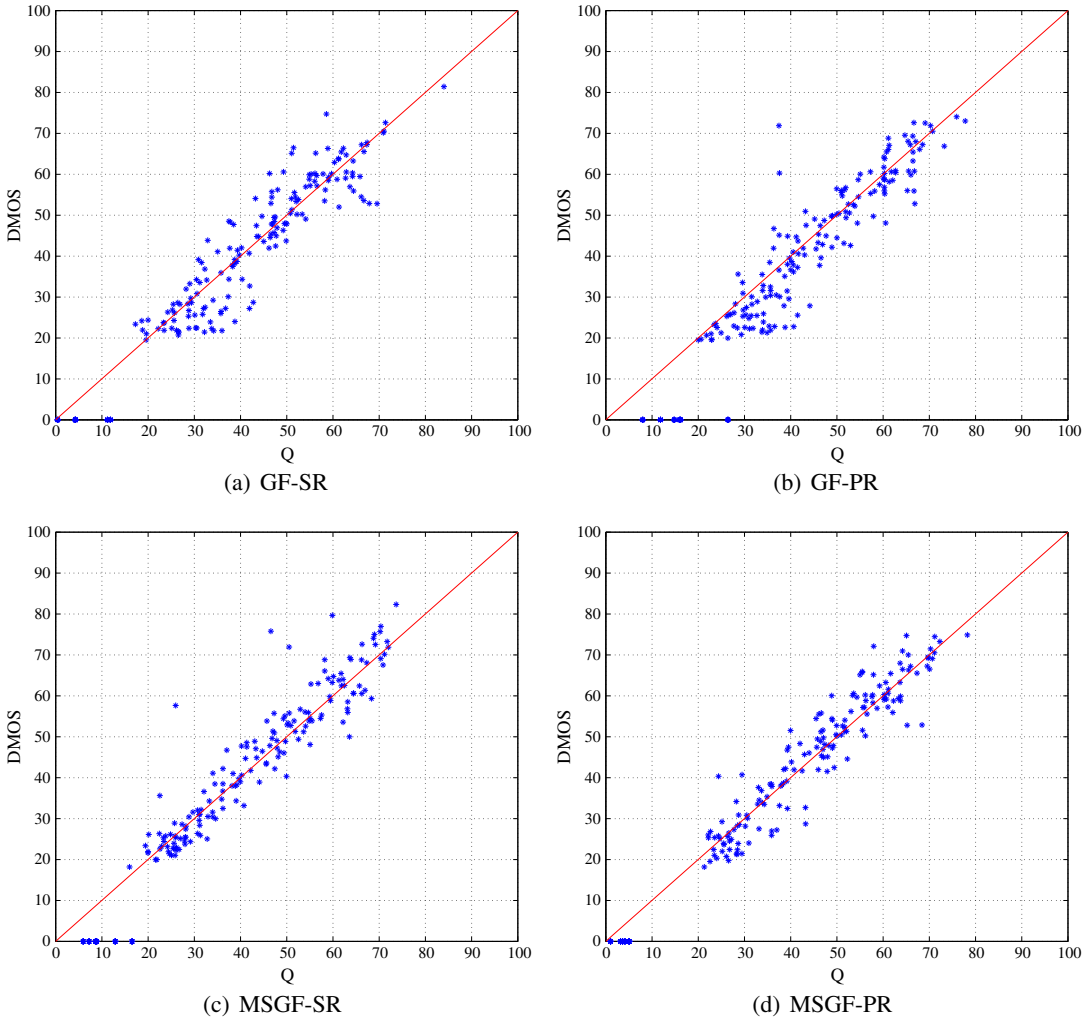

Fig. 7. The scatter plots of the predicted quality index Q vs. the DMOS.

Table 1
Median PLCC, SROCC, RMSE and MAE across 100 trials on the LIVE II IQA database.

Distortion		JPEG2000				JPEG			
Metric	Type	PLCC	SROCC	RMSE	MAE	PLCC	SROCC	RMSE	MAE
PSNR	FR	0.896	0.890	7.187	5.528	0.860	0.841	8.170	6.380
SSIM	FR	0.937	0.932	5.671	4.433	0.928	0.903	5.947	4.485
pLSA	NR	0.87	0.85	–	–	0.90	0.88	–	–
BIQI	NR	0.750	0.736	16.540	–	0.630	0.591	24.580	–
BLIINDS	NR	0.807	0.805	14.780	–	0.597	0.552	25.320	–
DIIVINE	NR	0.922	0.913	9.660	–	0.921	0.910	12.250	–
BLIINDS-II	NR	**0.963**	**0.951**	–	–	**0.979**	**0.942**	–	–
SF100	NR	0.929	0.924	–	–	0.940	0.928	–	–
NSS-TS	NR	0.947	0.931	5.792	7.169	0.933	0.915	6.333	7.912
GF-SR	NR	0.912	0.908	7.060	5.820	0.917	0.885	6.769	5.250
GF-PR	NR	0.909	0.902	7.184	5.915	0.933	0.911	6.010	4.281
MSGF-SR	NR	0.940	0.929	5.950	4.699	0.925	0.871	6.339	4.723
MSGF-PR	NR	0.942	0.935	5.796	4.603	0.939	0.902	5.758	4.154
		WN				Blur			
PSNR	FR	0.986	0.985	2.680	2.164	0.783	0.782	9.772	7.743
SSIM	FR	0.970	0.963	3.916	3.257	0.874	0.894	7.639	5.760
pLSA	NR	0.87	0.80	–	–	0.88	0.87	–	–
BIQI	NR	0.968	0.958	6.930	–	0.800	0.778	11.100	–
BLIINDS	NR	0.914	0.890	11.270	–	0.870	0.834	9.080	–
DIIVINE	NR	0.988	0.984	4.310	–	0.923	0.921	7.070	–
BLIINDS-II	NR	0.985	0.978	–	–	0.948	0.944	–	–
SF100	NR	0.978	0.962	–	–	**0.960**	**0.961**	–	–
NSS-TS	NR	0.963	0.971	4.464	6.018	0.950	0.939	5.481	6.863
GF-SR	NR	0.992	0.985	2.199	1.683	0.933	0.915	5.981	4.985
GF-PR	NR	0.992	0.985	2.169	1.716	0.935	0.912	5.954	5.041
MSGF-SR	NR	0.991	0.985	2.455	1.875	0.953	0.940	5.268	4.326
MSGF-PR	NR	**0.992**	**0.986**	2.213	1.749	0.956	0.946	4.902	4.002
		FF				Entire database			
PSNR	FR	0.890	0.890	7.516	5.800	0.824	0.820	9.124	7.325
SSIM	FR	0.943	0.941	5.485	4.297	0.863	0.851	8.126	6.275
pLSA	NR	0.84	0.77	–	–	0.79	0.80	–	–
BIQI	NR	0.722	0.700	19.480	–	0.740	0.726	18.360	–
BLIINDS	NR	0.743	0.678	18.620	–	0.680	0.663	20.010	–
DIIVINE	NR	0.888	0.863	12.930	–	0.917	0.916	10.900	–
BLIINDS-II	NR	**0.944**	0.927	–	–	0.923	0.920	–	–
SF100	NR	0.888	0.879	–	–	0.921	0.920	–	–
NSS-TS	NR	0.942	**0.935**	5.232	7.070	0.926	0.930	5.131	6.803
GF-SR	NR	0.817	0.824	10.147	7.197	0.907	0.914	7.075	5.106
GF-PR	NR	0.828	0.820	9.802	7.175	0.909	0.915	6.919	4.944
MSGF-SR	NR	0.926	0.903	6.477	4.981	0.938	0.938	5.718	4.211
MSGF-PR	NR	0.922	0.905	6.857	5.169	**0.944**	**0.942**	5.525	3.997

The performance of the proposed algorithms are validated on three largest publicly available databases, i.e., LIVE II [56], TID2008 [57] and CSIQ [58]. Here, the LIVE II IQA database consists of 29 pristine images and their 779 distorted versions under five distortion types, i.e., JPEG2000, JPEG, additive white noise (WN), Gaussian Blur and fast fading (FF). The TID2008 IQA database includes 25 reference images and their 1700 distorted versions under 17 distortion types with 4 levels. The CSIQ IQA database is composed of 30 reference images and their 900 distorted versions under 6 distortion types with 5 levels. Similar with many previous works [21,22,19,59,60], only 4 common distortion types (i.e., JPEG2000, JPEG, WN and Blur) in the TID2008 and CSIQ databases are involved in our comparison.

Following the same cross-validation criterion in [21,22,18,19], we randomly split the images of the LIVE II database into two non-overlapped sets in terms of the content, i.e., the training set and test set. The training set is composed of 80% reference images and their associated distorted versions. The rest 20% data is used for the test set. In addition, to determine an appropriate neighbor number for collecting the local training samples, we further divide the training set into the non-overlapped annotated set and validation set, which contain the 60% and 20% images in the LIVE II database. Here, the neighbor number can range from 5 to 100 at the interval of 5. For each distortion type, we search the optimal neighbor parameter by the cross validation, which achieves the best performance on the validation set. We conduct the random splitting evaluation 100 times on the LIVE II database. The median of the indexes across the 100 trials are reported here.

In this section, we employ four indices to measure the performance of different NR-IQA methods: the Pearson's linear correlation coefficient (PLCC) [61], the Spearman's rank ordered correlation coefficient (SROCC) [61], the root-mean-square error (RMSE) and mean absolute error (MAE) between the predicted quality Q and the ground truth DMOS.

4.2. Consistency experiment

In the consistency experiment, the LIVE II IQA database is used as the benchmark database. Here, we compare the proposed methods with some representative full-reference (i.e., PSNR, SSIM [62], VIF [63]) and no-reference (i.e., pLSA [64], BIQI [18], BLIINDS [21], BLIINDS-II [22], DIIVINE [19] and *SFk* [59]) IQA metrics.

Before giving the quantitative comparison results, the intuitive scatter plots of our proposed methods are first shown in Fig. 7. Here, the x-coordinate denotes the DMOS value and the y-coordinate denotes the predicted quality index Q. The red line represents the ideal linear correlation line "Q = DMOS". It can be seen that most samples compactly gather around the ideal linear corre-

lation line for each plot. This demonstrates that the predicted qualities of all our proposed methods are highly consistent with the human perception.

In the following, the detailed quantitative comparison on the five distortion types (i.e., JPEG2000, JPEG, WN, Blur and FF) and the entire database in shown in Table 1. For clarity, the best NR-IQA metrics in terms of PLCC and SROCC have been highlighted in boldface for each distortion type and the entire database. It can be seen that the proposed methods work well on the LIVE II database. For the WN, the MSGF-PR achieves the best performance whose PLCC and SROCC can be up to 0.992 and 0.986, respectively. For the Blur, the MSGF-PR achieves the suboptimal result in terms of the PLCC and SROCC. For the JPEG2000, the MSGF-PR achieves the second best performance in terms of the SROCC. For the JPEG and FF, the MSGF-PR performs as the third best in terms of PLCC and SROCC, respectively.

From the perspective of universal IQA, an efficient metric should work well across all distortion types. Based on this criterion, our MSGF-PR method outperforms all the other NR-IQA metrics, whose PLCC and SROCC reaches up to 0.944 and 0.942 in the entire database comparison. In addition, it is noted that the performances of our proposed four combinations are gradually improved by capturing local properties in the image representation and perceptual quality regression modules.

Firstly, under the same regression method, the perceptual correlation indexes raise by adding our MSGF feature. More particularly, the PLCC of MSGF-SR is 0.031 higher than the GF-SR, and the PLCC of MSGF-PR is 0.035 higher than the GF-PR. Secondly, under the same features, the perceptual correlation performance will also be improved by introducing the piecewise regression. For instance, the PLCC of GF-PR is 0.002 higher than GF-SR, and the PLCC of MSGF-PR is 0.009 higher than MSGF-SR. It proves that both our proposed MSGF features and the piecewise regression improve the performance of the NR-IQA metric.

To verify that the robustness of the proposed method, we further investigate the standard deviation of each measurement index across 100 trials as shown in Tables 2 and 3. It can be seen that the standard deviations for all of the proposed methods are very small, which indicates that the performance variation of our methods are very little across the 100 train-test trials.

Table 4
Median classification accuracy (%) across 100 train-test trials on the LIVE II IQA database.

	JPEG2000	JPEG	WN	Blur	FF	All
DIIVINE	80.00	81.10	100.00	90.00	73.30	83.75
GF	88.89	94.60	100.00	93.33	80.00	90.38
MSGF	86.11	100.00	100.00	96.67	90.00	94.12

Table 5
Standard derivation of the classification accuracy across 100 train-test trials on the LIVE II IQA database.

	JPEG2000	JPEG	WN	Blur	FF	All
GF	9.80	5.81	0.00	6.40	12.04	3.35
MSGF	10.00	2.31	0.00	2.46	7.91	2.71

4.3. Classification accuracy

Since we follow the two-step framework, the distortion type classification also plays an important role in our MSGF-PR method. To quantitatively evaluate the classification performance, we show the median classification accuracies of each individual distortion type and the entire test set across the 100 trials in Table 4. It can be seen that with the help of our MSGF features, the proposed MSGF-PR method achieves very high classification accuracy, which is better than the DIIVINE metric across all distortion types. Here, Table 5 shows the standard deviation of the our classification accuracies across the 100 train-test trials. It can be seen that the values in Table 5 are very small, which verifies that the results reported in Table 4 are robust and stable.

4.4. Database independence

To verify that the proposed method doesn't depend on some specific databases, we further conduct the data independency investigation in this section. Follow the protocol in [22,19], all the subject-rated images in the LIVE II database are used as the candidate training data for our piecewise regression. Both the TID2008 and CSIQ databases are used as the test sets.

Table 2
Standard derivation of the GF-SR/PR's performances across 100 train-test trials on the LIVE II IQA database.

	GF-SR				GF-PR			
	PLCC	SROCC	RMSE	MAE	PLCC	SROCC	RMSE	MAE
JPEG2000	0.027	0.028	1.055	0.893	0.028	0.028	0.987	0.853
JPEG	0.025	0.035	1.249	1.027	0.022	0.027	1.136	0.843
WN	0.002	0.005	0.321	0.285	0.002	0.006	0.285	0.239
Blur	0.026	0.037	1.089	0.884	0.027	0.046	1.038	0.849
FF	0.175	0.068	3.518	2.291	0.170	0.081	2.965	2.558
All	0.140	0.017	1.164	1.372	0.126	0.017	1.009	0.858

Table 3
Standard derivation of the MSGF-SR/PR's performances across 100 train-test trials on the LIVE II IQA database.

	MSGF-SR				MSGF-PR			
	PLCC	SROCC	RMSE	MAE	PLCC	SROCC	RMSE	MAE
JPEG2000	0.022	0.022	0.866	0.652	0.023	0.024	0.867	0.639
JPEG	0.022	0.034	1.155	0.779	0.019	0.034	1.140	0.829
WN	0.003	0.005	0.342	0.259	0.002	0.005	0.230	0.207
Blur	0.022	0.025	0.955	0.792	0.016	0.022	0.763	0.652
FF	0.036	0.039	1.382	0.918	0.038	0.048	1.540	1.011
All	0.009	0.009	0.533	0.358	0.011	0.010	0.588	0.403

Table 6
SROCC between the predicted quality index and MOS on the TID2008 database.

Metric	Type	JPEG2000	JPEG	WN	Blur	All
PSNR	FR	0.825	0.876	0.918	0.934	0.870
SSIM	FR	0.963	0.935	0.817	0.960	0.902
DIIVINE	NR	**0.924**	0.866	0.851	0.862	**0.889**
BLIINDS-II	NR	0.915	**0.889**	0.696	0.857	0.854
GF-SR	NR	0.776	0.781	0.883	0.835	0.808
GF-PR	NR	0.795	0.796	**0.911**	0.876	0.809
MSGF-SR	NR	0.842	0.800	0.885	0.884	0.845
MSGF-PR	NR	0.846	0.840	0.864	**0.893**	0.860

Table 7
SROCC between the predicted quality index and DMOS on the CSIQ database.

Metric	Type	JPEG2000	JPEG	WN	Blur	All
PSNR	FR	0.936	0.888	0.936	0.925	0.921
SSIM	FR	0.958	0.944	0.898	0.958	0.926
DIIVINE	NR	0.830	0.704	0.797	0.871	0.828
BLIINDS-II	NR	0.884	0.881	0.886	0.870	0.873
GF-SR	NR	0.840	0.891	0.940	0.859	0.868
GF-PR	NR	0.846	0.900	0.946	**0.889**	0.870
MSGF-SR	NR	**0.900**	0.915	0.926	0.831	0.867
MSGF-PR	NR	0.899	**0.923**	**0.952**	0.871	**0.881**

(a) QP=30,Q=38.36 (b) QP=36,Q=45.66

(c) QP=42,Q=50.19 (d) QP=48,Q=52.50

Fig. 8. The subjective results of the images compressed by HEVC. QP is the quantization parameter in HEVC, and Q is the predicted subjective score of the proposed method.

Tables 6 and 7 show the SROCC values between our predicted quality and the actual subjective score (i.e., MOS in TID2008 and DMOS in CSIQ) for the TID2008 and CSIQ databases, respectively. For comparison, the performances of two FR-IQA metrics (i.e., PSNR and SSIM) and two NR-IQA metrics (i.e., DIIVINE and BLIINDS-II) are also listed in Tables 6 and 7. For clarity, the highest SROCC results across all NR-IQA metrics have been highlighted in boldface.

It can be seen that all of the proposed methods perform well in the data independence experiment. In Tables 6 and 7, the MSGF-PR achieves comparable performance relative to the FR-IQA metrics. For the TID2008 database, the MSGF-PR is inferior to the DIIVINE

and superior to the BLIINDS-II. For the CSIQ database, our MSGF-PR method outperforms both of the NR-IQA metrics. In addition, the relative performances of the proposed methods still hold in both the TID2008 and CSIQ databases, i.e., GF-PR is superior to GF-SR, and MSGF-PR is better than MSGF-SR. It verifies that the superiorities of the MSGF features and the piecewise regression don't depend on specific datasets.

4.5. Investigation on HEVC still image profile

In this section, we further investigate the performance of the proposed method on the up-to-date HEVC compression distortion. Particularly, we select four standard color images, i.e., *Lena, Blonde Woman, Baboon* and *Avion* from USC-SIPI database[1] for testing. The resolutions of these test images are all 512×512. In our experiment, the HEVC reference software (HM13.0)[2] is used to encode each test image. Here, the intra main profile is used for coding and four distortion levels are generated under the setting of QP = {30, 36, 42, 48}. Similar with Section 4.4, we utilize all annotated images of LIVE II database [56] as the candidate training data for our piecewise regression.

For each test image, we compute the SROCC between the predicted scores of its four compressed versions and the QP for evaluation. Experimental results show that the SROCC of all test images are 1. That is, the proposed method could efficiently capture the monotonic decrease of perceptual quality when QP is increasing. For clarity, the four compressed versions of *Lena* and our predicted subjective scores are shown in Fig. 8. It can be seen that the predicted scores increase as the perceptual image qualities decrease from Fig. 8(a)–(d). Here, the predicted score corresponding to the DMOS value and a higher score indicates a wore subjective quality.

5. Discussion

Similar with most existing NR-IQA algorithms [18,21,19,22,59,20], the proposed method follows a supervised learning framework, which requires the human rated images to train a parametric quality prediction function. Although this framework could deliver high prediction accuracy, its flexibility is limited by the training procedure. That is, when some new training samples are present, the prediction function needs to be retrained. To overcome this limitation of flexibility, we could incorporate the proposed quality-aware features into an unsupervised method. Some possible solutions have been discussed in [64,29]. For example, we can compute our MSGF features for some undistorted natural images and the test images. Then, the mean feature distance between a test image and all natural images could be used as the predicted quality. Or, a natural fitting model (NFM) [29] could be first trained with a large number of natural images. Then, the MSGF feature similarity between the test image and the off-line trained NFM is used as the estimated quality. In our future work, some more flexible unsupervised method would be studied to extend our MSGF-PR method.

6. Conclusion

In this paper, we propose an efficient universal NR-IQA metric which tries to capture the local properties of the image representation and subjective quality regression. In the feature extraction module, both the spatial local binary pattern distribution and the frequency oriented gradient statistic are utilized to describe the image local structure, which is referred to as multi-domain structural information. In the perceptual quality prediction module, the piecewise regression scheme is used to simplify the complex global approximating task. Instead of training the single prediction model to fit all data's distribution, we learn a specific model for each test image, where the training set is composed of the local neighbors of the test image in the feature space. Since the distribution of the subjective scores is usually more homogeneous in the local region of the feature space, the piecewise regression could achieve more accurate approximating result. Through the extensive experiments on LIVE II, TID2008 and CSIQ databases, it is demonstrated that the proposed method is highly consistent with the human perception. Meanwhile, our proposed method also achieves comparable performance with respect to some FR-IQA metrics and outperforms many representative NR-IQA methods.

In our future work, we would try to develop more efficient local descriptor. In addition, a robust unsupervised grouping method for piecewise regression will also be studied.

[1] http://sipi.usc.edu/database/.

[2] https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-13.0/.

Acknowledgments

This work was supported in part by the National Basic Research Program of China (973 Program 2015CB351804), National Natural Science Foundation of China (No. 61271289), The program for Science and Technology Innovative Research Team for Young Scholars in Sichuan Province, China (No. 2014TD0006) and by The Fundamental Research Funds for the Central Universities (ZYGX2012YB007).

References

[1] Z. Wang, A.C. Bovik, Modern Image Quality Assessment, Morgan & Claypool, New York, 2006.
[2] W. Lin, C.-C.J. Kuo, Perceptual visual quality metrics: A survey, J. Vis. Commun. Image Represent. 22 (4) (2011) 297–312.
[3] Z. Wang, A. Bovik, Reduced- and no-reference image quality assessment, IEEE Signal Process. Mag. 28 (6) (2011) 29–40.
[4] S. Li, F. Zhang, L. Ma, K.N. Ngan, Image quality assessment by separately evaluating detail losses and additive impairments, IEEE Trans. Multimedia 13 (5) (2011) 935–949.
[5] A. Tanchenko, Visual-PSNR measure of image quality, J. Vis. Commun. Image Represent. 25 (5) (2014) 874–878.
[6] M. Narwaria, W. Lin, Objective image quality assessment based on support vector regression, IEEE Trans. Neural Networks 21 (3) (2010) 515–519.
[7] J. Wu, W. Lin, G. Shi, A. Liu, Perceptual quality metric with internal generative mechanism, IEEE Trans. Image Process. 22 (1) (2013) 43–54.
[8] L. Ma, S. Li, F. Zhang, K.N. Ngan, Reduced-reference image quality assessment using reorganized DCT-based image representation, IEEE Trans. Multimedia 13 (4) (2011) 824–829.
[9] M. Nauge, M. Larabi, C. Fernandez, A reduced-reference metric based on the interest points in color images, in: Picture Coding Symposium (PCS), 2010, pp. 610–613.
[10] R. Soundararajan, A. Bovik, RRED indices: reduced reference entropic differencing for image quality assessment, IEEE Trans. Image Process. 21 (2) (2012) 517–526.
[11] A. Rehman, Z. Wang, Reduced-reference image quality assessment by structural similarity estimation, IEEE Trans. Image Process. 21 (8) (2012) 3378–3389.
[12] X. Gao, W. Lu, D. Tao, X. Li, Image quality assessment based on multiscale geometric analysis, IEEE Trans. Image Process. 18 (7) (2009) 1409–1423.
[13] X. Kong, K. Li, Q. Yang, L. Wenyin, M.-H. Yang, A new image quality metric for image auto-denoising, in: IEEE International Conference on Computer Vision (ICCV), 2013, pp. 2888–2895.
[14] J. Yan, S. Lin, S.B. Kang, X. Tang, A learning-to-rank approach for image color enhancement, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2987–2994.
[15] Z. Chen, T. Jiang, Y. Tian, Quality assessment for comparing image enhancement algorithms, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 3003–3010.
[16] H.J. Seo, P. Milanfar, Visual saliency for automatic target detection, boundary detection, and image quality assessment, in: IEEE International Conference on Acoustics Speech and Signal Processing, 2010, pp. 5578–5581.
[17] B. Ghanem, E. Resendiz, N. Ahuja, Segmentation-based perceptual image quality assessment (SPIQA), in: IEEE International Conference on Image Processing, 2008, pp. 393–396.
[18] A. Moorthy, A. Bovik, A two-step framework for constructing blind image quality indices, IEEE Signal Process. Lett. 17 (5) (2010) 513–516.

[19] A.K. Moorthy, A.C. Bovik, Blind image quality assessment: from natural scene statistics to perceptual quality, IEEE Trans. Image Process. 20 (12) (2011) 3350–3364.
[20] X. Gao, F. Gao, D. Tao, X. Li, Universal blind image quality assessment metrics via natural scene statistics and multiple kernel learning, IEEE Trans. Neural Networks Learn. Syst. 24 (12) (2013) 2013–2026.
[21] M. Saad, A. Bovik, C. Charrier, A DCT statistics-based blind image quality index, IEEE Signal Process. Lett. 17 (6) (2010) 583–586.
[22] M. Saad, A.C. Bovik, C. Charrier, Blind image quality assessment: a natural scene statistics approach in the DCT domain, IEEE Trans. Image Process. 21 (8) (2012) 3339–3352.
[23] C. Tang, X. Yang, G. Zhai, Image quality/distortion metric based on α-stable model similarity in wavelet domain, J. Vis. Commun. Image Represent. 25 (7) (2014) 1746–1757.
[24] Q. Sang, H. Qi, X. Wu, C. Li, A.C. Bovik, No-reference image blur index based on singular value curve, J. Vis. Commun. Image Represent. 25 (7) (2014) 1625–1630.
[25] D.F. Specht, A general regression neural network, IEEE Trans. Neural Networks 2 (6) (1991) 568–576.
[26] C. Loader, Local Regression and Likelihood, vol. 47, Springer, New York, 1999.
[27] S. Marcelja, Mathematical description of the responses of simple cortical cells, J. Opt. Soc. Am. 70 (1980) 1297–1300.
[28] D.M. MacKay, Strife over visual cortical function, Nature 289 (5794) (1981) 117–118.
[29] A. Mittal, R. Soundararajan, A. Bovik, Making a "completely blind" image quality analyzer, IEEE Signal Process. Lett. 20 (3) (2013) 209–212.
[30] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24 (7) (2002) 971–987.
[31] G. Fink, J. Marshall, P. Halligan, R. Dolan, Hemispheric asymmetries in global/local processing are modulated by perceptual salience, Neuropsychologia 37 (1) (1998) 31–40.
[32] Q. Wu, H. Li, F. Meng, K.N. Ngan, B. Luo, C. Huang, B. Zeng, Blind image quality assessment based on multi-channel features fusion and label transfer, IEEE Trans. Circ. Syst. Video Technol. (2015), http://dx.doi.org/10.1109/TCSVT.2015.2412773. in press.
[33] G. Seber, C. Wild, Nonlinear rgression, in: Wiley Series in Probability and Statistics, Wiley, 2003.
[34] K.R. Rao, P. Yip, Discrete Cosine Transform: Algorithms, Advantages, Applications, Academic Press Professional, Inc., San Diego, CA, USA, 1990.
[35] S. Mallat, A Wavelet Tour of Signal Processing, 3rd ed., The Sparse Way, Academic Press, 2008.
[36] H. Li, K.N. Ngan, A co-saliency model of image pairs, IEEE Trans. Image Process. 20 (12) (2011) 3365–3375.
[37] J. Wu, F. Qi, G. Shi, Y. Lu, Non-local spatial redundancy reduction for bottom-up saliency estimation, J. Vis. Commun. Image Represent. 23 (7) (2012) 1158–1166.
[38] C. Wang, W. Liu, Z. Lai, H. Wang, Perceptually friendly shape decomposition by resolving segmentation points with minimum cost, J. Vis. Commun. Image Represent. 24 (3) (2013) 270–282.
[39] H. Li, F. Meng, B. Luo, S. Zhu, Repairing bad co-segmentation using its quality evaluation and segment propagation, IEEE Trans. Image Process. 23 (8) (2014) 3545–3559.
[40] H. Liu, M. Philipose, M.T. Sun, Automatic objects segmentation with RGB-D cameras, J. Vis. Commun. Image Represent. 25 (4) (2014) 709–718..
[41] H. Li, F. Meng, Q. Wu, B. Luo, Unsupervised multiclass region cosegmentation via ensemble clustering and energy minimization, IEEE Trans. Circ. Syst. Video Technol. 24 (5) (2014) 789–801.
[42] F. Meng, H. Li, K.N. Ngan, L. Zeng, Q. Wu, Feature adaptive co-segmentation by complexity awareness, IEEE Trans. Image Process. 22 (12) (2013) 4809–4824.
[43] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, 2005, pp. 886–893.
[44] K. Sharifi, A. Leon-Garcia, Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video, IEEE Trans. Circ. Syst. Video Technol. 5 (1) (1995) 52–56.
[45] M.J. Wainwright, E.P. Simoncelli, Scale mixtures of Gaussians and the statistics of natural images, in: S.A. Solla, T.K. Leen, K.-R. Müller (Eds.), Adv. Neural Information Processing Systems (NIPS'99), vol. 12, Cambridge, MA, 2000, pp. 855–861.
[46] J. Verbesselt, R. Hyndman, G. Newnham, D. Culvenor, Detecting trend and seasonal changes in satellite image time series, Remote Sens. Environ. 114 (1) (2010) 106–115.
[47] V.M.R. Muggeo, Estimating regression models with unknown break-points, Stat. Med. 22 (19) (2003) 3055–3071.
[48] A.K. Jain, Data clustering: 50 years beyond *k*-means, Pattern Recogn. Lett. 31 (8) (2010) 651–666.
[49] L. Kaufman, P.J. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis, vol. 344, Wiley Series in Probability and Statistics, 2009.
[50] L. Bottou, V. Vapnik, Local learning algorithms, Neural Comput. 4 (6) (1992) 888–900.
[51] M. Wu, B. Schölkopf, A local learning approach for clustering, in: Advances in Neural Information Processing Systems, 2006, pp. 1529–1536.
[52] M. Wu, B. Schölkopf, Transductive classification via local learning regularization, in: International Conference on Artificial Intelligence and Statistics, 2007, pp. 628–635.
[53] R.R. Sokal, F.J. Rohlf, Biometry, 3rd ed., W.H. Freeman, New York, 1995.
[54] B. Schölkopf, A.J. Smola, R.C. Williamson, P.L. Bartlett, New support vector algorithms, Neural Comput. 12 (5) (2000) 1207–1245.
[55] C.-C. Chang, C.-J. Lin, LIBSVM: A library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (3) (2011) 27:1–27:27.
[56] H.R. Sheikh, Z. Wang, L. Cormack, A.C. Bovik, LIVE Image Quality Assessment Database Release 2. <http://live.ece.utexas.edu/research/quality>.
[57] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, TID2008 – a database for evaluation of full-reference visual quality assessment metrics, Adv. Modern Radioelectron. 10 (2009) 30–45.
[58] E.C. Larson, D.M. Chandler, Categorical Image Quality (CSIQ) Database. <http://vision.okstate.edu/csiq>.
[59] P. Ye, J. Kumar, L. Kang, D. Doermann, Real-time no-reference image quality assessment based on filter learning, in: IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 987–994.
[60] W. Xue, L. Zhang, X. Mou, Learning without human scores for blind image quality assessment, in: IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 995–1002.
[61] M.B. Abdullah, On a robust correlation coefficient, The Statistician (1990) 455–460.
[62] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[63] H. Sheikh, A. Bovik, Image information and visual quality, IEEE Trans. Image Process. 15 (2) (2006) 430–444.
[64] A. Mittal, G.S. Muralidhar, J. Ghosh, A.C. Bovik, Blind image quality assessment without human training using latent quality factors, IEEE Signal Process. Lett. 19 (2) (2012) 75–78.