No-reference image quality assessment with visual pattern degradation☆

Jinjian Wu[a],[*], Man Zhang[a], Leida Li[a], Weisheng Dong[a], Guangming Shi[a], Weisi Lin[b]

[a] *State Key Laboratory of Integrated Services Networks, School of Artificial Intelligence, Xidian University, Xi'an 710071, China*
[b] *School of Computer Engineering, Nanyang Technological University, 639798, Singapore*

A R T I C L E   I N F O

A B S T R A C T

The technology of image quality assessment (IQA) is urgently demanded in perceptual-orientated image processing systems. Due to the lack of guidance from the reference information, it is challenging for the no-reference (NR) IQA to perform consistently with the human perception. It is well known that even though there exists no reference information, the human visual system can still effectively predict the image quality under the guidance of the prior knowledge. Thus, we try to build a visual content based prior knowledge database to guide the NR IQA. Cognitive researches state that the primary visual cortex presents substantially orientation selectivity, within which the structural information is extracted for visual perception. Inspired by this, the correlations among pixels in a local region are firstly analyzed for structure extraction. By mimicking the excitatory/inhibitory interactions among neurons, the relationships among pixels are estimated as their orientation similarities. Next, by aligning all of the correlations in a local region, a novel orientation similarity based pattern (OSP) is designed. Moreover, by merging the similar patterns with the clustering procedure, a set of fundamental OSPs are acquired. And then, the visual structure degradations are analyzed on these fundamental OSPs, and a prior knowledge database is learned. With the guidance of the knowledge database, a new visual pattern degradation (VPD) based NR-IQA model (NRVPD) is built. Finally, the performance of the proposed NRVPD model is verified on large benchmark IQA databases, and the experimental results demonstrate that the NRVPD performs highly consistent with the subjective perception.

© 2019 Elsevier Inc. All rights reserved.

1. Introduction

In the big-data multimedia era, we are facing explosive increase of visual signal. Thus, it becomes a new challenge to efficiently process/acquire the demanded visual signal. The technique of image quality assessment (IQA), which aims to automatically measure the quality of images, is useful to optimize the quality-oriented image/video processing systems [26,43], e.g., visual data acquisition, compression, and transmission.

During the past decades, there has been an increasing interest in developing objective IQA techniques, and a large amount of IQAs have been built [14,42,50]. According to the volume of the available reference image information, these existing models are classified into the following categories: (1) Full-Reference (FR) IQA, for which the full/whole information of the reference is demanded for quality prediction (such as the PSNR and SSIM [39], etc); (2) Reduced-Reference (RR) IQA, for which part reference information is used for degradation measurement (e.g., WNISM (wavelet-domain natural image statistic metric) [40] and RRED (Reduced Reference Entropic Differencing) [36]); (3) No-Reference (NR) IQA, which do not require any of the reference information during quality prediction (e.g., NFERM (NR free energy based robust metric) [12] and NSS-GS (natural scene statistics global scheme) [9], etc). With the help of the reference image, existing FR and RR IQAs can achieve high correlation with subjective quality assessment result. However, the reference image is not always available for most image/video processing systems. Therefore, the NR-IQA is greatly demanded in many intelligent image/video processing systems.

However, without the help from the reference image, it is extremely difficult for NR-IQA to perform consistently with the subjective perception. At the early stage, many distortion specific NR-IQA models had been designed for a certain type of distortion. In these models, the distortion type is used as a constraint to guide the quality prediction, for example, blockiness detection for JPEG noise [31], ringing artifact estimation for JPEG2000 coding [20], sharpness calculation for blur distortion [8], and so on. Because these models can only work for one specified distortion, their application domains are limited.

In the recent years, the non distortion specific NR IQA methods have attracted more and more attention. The contents of natural scenes follow some kind of statistical distribution (for example, the GGD (Generalized Gaussian Distribution)), and the distortions will destroy such distribution [40]. Following such assumption, the statistical property of wavelet coefficients from a distorted image is analyzed with GGD in [29], and a distortion identification based NR IQA method (i.e., DIIVINE) is proposed. By analyzing NSS features in the DCT domain, a new NR IQA model which uses DCT statistics (i.e., BLIINDS) method and its improved method (BLIINDS-II) are introduced in [34] and [33], respectively. Also, the NSS property is analyzed in the spatial domain and a referenceless IQA evaluator (i.e., BRISQUE) is proposed in [27]. Moreover, by integrating a set of NSS features, an integrated NR IQA evaluator (i.e., IL-NIQE) is proposed in [49]. Recently, a set of learning based non distortion specific NR IQA models are proposed, for example, a regression model is learned from image features (such as phase congruency [18], codebook [47], entropy [21], gradient magnitude [46], etc) for quality prediction. With the rapid development of deep learning (DL) technology, several DL-based NR IQA have been proposed [15]. For example, deep features are extracted with the convolution neural network for quality prediction in [19], and an end-to-end DL based neural network is built for NR IQA in [25]. However, the size limitation of the train dataset restrain the performance of these DL-based models. Therefore, there still exists a significant gap between these existing methods and the human perception for improvement.

Even without any help from the reference information, the human visual system (HVS) can efficiently and effectively predict the quality of an input scene. Thus, we turn to investigate the inner visual processing principle in the HVS for IQA modeling. Researches on cognition state that our human brain is proactive, which tries to link the input with an analogous representation in memory for prediction (i.e., linking the incoming features to existing representation in memory) [2]. During visual prediction, the regularity of the input scene will be analyzed with the help of the memory, and the visual patterns which represents the repeated structures are encoded [38]. In other words, the HVS is extremely adapted to extract patterns from the input scene for feature representation. Distortion will degrade visual patterns, which directly affect the accuracy of feature linking (with the analogous representations in memory). Therefore, we attempt to extract local visual patterns from images, and measure the pattern degradation for IQA.

In this work, we introduce an orientation similarity based pattern (OSP) for quality degradation measurement. The HVS presents substantial orientation selectivity (OS) in the local receptive field (especially in layer 4) for visual pattern extraction [13]. By mimicking the neuron interactions (i.e., excitation and inhibition [4]) among neurons for OS, correlations among neighbor pixels are analyzed with their preferred orientations. Next, according to the orientation similarity, the binarization correlations (similar/dissimilar) are acquired. Moreover, considering the arrangement of these binarization correlations in the local neighborhood, the OSP is built. Finally, by merging these similar patterns with the clustering procedure, a set of fundamental OSPs are acquired.

With the proposed OSP, a novel NR IQA method is proposed. Firstly, the visual information of an input scene is extracted with the OSP, and an image is mapped into an OSP based histogram for visual content representation. Next, the content degradation is analyzed according to changes on the OSP based histogram. By learning the degradation characteristic of different distortion types from IQA databases, a quality degradation based knowledge database is acquired. With the guidance of the knowledge database, the image quality is predicted under the support vector regression (SVR) procedure, and the novel NR IQA model with visual pattern degradation (NRVPD) is finally built.

The main contributions of this work are: 1) Inspired by the OS mechanism in the HVS, a new OSP is designed for visual information extraction. 2) A set of fundamental OSPs are selected for visual content representation. Experiments on different datasets demonstrate that these OSPs are universal and can effectively represent the common structures from a vast number of images. 3) A novel NRVPD model is introduced, which outperforms the-state-of-the-art NR-IQAs.
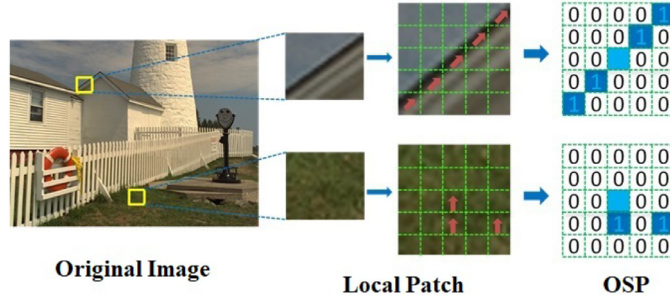
Fig. 1. An example of OSPs for image regions with individual visual contents, for which the above local region contains an obvious edge, and the bottom region contains disorderly texture.

2. Orientation similarity based pattern modeling

In this section, a brief introduction of the pattern extraction during the visual content perception is firstly given, especially for the OS in the primary visual cortex. Next, by imitating the OS mechanism, the novel OSP is designed.

Visual patterns, which represent these repeated information of images, play an important role in visual information perception. Though simply composed by qualitative gray-level differences of a local region, each pattern characterizes a certainty type of discernible regularity [11], and thus contains much high-level information of images [16]. Local visual patterns have been successfully used in a number of visual processing areas, e.g., the local binary pattern based texture classification [30], the scale invariant feature transform based keypoint detection [22], the local ternary patterns based object detection [37], the histogram of oriented gradients for face recognition [6], etc.

In this paper, a new set of visual patterns are designed for IQA task. Obviously, IQA is different from other visual processing tasks. For quality prediction, a type of visual pattern which should not only extract the visual content of the input scene, but also can capture the quality degradations from noise. The HVS is extremely adaptive to capture such type of pattern during quality prediction. Moreover, cognitive researches indicate that there exists substantial OS mechanism in the HVS for pattern extraction [13,41]. Inspired by this mechanism, we attempt to build a novel orientation similarity based visual pattern.

The OS arises from the arrangement of interactions among nearby cortical cells [4]. And there exits two different types of interactions, i.e., excitation vs. inhibition, which directly determine the preferred orientations of these cells: excitation is more likely to appear at cortical cells which have similar preferred orientations, whereas inhibition appears at those with dissimilar preferred orientations [7,48]. By imitating the interactions among a local receptive field, the correlations of pixels within a neighborhood ($\mathcal{R}$) are analyzed. The preferred orientation of pixel i is calculated as the gradient direction (θ) [41],

$$\theta(i) = \arctan \frac{G_y(i)}{G_x(i)}, \tag{1}$$

where $G_y(i)$ and $G_x(i)$ are the luminance changes alone the vertical and horizontal directions, respectively.

Since the type of interaction is related to the orientation similarity between cortical cells, we estimate the correlation between a pixel i and its neighbor $i+\Delta$ ($\Delta \in \mathcal{R}$) as

$$\mathcal{S}(i, i+\Delta) = \begin{cases} 1, & |\theta(i) - \theta(i+\Delta)| < T \\ 0, & \text{else} \end{cases}, \tag{2}$$

where $\mathcal{S}(i, i+\Delta)$ represents the correlation between i and $i+\Delta$, '1'/'0' means the two pixels have similar/dissimilar orientations, and T is the judging threshold. According to the subjective viewing test on the masking effect, the threshold T is set as $T=6°$ [3]. With Eq. (2), the correlation between two neighbor pixels is binarized (similar/dissimilar).

Inspired by the mechanism that the OS arises at a local receptive field, we design the novel pattern according to correlations among pixel in a local neighborhood. By considering the correlation between a pixel and its neighbors, an OSP is described as,

$$\mathcal{P}(i) = \{\mathcal{S}(i, j) | j=i+\Delta, \Delta \in \mathcal{R}\}, \tag{3}$$

where $\mathcal{P}(i)$ represents the OSP form of pixel i, and $\mathcal{R}$ represents the neighbor of pixel i. With Eq. (3), an OSP with binary orientation similarity is acquired. As shown in Fig. 1, different regions contain individual visual contents and present different OSPs. Moreover, the number of OSPs is exponentially increased with the size of $\mathcal{R}$. Thus, for a $r \times r$ local region, there exists $2^{(r \times r - 1)}$ different patterns. Considering the computational complexity, r is set as 5 in this work and there exists 2^{24} original patterns.
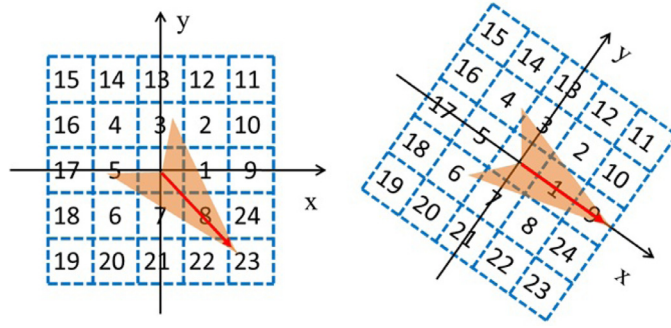
Fig. 2. An example of OSP realignment with its main orientation. The left one is the original OSP and the right one is the realigned OSP.
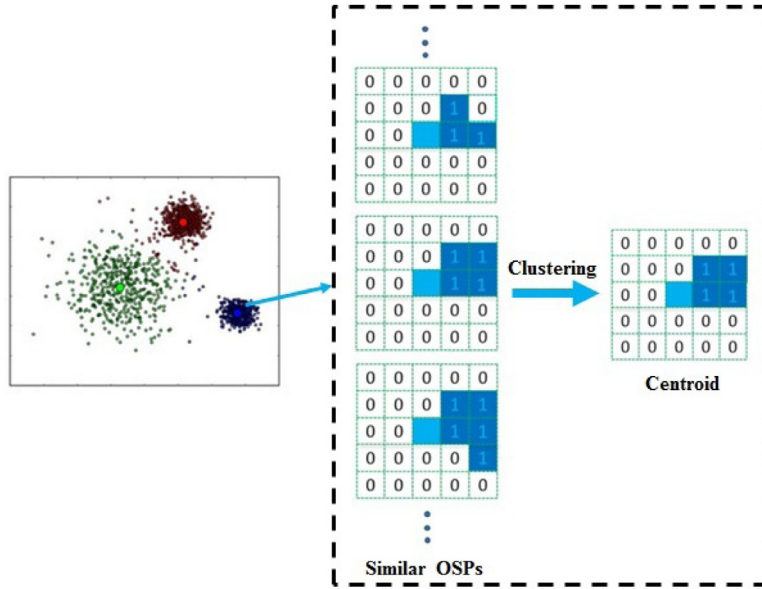

Fig. 3. An example of similar OSPs combination with the K-Means clustering algorithm.

3. Image quality assessment with OSP

In this section, by considering the rotation invariance and similarity of patterns, the fundamental patterns are firstly selected. With those fundamental patterns, the image content is extracted and represented by a pattern based histogram. Finally, the image quality is predicted with the degradations on the histogram with a learning procedure.

3.1. Fundamental OSPs selection

With Eq. (3), a large number of original OSPs is acquired (for a 5×5 local neighborhood, there are more than 16 millions different types of OSP). However, the significances of these OSPs are not equal for visual content representation (e.g., many types of patterns are infrequent). Thus, we try to reduce the pattern number and only reserve these fundamental ones. The rotation invariant characteristic is firstly considered for pattern reduction. In general, local regions with the same visual content while different rotated orientations can be captured by a same pattern. In order to merge such kinds of OSPs, the arrangement of correlations is realigned according to the main orientation (i.e., the preferred orientation of the central pixel), as shown in Fig. 2. With this step, the number of OSPs is reduced to $M=2,089,456$ (about 1/8 of the original size).

Besides, these OSPs which have similar shape present homogeneous visual contents, and they can also be merged for pattern reduction. As shown in Fig. 3, for these OSPs which have quite similar format, we can merge them and only keep one to represent such kind. In order to efficiently merge these OSPs with similar format, we use the K-Means procedure for clustering,

$$\{\hat{\mathcal{P}}_n, n = 1, \cdots, N\} = \arg \min \sum_{n=1}^{N} \sum_{m=1}^{M} ||w_m \cdot (\mathcal{P}_m - \hat{\mathcal{P}}_n)||^2, \tag{4}$$
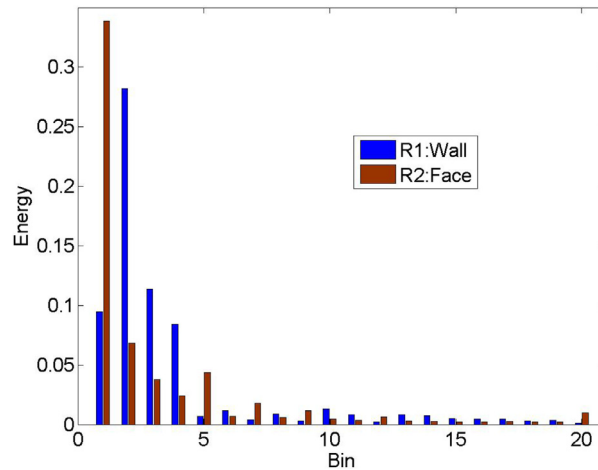
(a) R1: Wall   (b) R2: Face

Fig. 4. Two reference images for OSP based visual feature extraction, which contain extremely different visual contents (Wall vs. Face) and present obviously different OSP based histograms (as shown in Fig. 5).

Fig. 5. OSP based visual feature extraction on the two images (*Wall* and *Face* images) as shown in Fig. 4. Images with different visual contents present obviously different distributions on OSPs. In order to give a clear view of visual feature extraction, only the top 20 features are shown.

where $\hat{\mathcal{P}}_n$ is the nth centroid, N is the number of pixels, M is the centroid number, and w_m is the weighting factor (determined by the occurrence probability of $\mathcal{P}_m$).

In order to find out these fundamental OSPs ($\{\hat{\mathcal{P}}_n, n = 1, 2, \cdots, N\}$) which can efficiently represent the colorful images, a large dataset with natural images are required for K-Means clustering. In this work, a benchmark image dataset created by Achanta et al. [1] is chosen (which is designed for saliency detection and has no relationship with IQA). This dataset contains 1000 natural images with different visual objects, and is widely used for visual processing tasks. Moreover, images in the dataset are randomly divided into 2 classes (each with 500 images), and they are called as DS1 and DS2 here. The K-Means clustering is applied to DS1 according to the following steps:

1. The rotation invariant OSP values are firstly computed for pixels in an image.
2. Weight setting: all image pixels from DS1 are classified into the M types of OSPs, and the weight w_m for each type is set as the number of pixels belong to it.
3. Re-ranking the order of these OSPs based on to their weights.
4. Initial centroids setting: the first K OSP types are selected as the initial centroids.
5. Running the clustering procedure with Eq. (4)

Finally, K different centroids are acquired, which are set as the fundamental OSPs ($\{\hat{\mathcal{P}}_k, k = 1, 2, \cdots, K\}$) in this paper.

To demonstrate the universality of the learned K fundamental OSPs, the same procedure as that on DS1 is implemented on DS2. Then, the fundamental OSPs learned from the two databases are compared (we set the feature size K=100 here, and the effect of K on the performance is throughly analyzed in Section 4-C), and the comparison result is shown in the second

Table 1
Comparison among these basical OSPs from different image sets.

Datasets	Overlap proportion	Overlap number
DS1 VS. DS2	98.83%	94
DS1 VS. LIVE	96.51%	86
DS1 VS. TID	96.83%	90

and third rows of Table 1. Though the two datasets contains different images, their clustered centroids are extremely similar, and there exists 94 overlap centroids (with the same OSP types). Moreover, by counting all pixels from the two datasets on the 100 centroids, pixels belong to the overlap 94 centroids takes up 98.83%. Therefore, the fundamental OSPs learned from DS1 is universal for the basic visual content extraction from the colorful images.

In order to further illustrate the universality of the *K* fundamental OSPs learned from DS1, two IQA benchmark databases are chosen for verification, namely, LIVE [35] and TID2013 [32]. Since DS1 has no relationship with IQA, there is no similar image between it and LIVE/TID2013. However, their clustered centroids are highly related. As listed listed in Tab. 1, there are 86 overlap centroids (a total of 100 centroids) between DS1 and LIVE (90 between DS1 and TID2013). Moreover, more than 96% pixels belongs to the overlap centroids. According to the analysis above, these fundamental OSPs can extract the common structures from a vast number of images for visual information representation. And in this work, these fundamental OSPs learned from DS1 are used in all of the following experiments.

3.2. The OSP based IQA

Image quality degradation is caused by the distortions on visual contents. With the proposed OSPs ($\hat{\mathcal{P}}_k$), the spatial organizational structure of the visual content can be efficiently represented. Moreover, the HVS is extremely sensitive to changes on the luminance (rather than the luminance itself). Thus, luminance contrast is another important factor for visual content representation. The gradient magnitude is adopted to calculate the luminance contrast $\mathcal{C}_l$,

$$\mathcal{C}_l(i) = \sqrt{G_y(i)^2 + G_x(i)^2}, \tag{5}$$

where $G_y(i)$ and $G_x(i)$ are the luminance changes alone the vertical and horizontal directions, respectively.

By considering both the visual pattern and the luminance contrast, the visual information of each pixel is extracted. And then, the visual content that an image presented is mapped into a histogram (according to the values of visual pattern and luminance contrast) $\mathcal{H}$,

$$\mathcal{H}(y) = \sum_{x=1}^{N} \mathcal{C}_l(x) \cdot \delta(\hat{\mathcal{P}}(x), \hat{\mathcal{P}}_y) \tag{6}$$

$$\delta(\hat{\mathcal{P}}(x), \hat{\mathcal{P}}_y)) = \begin{cases} 1, & \text{if } \hat{\mathcal{P}}(x) = \hat{\mathcal{P}}_y \\ 0, & \text{else} \end{cases}. \tag{7}$$

An example of OSP based image information extraction (i.e., with Eq. 6) is shown in Figs. 4 and 5, where Fig. 4(a) and (b) are two reference images, and Fig. 5 is their corresponding histograms. As can be seen, Fig. 4(a) is with unsmooth brick wall, while Fig. 4(b) contains a large smooth area. Meanwhile, the two images represent distinguished histograms, as the two types of bars shown in Fig. 5. Therefore, the proposed OSP based histogram can efficiently represent images with different structures/contents.

Besides, different distortion types will cause individual quality degradation on the OSP based histograms (we will demonstrate this in the experimental section). Thus, the quality is measured as the degradation on the OSP based histogram. By taking both the OSP based histograms $\{\mathcal{H}_m | m = 1, \cdots, M\}$ and the subjective qualities $\{\mathcal{Q}_m | m = 1, \cdots, M\}$ of images from these existing IQA databases, a prior knowledge database $\mathcal{B}_n$ about the visual structural degradation can be learned,

$$\mathcal{B}_n = \text{Learn}(\mathcal{H}, \mathcal{Q}), \tag{8}$$

where Learn($\cdot$) represents a learning approach.

Then, with the guidance of the knowledge database $\mathcal{B}_n$, the perceptual quality of an image is predicted,

$$\hat{\mathcal{Q}}(I) = \text{Predict}(\mathcal{H}(I), \mathcal{B}_n), \tag{9}$$

where $\hat{\mathcal{Q}}$ is the predicted quality, and Predict($\cdot$) represents a predicting approach. As an efficient mapping procedure (regression module) from the high dimensional feature data (e.g., histogram) to the target score, the support vector regression (SVR) procedure is often adopted for quality prediction [12,27]. Here, we adopt the LIBSVM package [5] for IQA, and the learning and predicting procedures are implemented with the SVR under a radial basis function kernel ($K(\mathcal{H}_i, \mathcal{H}_j) = \exp(-\gamma \|\mathcal{H}_i - \mathcal{H}_j\|)$). For LIBSVR, there are three key parameters during optimization, i.e., the precision parameter γ, the balance parameter C, and the tolerance of termination criterion ϵ. Here, we search for the optimal γ and C by the grid search method, and select ϵ by ten-fold cross validation.

4. Experimental result demonstration

In this section, we firstly verify the efficiency of the OSPs for visual content extraction on different images with individual distortion types. Next, a set of state-of-the-art NR IQAs are chosen, and a comprehensive comparison against the proposed model on several publicly available databases is given.

4.1. Visual content extraction

The proposed OSPs can efficiently represent quality degradation caused by different distortion types. An intuitive example of OSP based quality degradation is given. As shown in Fig. 6(a)–(c) are three distorted images contaminated by white Gaussian noise (WGN), Gaussian blur (GBN), and JPEG2000 (J2K) compression, respectively. The number of features (the number of centroids) is set as 800 (which will be further analyzed in Section 4-C). For a clear view, we only draw the first 20 features (which takes up the majority) in Fig. 6. Moreover, Fig. 4(a) shows the corresponding reference image, and Fig. 6(d) shows the OSP based histograms of these distorted images.

Though under a similar level of distortion (the PSNR values for Fig. 6(a)–(c) are 24.5 dB, 21.0 dB, and 21.1 dB, respectively), different distortion types cause individual changes on OSP features. As can be seen, the disturbance of Fig. 6(a) is obviously increased with WGN. As a result, the energies of most regions in this image are increased (shown as the yellow bars in Fig. 6(d)). However, the energy for the first bar is greatly decreased (almost equal to zero). With further analysis, we have found that the first OSP represents the smooth pattern (i.e., bin 1 as shown in Fig. 7), and such a pattern mainly appears at the plain region. With the disturbance from the WGN, the smooth pattern is changed into other types of patterns (e.g., disorderly textural patterns as bins 2–4 shown in Fig. 7). While for the GBN, it smooths the luminance change of the whole image, and thus the energies of most regions are greatly decreased. As a result, the GBN causes almost opposite changes on the OSP features against the WGN. As shown in the light blue bars, the energy for the first bin is increased, while that for the other bins are obviously decreased (especially for the second, third and forth bins). For the J2K, many structures are removed due to the quantization during compression. As a result, the energies for bars 2–4 (corresponding to these textural patterns) are obviously decreased (shown as the red bars shown in Fig. 6(d)). At the meantime, the ringing effect from the J2K compression creates some new structures, which increases the energies of part bars (such as the 5–8 bins). In summary, the OSP based histogram can efficiently capture the individual quality degradations caused by different types of distortion.

Besides, the changes of OSP features under the same distortion type but different levels are analyzed. As an example, Fig. 4(a) is distorted by the GNB under three different levels, i.e., Fig. 8(a) with PSNR=34.0 dB, Fig. 8(b) with 25.3 dB, and Fig. 6(b) with 21.0 dB. The changes on the OSP features are shown in Fig. 9. By comparing these bars, we can see that with the increase of distortion level, the energy of the first bin (also for 5–8 bins) is gradually increased, while the energy of the second bin (also for 3–4 bins) is gradually decreased. Therefore, the OSP based histogram can also efficiently represent the degradation from a same distortion type but different levels.

Moreover, each type of distortion causes some identified changes on the OSP features. An example of JPEG compression noise (JPG for short) is given in Figs. 10 and 11, where Fig. 10(a) and (b) are two contaminated images, and Fig. 11 shows the OSP based histograms for them. Since the two images possess individual structures/contents, their OSP based histograms are obviously different. However, the degradation from the JPG causes some common changes on the OSP features of these two contaminated images. Shown as Fig. 10(a) and (b), there is obvious blockiness which greatly degrades their qualities. At the meantime, the energies of the 6, 8 and 18 bins are obviously increased for both histograms.

With the through analysis above, it can be concluded that our OSPs can efficiently represent quality degradations from different distortions and different levels.

4.2. Performance comparison

In this subsection, the effectiveness of our NRVPD method is verified. And a comprehensive comparison with the existing NR IQAs on several public benchmark databases is given.

4.2.1. Database and protocol

In this work, four large-scale benchmark databases in IQA society are chosen, namely, LIVE [35], CSIQ [17], TID2013 [32], and Wild-LIVE [10], in which the quality of each image has predicted by observers and marked with the mean opinion score (i.e., MOS, or differential MOS (DMOS)). Detailed information of the four benchmark databases is listed in Table 2

For a comprehensive analysis, nine state-of-the-art NR IQAs (which are IMNSS [44], HOSA [45], CNN-SVR [19], DL-IQA [15], IL-NIQE [49], NIQE [28], BRISQUE [27], BLIINDS-II [33], and DIIVINE [29]) are chosen for comparison.

To evaluate the prediction accuracy of these IQA methods, three metrics are employed: the Spearman rank order correlation coefficient (SRCC) between the predicted quality and the subjective quality (i.e., MOS/DMOS), which measures the prediction monotonicity of each method; The Pearson linear correlation coefficient (PLCC) and the root mean squared error (RMSE). All of them represent the accuracy of prediction. A good IQA model returns a high SRCC/PLCC value and a low RMSE value. Before calculating the three values, a nonlinear regression function (which is introduced by the Video Quality Expert
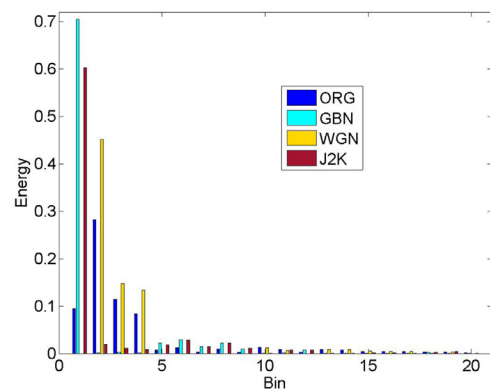
(a) WGN with PSNR=24.5

(b) GBN with PSNR=21.0

(c) J2K with PSNR=21.1

(d) OSP based histogram

Fig. 6. OSP based visual feature degradation under different distortion types. (a)–(c) are three distorted images under WGN, GBN, and J2K, respectively. (d) is their corresponding OSP based histograms. Though under a similar level of distortion, different distortion types cause individual changes on OSP features.

Table 2
The four benchmark databases for IQA.

DB	Ref. no.	Dis. type	Dis. level	Dis. no.
LIVE	29	5	–	779
CSIQ	30	6	5	866
TID2013	25	24	5	3000
Wild-LIVE	0	–	–	1162

Group) is used for score mapping,

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

where Q_r is the regressed quality value, Q is the predicted quality score, and $\{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5\}$ are the fitted regressive parameters.

4.2.2. Performance on individual databases

We firstly verify the proposed NRVPD method on LIVE database. Since the SVR is used to predict the quality score, a training and testing procedure is demanded. Thus, 80% reference images are randomly selected, and then their distorted images are used to training the SVR model. Moreover, The overall training and testing procedure is repeated 1000 times for performance bias eliminating. Finally, the median performance is adopted.

The performances of individual distortion types on the LIVE database are listed in Table 3, in which the best two results of the three metrics (i.e., PLCC, SRCC, and RMSE) are highlighted (with bold font). As shown in Table 3, the NRVPD performs
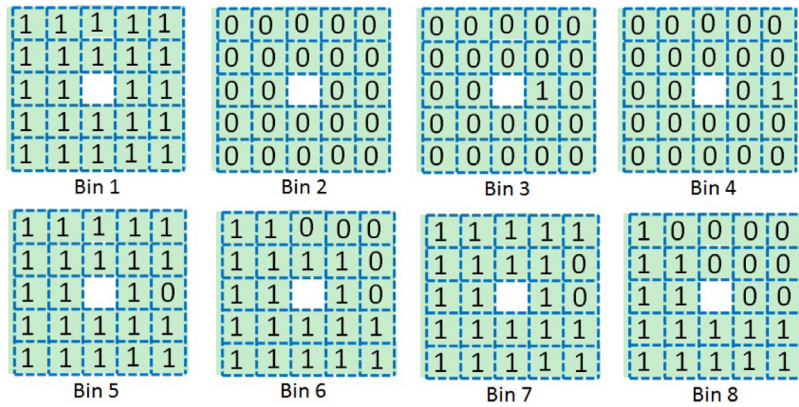
Fig. 7. The top eight representative OSPs, which represent different visual contents.

(a) GBN with PSNR=34.0   (b) GBN with PSNR=25.3

Fig. 8. Visual degradation (contaminated by the GBN) under different levels.

Fig. 9. OSP feature degradation under the same distortion type but different levels, where ORG refers to Fig. 4 (a), GBN1 refers to Fig. 8(a), GBN2 refers to Fig. 8(b), and GBN3 refers to Fig. 6(b). With the increases of distortion level, the energies of these OSP bins present disciplinary changes.

(a) D1   (b) D2

Fig. 10. Visual degradation (contaminated by the JPEG compression) on two different images.

Fig. 11. OSP feature degradation on two different images (as shown in Fig. 10) while under the same distortion (JPG), which present some identified changes on the OSP features.
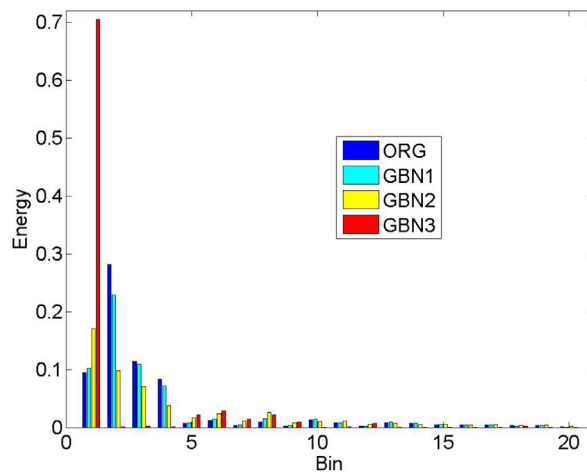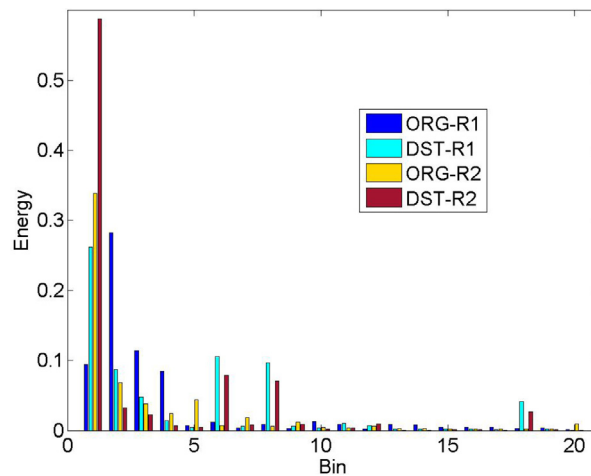
Table 3
Performances comparison on each type of distortion on the LIVE database, and the best two methods are emphasized with bold.

Distortion	Crit.	NRVPD	IMNSS	HOSA	CNN-SVR	DL-IQA	IL-NIQE	NIQE	BRISQUE	BLIINDS-II	DIIVINE
	PLCC	**0.955**	0.950	0.943	**0.955**	0.947	0.915	0.950	0.951	0.943	0.847
J2K	SRCC	**0.941**	**0.934**	0.928	0.931	0.928	0.895	0.935	0.930	0.927	0.822
	RMSE	**7.356**	**7.580**	8.505	–	–	10.35	8.015	7.902	8.524	14.09
JPG	PLCC	**0.971**	0.951	0.970	0.968	0.940	0.942	0.955	**0.971**	0.937	0.907
	SRCC	0.953	0.933	**0.960**	0.953	0.912	0.942	0.945	**0.965**	0.923	0.900
	RMSE	**7.565**	7.877	**7.494**	–	–	10.17	9.402	7.648	11.20	12.83
	PLCC	0.975	0.982	**0.984**	**0.985**	0.955	0.906	0.882	0.983	0.961	0.963
WGN	SRCC	0.968	**0.986**	0.972	**0.983**	0.968	0.979	0.977	0.980	0.950	0.965
	RMSE	6.242	**4.419**	**4.862**	–	–	11.84	12.76	5.098	7.731	7.346
	PLCC	0.938	0.948	**0.977**	0.963	0.944	0.922	0.958	**0.965**	0.930	0.888
GBN	SRCC	0.910	0.949	**0.966**	**0.967**	0.946	0.908	0.949	0.961	0.911	0.831
	RMSE	6.600	6.943	**4.042**	–	–	7.151	5.396	**4.854**	6.698	8.165
	PLCC	0.933	0.922	0.895	**0.952**	0.890	0.865	0.886	0.911	0.929	0.805
FFN	SRCC	**0.917**	0.895	0.874	**0.979**	0.861	0.827	0.844	0.886	0.911	0.763
	RMSE	**10.42**	10.560	12.18	–	–	14.35	13.12	11.33	**9.886**	14.83

Table 4

Overall NR-IQA performance comparison on four databases (i.e., LIVE, CSIQ, TID2013, and Wild-LIVE), and the best two method are emphasized with bold.

Distortion	Crit.	NRVPD	IMNSS	HOSA	CNN-SVR	DL-IQA	IL-NIQE	NIQE	BRISQUE	BLIINDS-II	DIIVINE
	PLCC	0.947	0.943	**0.949**	**0.974**	0.930	0.865	0.911	0.946	0.892	0.881
LIVE	SRCC	0.943	0.944	**0.948**	**0.964**	0.927	0.902	0.914	0.941	0.906	0.879
	RMSE	**8.636**	8.705	8.706	–	–	13.50	11.31	**8.671**	11.77	13.18
	PLCC	**0.889**	0.835	**0.836**	–	–	0.816	0.715	0.831	0.782	0.800
CSIQ	SRCC	**0.840**	0.789	0.776	–	–	**0.807**	0.629	0.768	0.709	0.769
	RMSE	**0.119**	**0.142**	0.144	–	–	0.147	0.181	0.143	0.163	0.149
	PLCC	**0.768**	0.598	**0.764**	–	–	0.640	0.415	0.626	0.601	0.672
TID2013	SRCC	**0.683**	0.522	**0.688**	–	–	0.519	0.298	0.520	0.467	0.583
	RMSE	0.790	0.997	**0.805**	–	–	0.954	1.137	0.956	0.996	0.924
Wild-LIVE	PLCC	**0.775**	0.530	**0.675**	–	–	0.505	0.487	0.647	0.415	0.596
	SRCC	**0.759**	0.520	**0.661**	–	–	0.425	0.445	0.611	0.388	0.576
W-Mean	PLCC	**0.812**	0.667	**0.782**	–	–	0.669	0.541	0.704	0.630	0.704
	SRCC	**0.757**	0.619	**0.731**	–	–	0.594	0.459	0.632	0.546	0.649

Table 5

F-test based performance comparison.

DB	IMNSS	HOSA	CNN-SVR	DL-IQA	IL-NIQE	NIQE	BRISQUE	BLIINDS-II	DIIVINE
LIVE	0	0	−1	1	1	1	1	1	1
CSIQ	1	1	–	–	1	1	1	1	1
TID2013	1	0	–	–	1	1	1	1	1
Wild-LIVE	1	1	–	–	1	1	1	1	1

highly consistent with the human perception, whose PLCC and SRCC values are always larger than 0.91 on the five distortion types. Moreover, by comparing with the other NR-IQA methods, the NRVPD outperforms DL-IQA, IL-NIQE, and DIVINE on all of these distortions (has larger PLCC/SRCC values and smaller RMSE value than the others). Meanwhile, the proposed NRVPD performs the best on J2K (with the highest PLCC (0.955) and SRCC (0.941) values, as well as the lowest RMSE (7.356) value) and JPG (with the largest PLCC (0.971) value), performs comparably with the best method on WGN (IL-NIQW) and FFN (CNN-SVR), while a slightly worse than HOSA (who performs the best) on GBN.

Meanwhile, the overall performance comparison against nine latest NR-IQA methods on the four databases are also given. As shown in Table 4, the NRVPD performs better than the other NR-IQA methods on three databases, i.e., CSIQ, TID2013, and Wild-LIVE. Especially for the challenging database (Wild-LIVE contains widely diverse authentic image distortions), the proposed NRVPD performs much better (a remarkable improvement) than these existing methods (NRVPD vs. the second best one on Wild-LIVE: 0.775 VS. 0.675). On the LIVE, the NRVPD performs slightly worse than CNN-SVR and HOSA, while better than the other eight methods. Meanwhile, the weighted average (weighting by the size of each database) performance on the four databases is listed at the last row of Table 4. The average values further prove that the NRVPD return higher PLCC (0.812) and SRCC (0.757) values than the other methods. Therefore, the NRVPD outperforms the other state-of-the-art NR-IQA methods.

Besides, the statistical significance is estimated for further comparison. In this work, the *F*-test criterion, which calculates the residuals between ground truth values and the predicted qualities for comparison, is employed. And the confidence is set as 95%. The *F*-test results between the proposed NRVPD and these existing NR-IQA methods are listed in Table 5. There are three types of *F*-test result, where '1' means the proposed NRVPD performs statistically better than the compared method, '−1' means statistically worse, and '0' means indistinguishable. As can be seen, most of the values are positive (with '1'). For more detailed comparison on the results of the *F*-test, we can see that the NRVPD performs significantly better than the other ones on CSIQ, TID2013 and Wild-LIVE databases. For the LIVE database, the NRVPD is a little worse than CNN-SVR, and performs similar with HOSA, while better than the other methods. Therefore, the comparison result from the *F*-test is consistent with the analysis above (from Table 4), which further confirms the prediction accuracy of our NRVPD.

4.2.3. Cross-validation

Having evaluated the proposed NRVPD on each individual database, we try to prove that the NRVPD is not limited by the testing database. Since the three benchmark databases (i.e., LIVE, CSIQ, and TID2013) contain 4 common types of distortions (i.e., J2K, JPG, WGN, and GBN), these common distortions are chosen from the three databases for cross validation. In this experiment, the NRVPD is firstly trained on the whole CSIQ, and then tested on the other two databases. The comparison results with the other NR-IQA methods are shown in Table 6. The results on the TID2013 database show that the NRVPD obtains the largest PLCC/SRCC values against the other NR-IQAs, and the smallest RMSE. In other words, the proposed NRVPD performs obviously better than the other methods. On the LIVE database, the NRVPD still performs better than the other methods (has the best performances on all of the three measurement metrics). Meanwhile, we also train on the LIVE (TID2013) and test on the other two databses. Experimental results are shown in Tables 7 and 8, respectively. As

Table 6
Cross validation with training on CSIQ and testing on TID2013/LIVE.

DB	Crit.	NRVPD	IL-NIQE	NIQE	BRISQUE	DIIVINE
TID2013	PLCC	**0.9166**	**0.8727**	0.8223	0.5830	0.8127
	SRCC	**0.8882**	**0.8768**	0.8143	0.5706	0.7642
	RMSE	**0.5589**	**0.6825**	0.7955	1.1358	0.8146
LIVE	PLCC	**0.9191**	0.9136	**0.9179**	0.6436	0.5224
	SRCC	**0.9265**	0.9156	**0.9185**	0.6325	0.5209
	RMSE	**10.650**	10.991	**10.727**	12.253	13.652

Table 7
Cross validation with training on LIVE and testing on TID2013/CSIQ.

DB	Crit.	NRVPD	IL-NIQE	NIQE	BRISQUE	DIIVINE
TID2013	PLCC	**0.9080**	**0.8727**	0.8223	0.7211	0.8597
	SRCC	**0.9035**	**0.8768**	0.8143	0.7262	0.8489
	RMSE	**0.5856**	**0.6825**	0.7955	0.9685	0.7141
CSIQ	PLCC	**0.8977**	**0.9064**	0.8907	0.8404	0.8747
	SRCC	0.8617	**0.8801**	**0.8669**	0.8268	0.8537
	RMSE	**0.1245**	**0.1193**	0.1284	0.1532	0.1370

Table 8
Cross validation with training on TID2013 and testing on LIVE/CSIQ.

DB	Crit.	NRVPD	IL-NIQE	NIQE	BRISQUE	DIIVINE
LIVE	PLCC	0.8917	**0.9136**	**0.9179**	0.7898	0.6273
	SRCC	0.9019	**0.9156**	**0.9185**	0.7954	0.6211
	RMSE	11.233	10.990	**10.726**	**9.822**	12.468
CSIQ	PLCC	**0.8928**	**0.9064**	0.8907	0.8391	0.6589
	SRCC	**0.8709**	**0.8801**	0.8669	0.8081	0.6414
	RMSE	**0.1228**	**0.1193**	0.1284	0.1537	0.2126

Table 9
The D-Test, L-Test and P-test results on the Waterloo exploration database.

Algo.	D	L_s	P
NIQE	0.9109	0.9885	0.9937
IL-NIQE	0.9080	0.9926	0.9927
dipIQ	0.9346	0.9846	0.9999
NRVPD	0.9273	0.9917	0.9972

can be seen, the proposed NRVPD also performs well. Therefore, we can conclude that the performance of the NRVPD is more consistent with the HVS than the other NR-IQAs.

Moreover, the proposed NRVPD is trained on the LIVE database and test on the Waterloo exploration database [23]. The Waterloo exploration database contains 4744 pristine images, and 94,880 distorted images with four distortion types (i.e., J2K, JPG, WGN, and GBN) under 5 distortion levels (but each distorted image has no MOS/DMOS value). Meanwhile, three test criteria have been designed to measurement the performance on such database: 1) the pristine/distorted image discriminability test (*D*-test), which measures whether an IQA metric can distinguish the pristine image from the distorted one; 2) Listwise ranking consistency test (*L*-test), which evaluates the robustness of the IQA metric when rating distorted images (which have the same reference image and degraded by the same distortion type, but under different distortion levels); and 3) Pairwise preference consistency test (*P*-test), which measures the IQA metric on pairs of images whose quality is clearly discriminable. More details about the three test criteria can be found in [23]. The results are listed in Table 9. As can be seen, the proposed NRVPD performs well on this large database (whose *D*-test, *L*-test and *P*-test values are 0.9273, 0.9917, and 0.9972), which is comparable with the state-of-the-art algorithm (i.e., dipIQ [24]).

4.3. Analysis on the feature size

In this subsection, the effect of the OSP feature size on the quality prediction performance is demonstrated. As illustrated in Section 3.1, the image contents are represented by these fundamental OSPs, which are clustered by the K-means algorithm with a set of initial centroids. Here, we set the number of initial centroids with different values, and then verify the performance on TID2013 database (with four commonly used distortion types).

The relation between the performance (represented by PLCC and SRCC values) and the size of the OSPs (from 100 to 1200 bins) is shown in Fig. 12. From the two curves we can see that with the increase of the size from 100 to 600, the
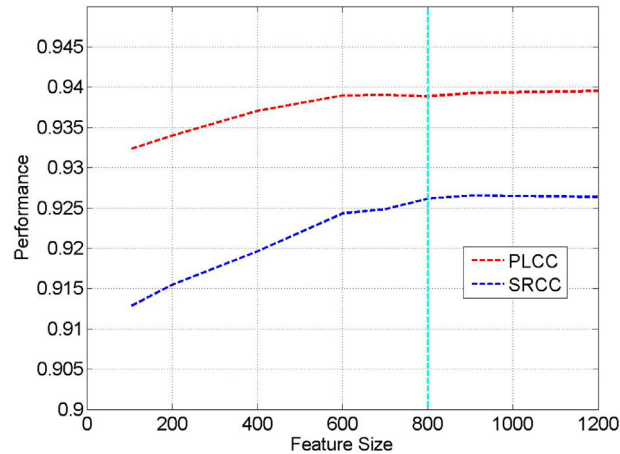
Fig. 12. The influence of the feature size on IQA performance, which comes from the TID2013 database (with four commonly used distortion types).

performance is gradually improved. After that, the increase is limited. Considering the trade between the prediction accuracy and the computational complexity, we finally set the size of the OSPs as 800 for all of the experiments.

5. Conclusion

In this work, we have designed a novel pattern to build a prior knowledge database for NR-IQA measurement. With the inspiration from the OS mechanism for visual structure extraction, we have designed an OS based visual pattern for image feature extraction. Moreover, by considering the rotation invariant and pattern similarity, a set of fundamental OSPs have been acquired. We have verified those fundamental OSPs with different image databases, and have proved that those proposed OSPs extract common visual features for content representation. Next, we have analyzed the structure degradations from different distortion types on those patterns, and found out that different distortions cause identifying degradations on them. According to this, a visual structure based knowledge database has been built to guide the quality measurement, and a novel learning based NRVPD methods has been introduced. Experimental results on these benchmark databases have confirmed that the NRVPD outperforms the stat-of-the-art NR-IQAs, and is highly consistent with the HVS.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, Frequency-tuned salient region detection, in: IEEE CVPR 2009, 2009, pp. 1597–1604.
[2] M. Bar, The proactive brain: memory for predictions, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 364 (1521) (2009) 1235–1243.
[3] F.W. Campbell, J.J. Kulikowski, Orientational selectivity of the human visual system, J. Physiol. 187 (2) (1966) 437–445.
[4] J.A. Cardin, L.A. Palmer, D. Contreras, Stimulus feature selectivity in excitatory and inhibitory neurons in primary visual cortex, J. Neurosci. 27 (39) (2007) 333–344.
[5] C.C. Chang, C.J. Lin, Libsvm: a library for support vector machines, 2001,
[6] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Conference on Computer Vision and Pattern Recognition, 1, IEEE, 2005, pp. 886–893.
[7] D. Ferster, K.D. Miller, Neural mechanisms of orientation selectivity in the visual cortex, Annu. Rev. Neurosci. 23 (2000) 441–471.
[8] R. Ferzli, L.J. Karam, A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB), IEEE Trans. Image Process. 18 (4) (2009) 717–728.
[9] 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 Netw. Learn.Syst. 24 (12) (2013) 2013–2026.
[10] D. Ghadiyaram, A.C. Bovik, Massive online crowdsourced study of subjective and objective picture quality, IEEE Trans. Image Process. 25 (1) (2016) 372–387.
[11] U. Grenander, M. Miller, Pattern Theory: From Representation to Inference, Oxford University Press, Inc., New York, NY, USA, 2007.
[12] K. Gu, G. Zhai, X. Yang, W. Zhang, Using free energy principle for blind image quality assessment, IEEE Trans. Multimed. 17 (1) (2015) 50–63.
[13] D. Hansel, C. Vreeswijk, The mechanism of orientation selectivity in primary visual cortex without a functional map, J. Neurosci. 32 (12) (2012) 4049–4064.
[14] R. Hong, J. Pan, S. Hao, M. Wang, F. Xue, X. Wu, Image quality assessment based on matching pursuit, Inf. Sci. 273 (2014) 196–211.
[15] W. Hou, X. Gao, D. Tao, X. Li, Blind image quality assessment via deep learning, IEEE Trans. Neural Netw. Learn. Syst. 26 (6) (2015) 1275–1286.
[16] S.u. Hussain, B. Triggs, Visual recognition using local quantized patterns, in: ECCV 2012, Springer, Berlin Heidelberg, 2012, pp. 716–729.
[17] E.C. Larson, D.M. Chandler, Categorical image quality (csiq) database, 2004,
[18] C. Li, A. Bovik, X. Wu, Blind image quality assessment using a general regression neural network, IEEE Trans. Neural Netw. 22 (5) (2011) 793–799.
[19] J. Li, J. Yan, D. Deng, W. Shi, S. Deng, No-reference image quality assessment based on hybrid model, Signal Image Video Process. 11 (6) (2017) 985–992.

[20] H. Liu, N. Klomp, I. Heynderickx, A no-reference metric for perceived ringing artifacts in images, IEEE Trans. Circt. Syst. Video Technol. 20 (4) (2010) 529–539.
[21] L. Liu, B. Liu, H. Huang, A.C. Bovik, No-reference image quality assessment based on spatial and spectral entropies, Signal Process. 29 (8) (2014) 856–863.
[22] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110.
[23] K. Ma, Z. Duanmu, Q. Wu, Z. Wang, H. Yong, H. Li, L. Zhang, Waterloo exploration database: new challenges for image quality assessment models, IEEE Trans. Image Process. 26 (2) (2017) 1004–1016.
[24] K. Ma, W. Liu, T. Liu, Z. Wang, D. Tao, Dipiq: blind image quality assessment by learning-to-rank discriminable image pairs, IEEE Trans. Image Process. 26 (8) (2017) 3951–3964.
[25] K. Ma, W. Liu, K. Zhang, Z. Duanmu, Z. Wang, W. Zuo, End-to-end blind image quality assessment using deep neural networks, IEEE Trans. Image Process. 27 (3) (2018) 1202–1213.
[26] R.A. Manap, L. Shao, Non-distortion-specific no-reference image quality assessment: a survey, Inf. Sci. 301 (2015) 141–160.
[27] A. Mittal, A. Moorthy, A. Bovik, No-reference image quality assessment in the spatial domain, IEEE Trans. Image Process. 21 (12) (2012) 4695–4708.
[28] A. Mittal, R. Soundararajan, A. Bovik, Making a completely blind image quality analyzer, IEEE Signal Process. Lett. 20 (3) (2013) 209–212.
[29] 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.
[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] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, S. Yao, Z. Lu, X. Yang, A locally adaptive algorithm for measuring blocking artifacts in images and videos, Signal Process. 19 (6) (2004) 499–506.
[32] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Kuo, Color image database TID2013: peculiarities and preliminary results, in: 2013 4th European Workshop on Visual Information Processing (EUVIP), 2013, pp. 106–111.
[33] M. Saad, A. 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.
[34] M.A. Saad, A.C. Bovik, C. Charrier, A DCT statistics-based blind image quality index, IEEE Signal Process. Lett. 17 (6) (2010) 583–586.
[35] H.R. Sheikh, K. Seshadrinathan, A.K. Moorthy, Z. Wang, A.C. Bovik, L.K. Cormack, Image and video quality assessment research at live, 2006,
[36] R. Soundararajan, A. Bovik, RRED Indices: reduced reference entropic differencing for image quality assessment, IEEE Trans. Image Process. 21 (2) (2012) 517–526.
[37] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process. 19 (6) (2010) 1635–1650.
[38] P.A. Van Der Helm, E.L.J. Leeuwenberg, Accessibility: a criterion for regularity and hierarchy in visual pattern codes, J. Math. Psychol. 35 (2) (1991) 151–213.
[39] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[40] Z. Wang, E.P. Simoncelli, Reduced-Reference Image Quality Assessment Using a Wavelet-Domain Natural Image Statistic Model, 5666, SPIE, 2005, pp. 149–159.
[41] J. Wu, W. Lin, G. Shi, L. Li, Y. Fang, Orientation selectivity based visual pattern for reduced-reference image quality assessment, Inf. Sci. 351 (2016) 18–29.
[42] J. Wu, W. Lin, G. Shi, A. Liu, Perceptual quality metric with internal generative mechanism, IEEE Trans. Image Process. 22 (1) (2013) 43–54.
[43] J. Wu, Y. Liu, W. Dong, G. Shi, W. Lin, Quality assessment for video with degradation along salient trajectories, IEEE Trans. Multimed. (2019), doi:10.1109/TMM.2019.2908377.
[44] X. Xie, Y. Zhang, J. Wu, G. Shi, W. Dong, Bag-of-words feature representation for blind image quality assessment with local quantized pattern, Neurocomputing 226 (2017) 176–187.
[45] J. Xu, P. Ye, Q. Li, H. Du, Y. Liu, D. Doermann, Blind image quality assessment based on high order statistics aggregation, IEEE Trans. Image Process. 25 (9) (2016) 4444–4457.
[46] W. Xue, X. Mou, L. Zhang, A.C. Bovik, X. Feng, Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features, IEEE Trans. Image Process. 23 (11) (2014) 4850–4862.
[47] P. Ye, D. Doermann, No-reference image quality assessment using visual codebooks, IEEE Trans. Image Process. 21 (7) (2012) 3129–3138.
[48] O. Yizhar, L.E. Fenno, M. Prigge, F. Schneider, T.J. Davidson, D.J. OShea, V.S. Sohal, I. Goshen, J. Finkelstein, J.T. Paz, K. Stehfest, R. Fudim, C. Ramakrishnan, J.R. Huguenard, P. Hegemann, K. Deisseroth, Neocortical excitation/inhibition balance in information processing and social dysfunction, Nature 477 (7363) (2011) 171–178.
[49] L. Zhang, L. Zhang, A.C. Bovik, A feature-enriched completely blind image quality evaluator, IEEE Trans. Image Process. 24 (8) (2015) 2579–2591.
[50] W. Zhou, L. Yu, W. Qiu, Y. Zhou, M. Wu, Local gradient patterns (lgp): an effective local-statistical-feature extraction scheme for no-reference image quality assessment, Inf. Sci. 397 (2017) 1–14.