Journal of Imaging Science and Technology® 60(6): 060503-1–060503-8, 2016.

No-Reference Image Quality Assessment Using Image Saliency for JPEG Compressed Images

Zengjie Song and Jiangshe Zhang
Department of Statistics, School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, 710049, China
E-mail: jszhang@mail.xjtu.edu.cn

Junmin Liu
Department of Information Science, School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, 710049, China

Abstract. *Image quality assessment (IQA) is one of the most important issues in the field of image processing. Traditional IQA methods usually assume that the "reference" or "perfect" image is given. Obviously, this assumption is limited because the reference image may not be available in most practical applications. In addition, the mechanisms of human visual system (HVS) have not been explicitly exploited in the majority of existing IQA methods. In this article, to reduce dependence on reference image while introducing one mechanism of HVS, we propose a new method, referred to as Image-Saliency-based No-reference Image Quality Index (ISNIQI), by incorporating image saliency derived from visual attention models into the quality assessment of JPEG compressed images. Since the high saliency image region attracts more visual attention, ISNIQI assigns larger weight to the quality of image region with higher saliency. Experimental results on JPEG compressed images demonstrate the effectiveness of the proposed method in terms of the correlation with subjective perception.* © *2016 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2016.60.6.060503]

INTRODUCTION

Assessing the visual quality of images is of fundamental importance to numerous image processing applications. Up to now, many approaches have been developed to measure the visual quality of images, including *subjective* and *objective* methods. For applications in which the human is the ultimate receiver of the visual signal, the only "correct" method of assessing image quality is the subjective methods. However, subjective methods are usually inconvenient, time consuming and expensive. Therefore, designing the objective *image quality assessment* (IQA) methods is a very important but challenging task.[1,2]

Based on the availability of a reference image, the objective methods can be classified into three categories. Most of the existing methods[3–6] are known as *full reference* (FR), meaning that the evaluation process requires all pixel information of a reference image. The second type of methods[7,8] is *reduced-reference* (RR) quality assessment, which only uses partial features of the reference image to help evaluate the quality of the distorted image. Although these FR and RR methods correlate highly with the human subjective judgments of quality, they are limited in practical applications due to the fact that the reference images, even the partial features of the reference images, are usually unavailable. Therefore, a *no-reference* (NR) or "blind" quality metric, which aims at predicting the quality of distorted images accurately and automatically without any knowledge of the reference images, is desirable. In this article, the discussion is confined to no-reference image quality assessment (NR-IQA) methods.

According to whether examining the exact prior knowledge of distortion, previous NR-IQA methods can be broadly categorized into two main classes: the *distortion-specific* methods[9–12] and the *distortion-generic* methods.[13–16] Due to the extensive applications of JPEG image format, many of these methods focus on or can be applied to the quality assessment of JPEG compressed images. For example, Wang et al.[9] proposed an NR method based on image features in pixel domain, which uses differential signals to estimate the blockiness and blur effects of JPEG compressed image. Mittal et al.[16] introduced a natural scene statistic (NSS)-based distortion-generic NR-IQA method in which a mapping is learned from feature space to quality scores using a regression module. However, these methods depend only on features in pixel domain or transform domains, and consider no perceptual property of *human visual system* (HVS) for image distortion. To imitate the evaluation of HVS, Liu et al.[17] designed an NR method named NPBM (no-reference perceptual blockiness metric) that not only extracts the blockiness features but also considers the *visual masking effect* on perceiving blocking artifacts, improving the consistency between IQA values and subjective scores. In our proposed method, we utilize techniques from the NPBM method to accurately measure the blocking artifacts of JPEG compressed images.

In order to design objective metrics that predict image quality consistent with what humans perceive, the mechanisms of HVS should be exploited.[18] It has been proven that the *image saliency* plays an important role in the HVS perception, that is, the image regions with high saliency have a significant impact on the perception process. As Itti et al. proposed in visual attention model IT,[19] the salient parts of the scene first attract our attention and then our

Received Dec. 13, 2015; accepted for publication Mar. 28, 2016; published online Nov. 17, 2016. Associate Editor: Marius Pedersen.
1062-3701/2016/60(6)/060503/8/$25.00

gaze. Similarly, Zhang et al.[20] argue that as a consequence of evolution, most vertebrates, including human, have a remarkable ability to automatically pay more attention to salient regions of the visual scene. Therefore, an effective IQA metric should incorporate the image saliency information. However, the above-mentioned methods take no account of this crucial mechanism of HVS.

Recently, increased awareness to the close relationship between image saliency and quality perception has led to a number of approaches that try to integrate image saliency into IQA metrics to potentially improve their prediction performance. According to previous classification criteria for objective IQA methods, these methods can be broadly categorized into FR and NR schemes. Generally speaking, FR methods[21–24] utilize visual attention data or ROI (region of interest) data recorded during eye tracking experiments to obtain saliency maps which are then added to several common fidelity metrics (such as SSIM, PSNR, VIF, etc.). Such elegant works provided a good demonstration that appropriately including image saliency information in the objective metrics can improve prediction performance of original IQA methods. However, it should be noted that the visual attention data or the ROI data used in these works are derived from eye tracking experiments on reference images, meaning that these methods cannot be used in situations where a fully automatic NR-IQA metric is needed. Instead of using specific data, NR methods attempt to incorporate image saliency information computed by visual attention models into IQA methods. One representative method belonging to this category is the training-based method IQVG.[25] This method first randomly samples a sufficient number of image patches guided by image saliency map and convolves each patch with Gabor filters to get a bag of features. Then, the image is represented by using a histogram to encode the bag of features. Support vector regression (SVR) is used to learn the mapping from feature space to image quality. While IQVG can be adaptable to different distortion type and achieve a good performance for quality estimation on the whole LIVE IQA database,[26] there are several aspects that should be remarked. Firstly, in order to ultimately obtain a trained model having better prediction performance, IQVG needs to empirically determine a number of parameters (such as patch number, patch size, five frequencies and four orientations used in Gabor filter, etc.), which causes the inconvenience of application. Secondly, IQVG consumes about 60 s to extract local features from 5000 patches of which the size is 11×11 on an Intel Pentium 2.13 GHz machine. That is to say, IQVG is time consuming relative to other common IQA methods which only cost a few seconds to compute a quality score. More importantly, IQVG cannot achieve a good prediction performance for some specific distortions, especially for JPEG compression.

In this article, we propose an Image-Saliency-based No-reference Image Quality Index (ISNIQI) that could achieve better prediction performance for JPEG compressed images while possessing a low computational complexity.

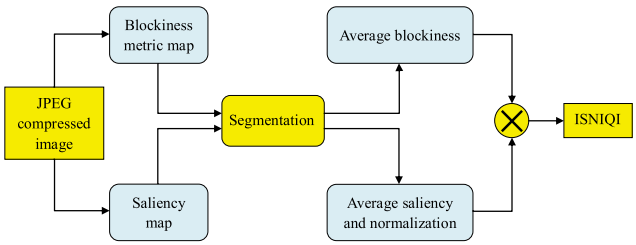

Figure 1. Schematic overview of the proposed method ISNIQI.

Since the high saliency image region attracts more visual attention, our proposed ISNIQI assigns larger weight to the quality of image region with higher saliency. The procedure to compute a prediction value of ISNIQI can be summarized as three steps: (1) computing the quality map of blocking artifacts by utilizing techniques from NPBM;[17] (2) employing a visual attention model to generate the saliency map as a weighting function; (3) integrating these two maps based on blocks to obtain a final quality score. The schematic overview of the proposed approach (ISNIQI) is summarized in Figure 1. Compared with the other state-of-the-art NR methods, our approach ISNIQI has following advantages for the quality assessment of JPEG compressed images. First, the prediction values of ISNIQI correlate better with subjective scores than other IQA indices evaluated. Second, only three parameters (i.e. two adjustment parameters and the number of sub-blocks) need to be tuned in ISNIQI, resulting in a more convenient combination with existing visual attention models. Third, ISNIQI has a relatively low computational complexity than other NR methods which need a domain transform or extract multiple features.

The rest of this article is organized as follows. We review the work of NPBM relevant to this article in Measurement of the Blocking Artifacts. Assessment Using Image Saliency first introduces the image saliency derived from several visual attention models and then presents in detail our proposed method. Following Experiment Results and Discussions, the Conclusion summarizes our work and gives some specific recommendations on the application scope of our method.

MEASUREMENT OF THE BLOCKING ARTIFACTS

The original NPBM method is built upon the specific structure information of the artifact itself combined with the properties of HVS by means of a simple and efficient model of visual masking.[17] It involves three steps to measure the blocking artifacts: (1) identifying the position of blocking artifacts; (2) calculating the degree of blockiness distortion; and (3) estimating the local visibility of the artifact to the human eye. For measuring the blocking artifacts of JPEG compressed image, we only need to implement the last two steps because the 8×8 non-overlapping blocks regularly arise along horizontal and vertical directions. The following discussion is carried out in the horizontal direction, but it can be done in the same way for the vertical direction.

1	1	0	1	1
1	2	0	2	1
1	2	0	2	1
1	2	0	2	1
1	1	0	1	1

Figure 2. The low-pass filter $\boldsymbol{L}$ for local luminance detection.

Local Pixel-based Blockiness Measure

In NPBM method, the local blockiness metric (LBM) is locally characterized as a blocking edge that stands out from its spatial vicinity, and is defined as the local gradient energy normalized by its neighboring pixels. Specifically, for a grayscale image $\boldsymbol{I}$, LBM at location (i, j) is estimated as

$$\mathrm{LBM}_h(i,j) = \begin{cases} \mathrm{BG}_h & \text{if } \mathrm{NBG}_h = 0,\ BG_h \neq 0 \\ \dfrac{\mathrm{BG}_h}{\mathrm{NBG}_h} & \text{if } \mathrm{NBG}_h \neq 0 \\ 0 & \text{if } \mathrm{NBG}_h = 0,\ BG_h = 0, \end{cases} \tag{1}$$

where BG_h denotes the local gradient energy at the blocking artifact, and NBG_h denotes the averaged gradient energy over its direct vicinity, which are given by

$$\mathrm{BG}_h = |\boldsymbol{I}(i,j+1) - \boldsymbol{I}(i,j)| \tag{2}$$

$$\mathrm{NBG}_h = \frac{1}{8} \sum_{x=-4,\ldots,4, x\neq 0} |\boldsymbol{I}(i,j+x+1) - \boldsymbol{I}(i,j+x)|. \tag{3}$$

Visual Masking Effect Measure

Visual masking effect, such as texture masking effect and luminance masking effect in NPBM,[17] is a phenomenon that the human eye's perception of one incitation can be changed by another incitation. By doing experiments, we find that the texture masking effect has little influence on the resulting performance of our method. Therefore, we only use the luminance masking effect estimation procedure of NPBM to imitate visual masking effect in the proposed method ISNIQI.

NPBM utilizes the visibility coefficient (i.e. VC_l) to quantify the luminance masking effect. For simplicity, the relationship between the visibility coefficient VC_l and the local luminance $\boldsymbol{I}_l$ is modeled by a nonlinear function for low background luminance and is approximated by a linear function at higher background luminance, i.e.,

$$\mathrm{VC}_l(i,j) = \begin{cases} \dfrac{\sqrt{\boldsymbol{I}_l(i,j)}}{9} & \text{if } 0 < \boldsymbol{I}_l(i,j) < 81 \\ \dfrac{(1-\beta)\cdot(81-\boldsymbol{I}_l(i,j))}{174} + 1 & \text{otherwise,} \end{cases} \tag{4}$$

where $0 < \beta < 1$ ($\beta = 0.7$ in Ref. 17) is used to adjust the slope of the linear part of this function, and the local luminance $\boldsymbol{I}_l(i,j)$ is calculated via a weighted low-pass filter $\boldsymbol{L}$ (see Figure 2) and defined as

$$\boldsymbol{I}_l(i,j) = \frac{1}{26} \sum_{x=1}^{5} \sum_{y=1}^{5} \boldsymbol{I}(i-3+x, j-3+y) \cdot \boldsymbol{L}(x,y). \tag{5}$$

ASSESSMENT USING IMAGE SALIENCY

The major advantage of the NPBM method is its effectiveness of measuring the blocking artifacts. However, it ignores the important influence of human eye's attention on perception of distortion, which will be further explained as follows.

For general image analysis tasks, we are usually interested in only a fraction of the whole image. For example, when people observe Figure 3(a), a top priority of observation is given to the region A (i.e. the propeller of airplane), followed by regions B and C. This results in that the distortion of region A is most likely to be perceived, while the noise of region C is almost neglected (see Fig. 3(b)–(d)). Consequently, the subjective perception quality of Fig. 3(d) is plausibility better than that of Fig. 3(c), and the quality of Fig. 3(b) is the worst one. This phenomenon may validate the speculation that the quality of image region attracting most attention accounts for the major part of the whole image quality.

In order to give different priorities to the quality of different image areas, ISNIQI employs image saliency to help

(a) (b) (c) (d)

Figure 3. Effects of visual attention on IQA: (a) different regions that human beings pay attention to (the propeller region A attracts the most visual attention, followed by regions B and C, which can also be found from the corresponding saliency maps in Fig. 4); (b)–(d) distorted images generated by adding the same intensity noise to marked regions in image (a), respectively.

assess image quality. Before we establish the ISNIQI method, it is necessary to make a brief introduction on image saliency.

Image Saliency

Studying on the image saliency derived from visual attention models has attracted much interest recently and there are now several frameworks and computational approaches available. According to Ref. 27, the visual attention models can be classified into object-based models and space-based models. Object-based models try to segment or detect objects to predict salient regions. In contrast, for those space-based models, the goal is to create image saliency maps that may predict which locations have higher probability of attracting human attention. Here, we mainly focus on the space-based models for obtaining the saliency map as a weighting function in IQA method.

Borj et al.[27] performed an exhaustive comparison of 35 state-of-the-art attention models over 54 challenging synthetic patterns, three natural image datasets, and two video datasets, using three evaluation scores. They find that the attention model IT[19] works better in locating a target over synthetic patterns; and the model SR[28] is fast and effective to provide a tradeoff between accuracy and speed necessary for many applications. Besides, another attention model MSS[29] not mentioned in Ref. 27 emphasizes well-defined borders, highlights whole object regions, and thus suppresses the background better than the above-mentioned models. Therefore, we select these eminent visual attention models to generate the saliency maps in our ISNIQI method. The visual comparison of saliency maps generated by these models for several images in LIVE database[26] is shown in Figure 4(b)–(d). We also exhibit the corresponding region-of-interest (ROI) maps in ROI-D database[30] as a ground truth for these visual attention models (see Fig. 4(e)).

Proposed ISNIQI Method

As illustrated in Fig. 4, the image saliency map can reflect the degree of human eye's attention on different image regions. Specifically, the higher the saliency value is, the more of visual attention the image region attracts, and hence, the more important the quality of corresponding image region should be. So, in order to improve the consistency between the objective IQA values and subjective scores, we propose the ISNIQI method employing image saliency to help assess the quality of JPEG compressed images, which aims at giving more importance to the qualities of conspicuous regions and penalizing measures in less salient regions. In other words, ISNIQI uses image saliency map as a weighting function to spatially integrate quality of different image regions.

The implementation of ISNIQI contains the following four steps: (1) computing the blocking artifacts quality map and the image saliency map, respectively; (2) partitioning these two maps into equally N non-overlapping sub-blocks along horizontal and vertical directions, respectively; (3) averaging the blockiness and saliency values over each block; and (4) integrating the two types of averaged values to produce an overall score of IQA. The detailed procedure of ISNIQI to assess quality of a JPEG compressed image with size $H \times W$ is described as follows.

1. The blockiness metric map $\boldsymbol{Q}_h$ along the horizontal direction is computed according to Eqs. (1) and (4), and each component of $\boldsymbol{Q}_h$ is given by
$$\boldsymbol{Q}_h(i,j) = \mathrm{LBM}_h(i,8j)^{\gamma_1} \cdot \mathrm{VC}_l(i,8j)^{\gamma_2} \tag{6}$$
where $i = 1, 2, \ldots, 8 \times (\lfloor \frac{H}{8} \rfloor - 1)$, and $j = 1, 2, \ldots, \lfloor \frac{W}{8} \rfloor - 1$. The superscripts γ_1 and γ_2 are used to adjust importance of different items in $\boldsymbol{Q}_h$.
2. Divide the quality map $\boldsymbol{Q}_h$ into N non-overlapping sub-blocks along the horizontal direction, and then define the averaged quality of each sub-block as
$$\boldsymbol{q}_h(k) = \frac{1}{|\boldsymbol{B}_{h,k}|} \sum_{(i,j)\in \boldsymbol{B}_{h,k}} \boldsymbol{Q}_h(i,j), \tag{7}$$
where $k = 1, 2, \ldots, N$, $\boldsymbol{B}_{h,k}$ denotes the kth sub-block of $\boldsymbol{Q}_h$, and $|\boldsymbol{B}_{h,k}|$ indicates the number of elements contained in $\boldsymbol{B}_{h,k}$.
3. Employ a visual attention model to generate the saliency map (i.e. $\boldsymbol{SM}$).
4. Divide the image saliency map $\boldsymbol{SM}$ into N non-overlapping sub-blocks along the horizontal direction, and then average the saliency of each sub-block as
$$\boldsymbol{SM}_h(k) = \frac{1}{|\boldsymbol{D}_{h,k}|} \sum_{(i,j)\in \boldsymbol{D}_{h,k}} \boldsymbol{SM}(i,j), \tag{8}$$
where $\boldsymbol{D}_{h,k}$ represents the kth sub-block of $\boldsymbol{SM}$, and then calculate the weight factor
$$\boldsymbol{\omega}_h(k) = \frac{\boldsymbol{SM}_h(k)}{\sum_{i=1}^{N} \boldsymbol{SM}_h(i)}. \tag{9}$$
5. Similar to $\boldsymbol{q}_h$ and $\boldsymbol{\omega}_h$, we can define the average qualities $\boldsymbol{q}_v$ and the weight factors $\boldsymbol{\omega}_v$ along the vertical direction, respectively. Finally, we can obtain the NR-IQA index by
$$\mathrm{ISNIQI} = \frac{1}{2} \sum_{k=1}^{N} [\boldsymbol{\omega}_h(k) \cdot \boldsymbol{q}_h(k) + \boldsymbol{\omega}_v(k) \cdot \boldsymbol{q}_v(k)]. \tag{10}$$

EXPERIMENTS AND DISCUSSIONS

Test Protocol

We conduct experiments on the JPEG subsets of LIVE database[26] and CSIQ database,[31] respectively. There are 29 reference images and 233 JPEG compressed images in LIVE database, and 30 reference images together with 150 JPEG compressed images in CSIQ database. The difference mean opinion score (DMOS) associated with distorted images is provided in both IQA databases. Four criteria are used to benchmark the comparison methods. The Pearson linear correlation coefficient (PLCC) provides an evaluation of prediction accuracy; the Spearman rank-order correlation coefficient (SROCC) and the Kendall rank-order correlation coefficient (KROCC) measure the prediction monotonicity;

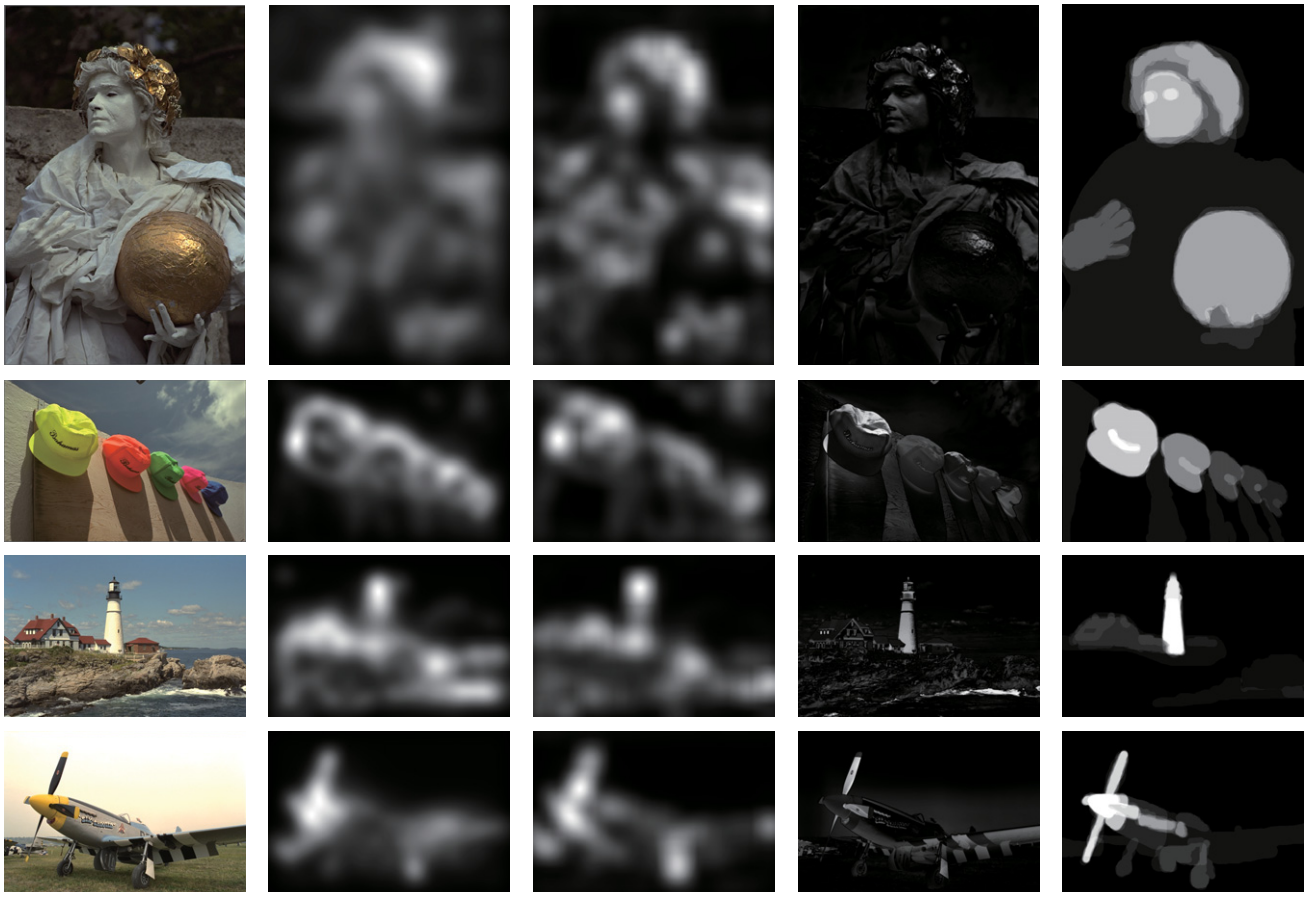

Figure 4. Visual comparison of saliency maps: (a) original images in LIVE database;[26] (b) saliency maps generated by IT attention model;[19] (c) saliency maps generated by SR attention model;[28] (d) saliency maps generated by MSS attention model;[29] (e) ROI maps in ROI-D database.[30] The highlight regions indicate the salient objects.

the root mean square error (RMSE) is also computed as an auxiliary comparison criterion. Definitions, explanations, and ways for calculating these four performance metrics can be found in Ref. 32.

To implement the ISNIQI method, we employ visual attention models IT,[19] SR[28] and MSS[29] to generate the image saliency map, respectively. Our method is compared with other 5 competitive NR-IQA methods, namely Wang's,[9] BLIINDS-II,[15] BRISQUE,[16] IQVG[25] and NPBM.[17] In what follows, the notation "ISNIQI-$*$" means that the attention model "$*$" is used to generate saliency map in ISNIQI method.

Only three parameters need to be determined in ISNIQI, that is, the adjustment parameters γ_1, γ_2 in Eq. (6) and the number of sub-blocks N. We select one-fifth of JPEG images from two IQA databases, respectively to conduct parameter estimation, and then apply resulting estimations to the whole LIVE and CSIQ JPEG subsets, respectively. To determine the adjustment parameters, we use grid search technique, in which the search interval is $[1, 2]$ for γ_1 and $[0, 1]$ for γ_2. It is emphasized that the mean value of blockiness metric maps $\boldsymbol{Q}_h$ and $\boldsymbol{Q}_v$ is taken as an IQA index during grid searching, meaning that the value of N does not have to be determined at this estimation step. Given estimations of γ_1 and γ_2, the final IQA indexes ISNIQI of selected JPEG images are computed with various value of N, ranging from 5 to 100. The value of N, with higher PLCC and SROCC, is accepted as the final estimation. In practice, we can only use ISNIQI-SR, the fastest version of ISNIQI, to determine the appropriate value of N, because of the same variation tendency of PLCC and SROCC with the increase of N for different versions of ISNIQI (see Figure 5). The estimations of all parameters are summarized in Table I, and are fixed among visual attention models.

Table I. Estimations of parameters for two IQA databases.

Database	No. of selected JPEG images	γ_1	γ_2	N
LIVE	46	1	0	10
CSIQ	30	1.5	0.3	35

Experimental Results on the Whole JPEG Subsets

The prediction performances of NR-IQA methods on two JPEG subsets are given in Table II. For each performance

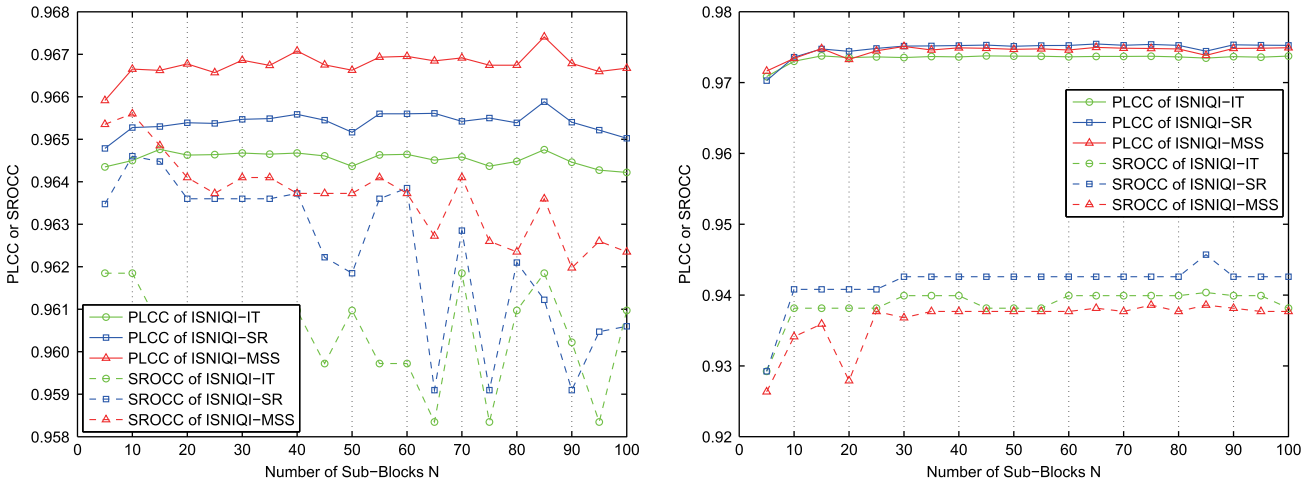

Figure 5. Variation tendency of PLCC and SROCC with the increase of N for different versions of ISNIQI: (a) LIVE; (b) CSIQ.

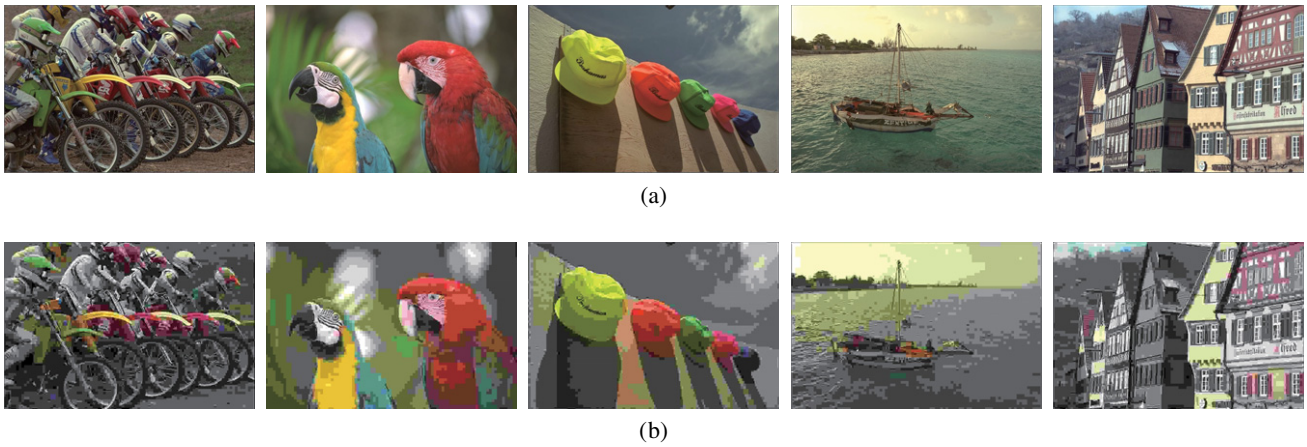

Figure 6. Two distorted versions of 5 reference images in LIVE IQA database: (a) the least apparent JPEG distortion; (b) the most apparent JPEG distortion.

measure, the top three results are highlighted in boldface. From Table II, we have the following findings. Firstly, all three versions of ISNIQI obtain competitive performances compared to the existing NR-IQA methods, especially outperform the comparison methods on the LIVE JPEG subset. Secondly, compared with NPBM method, ISNIQI with each attention model achieves better predictive results on both JPEG subsets. We believe that the success results from the consideration for the important influence of image saliency on IQA process, which makes the IQA values correlate better with the subjective scores. Thirdly, across both JPEG subsets, the prediction performances of ISNIQI-IT and ISNIQI-MSS are more consistent than that of ISNIQI-SR. While Wang's method gets the best results over other competing methods on CSIQ JPEG subset, its prediction performances are not even in the top four on LIVE JPEG subset. These results indicate that ISNIQI-IT and ISNIQI-MSS are more stable than other NR-IQA methods.

Influence of Distortion Levels on Performance

In this experiment, we investigated the influence of distortion levels on the saliency weights and therefore on the performance of ISNIQI method. Two distorted versions of each of the 29 reference images in LIVE IQA database[26] were used. The two distorted versions were chosen such that one of the images contained the least apparent JPEG distortion (with highest DMOS, see Figure 6(a)) and the other image contained the most apparent (suprathreshold) JPEG distortion (with lowest DMOS, see Fig. 6(b)).

As can be seen from Table III, in contrast with the most apparent distortion situation (i.e. highest level of distortion), ISNIQI achieves bigger improvements on NPBM method in the situation of the least apparent distortion (i.e. lowest level of distortion). For the quality assessment of JPEG images with suprathreshold distortion, ISNIQI has no obvious advantage compared to NPBM. These results reveal that incorporating image saliency information into objective quality models is very useful when the distortion is unapparent, and less helpful when the distortion is severe. We attribute the invalidation to the limited capacity of visual attention models to extract salient regions from images with suprathreshold distortion. As Vu et al. mentioned in Ref. 21, for spatially localized distortion, such as JPEG

Table II. Performance comparisons of NR-IQA methods on the whole JPEG subsets of LIVE database and CSIQ database.

Method	LIVE				CSIQ			
	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE
Wang's	0.9332	0.9211	0.7552	8.7198	**0.9742**	**0.9526**	**0.8084**	**0.0690**
BLIINDS-II	0.9235	0.8978	0.7146	9.3072	0.9401	0.8868	0.7023	0.1043
BRISQUE	0.9295	0.9145	0.7497	8.9473	0.9457	0.9038	0.7291	0.0995
IQVG	0.9089	0.8940	0.7115	10.1182	0.9449	0.9089	0.7334	0.1002
NPBM	0.9599	0.9523	0.8089	6.8014	0.9598	0.9351	0.7710	0.0859
ISNIQI-IT	**0.9623**	**0.9537**	**0.8127**	**6.5999**	**0.9637**	**0.9410**	**0.7850**	**0.0817**
ISNIQI-SR	**0.9634**	**0.9542**	**0.8133**	**6.5013**	0.9614	**0.9405**	0.7834	0.0842
ISNIQI-MSS	**0.9640**	**0.9536**	**0.8135**	**6.4498**	**0.9628**	0.9400	**0.7841**	**0.0826**

Table III. Performance comparisons of ISNIQI on the LIVE JPEG subset with different distortion levels.

Method	lowest level of distortion				highest level of distortion			
	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE
NPBM	0.8309	0.8457	0.6634	0.3834	0.5426	0.6773	0.5320	0.0282
ISNIQI-IT	0.9164	0.9073	0.7472	0.2763	0.5920	0.6576	0.5271	0.0270
ISNIQI-SR	0.9073	0.8925	0.7127	0.2897	0.5960	0.6768	0.4926	0.0269
ISNIQI-MSS	0.9171	0.8996	0.7275	0.2748	0.5656	0.6576	0.5271	0.0277

and JPEG2000, visual attentions indeed depend on the amount of distortion (near threshold versus suprathreshold). Therefore, developing a visual attention model which can effectively extract saliency information from images with various amount of distortion, is another important but challenging task.

Computational Complexity

The time complexity of each selected NR-IQA method on LIVE JPEG subset was also evaluated. Table IV shows the average time taken by each method on an Intel i2 2.66 GHz CPU and 4 GB RAM PC with the platform MATLAB R2010a. On the one hand, ISNIQI has a lower computational complexity than other NR-IQA methods which need a domain transform (i.e. BLIINDS-II) or extract a number of features (i.e. IQVG). And ISNIQI-SR is the fastest one compared to other versions of ISNIQI. On the other hand, the time cost of obtaining an IQA index by ISNIQI is a little higher than that of NPBM due to the additional process generating saliency map in ISNIQI. Hence, our future work will involve simplifying the entire method to further reduce computational complexity.

Table IV. Time cost of each NR-IQA method on LIVE JPEG subset.

Method	Time (seconds)
Wang's	0.0782
BLIINDS-II	207.3243
BRISQUE	0.3465
IQVG	43.7127
NPBM	0.7556
ISNIQI-IT	1.4509
ISNIQI-SR	0.8322
ISNIQI-MSS	3.1311

CONCLUSION

In this article, we propose a novel, efficient and effective NR-IQA method, ISNIQI, to measure the quality of JPEG compressed images. Based on the fact that the image saliency map can reflect the priority of human eye's attention on image regions, ISNIQI regards image saliency as the weight factor of image quality, and assigns larger weight to the quality of image region with higher saliency. Experiments performed on JPEG subsets of LIVE and CSIQ databases have shown that ISNIQI could achieve more accurate prediction performance compared with other NR-IQA methods evaluated in terms of the correlation with human subjective judgments of quality. Moreover, ISNIQI also has the advantage of convenient combination with existing visual attention models. In practical application, ISNIQI-IT and ISNIQI-MSS can be used to obtain the best overall performance, and ISNIQI-SR will be an appropriate choice if one pursues efficiency.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to the Editor (Dr. Chunghui Kuo) and two referees for their

valuable comments and suggestions, which greatly improved the presentation of this article. This work was supported in part by the National Basic Research Program of China (No. 2013CB329404), in part by the Natural Science Foundation of China (Nos. 11401465, 91230101 and 61572393), in part by the Key Project of National Natural Science Foundation of China (No. 11131006), in part by the Fundamental Research Funds for the Central Universities (No. xjj20140101), and in part by the Projects funded by China Postdoctoral Science Foundation (No. 2014M560781) and Shaanxi Postdoctoral Science Foundation.

REFERENCES

[1] Z. Wang, "Objective image quality assessment: facing the real-world challenges," *Proc. IS&T's Electron. Imag. Image and Qual. Syst. Perform. XIII* (IS&T, Springfield, VA, 2016), pp. 205.1–205.6.
[2] J. M. Sung, B. S. Choi, and Y. H. Ha, "Comparative display image quality evaluation based on an analytic network process for mobile devices," J. Imaging Sci. Technol. **60**, 20501.1–20501.7 (2016).
[3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from erroe visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).
[4] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," IEEE Trans. Image Process. **15**, 430–444 (2006).
[5] D. M. Chandler and S. S. Hemami, "VSNR: a wavelet-based visual signal-to-noise ratio for natural images," IEEE Trans. Image Process. **16**, 2284–2298 (2007).
[6] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment," IEEE Trans. Image Process. **20**, 2378–2386 (2011).
[7] Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," IEEE J. Sel. Top. Signal Process. **3**, 202–211 (2009).
[8] R. Soundararajan and A. C. Bovik, "RRED indices: reduced reference entropic differencing for image quality assessment," IEEE Trans. Image Process. **21**, 517–526 (2011).
[9] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," *Proc. IEEE Int'l. Conf. Image Process.* (IEEE, Piscataway, NJ, 2002), pp. I-477–I-480.
[10] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: application to JPEG2000," Signal Process., Image Commun. **19**, 163–172 (2004).
[11] H. R. Sheikh, A. C. Bovik, and L. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," IEEE Trans. Image Process. **14**, 1918–1927 (2005).
[12] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," IEEE Trans. Image Process. **18**, 717–728 (2009).
[13] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," IEEE Signal Process. Lett. **17**, 513–516 (2010).
[14] P. Ye and D. Doermann, "No-reference image quality assessment using visual codebooks," IEEE Trans. Image Process. **21**, 3129–3138 (2012).
[15] M. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: a natural scene statistics approach in the DCT domain," IEEE Trans. Image Process. **21**, 3339–3352 (2012).
[16] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Trans. Image Process. **21**, 4695–4708 (2012).
[17] H. Liu and I. Heynderickx, "A perceptually relevant no-reference blockiness metric based on local image characteristics," EURASIP J. Adv. Signal Process. **2009**, 1–14 (2009).
[18] C. Fredembach, J. Wang, and G. J. Woolfe, "Saliency, visual attention and image quality," *Proc. IS&T's Eighteenth Color and Imag. Conf.* (IS&T, Springfield, VA, 2010), pp. 128–133.
[19] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Mach. Intell. **20**, 1254–1259 (1998).
[20] L. Zhang, Y. Shen, and H. Li, "VSI: a visual saliency-induced index for perceptual image quality assessment," IEEE Trans. Image Process. **23**, 4270–4281 (2014).
[21] C. T. Vu, E. C. Larson, and D. M. Chandler, "Visual fixation pattern when judging image quality: effects of distortion type, amount, and subject experience," *Proc. IEEE Southwest Symp. Image Anal. Interp.* (IEEE, Piscataway, NJ, 2008), pp. 73–76.
[22] E. C. Larson, C. T. Vu, and D. M. Chandler, "Can visual fixation patterns improve image fidelity assessment?," *Proc. IEEE Int'l. Conf. Image Process.* (IEEE, Piscataway, NJ, 2008), pp. 2572–2575.
[23] U. Engelke and H.-J. Zepernick, "Framework for optimal region of interest-based quality assessment in wireless imaging," J. Electron. Imaging **19**, 011005 (2010).
[24] H. Liu and I. Heynderickx, "Visual attention in objective image quality assessment: based on eye-tracking data," IEEE Trans. Circuits Syst. Video Technol. **21**, 971–982 (2011).
[25] Z. Gu, L. Zhang, and H. Li, "Learning a blind image quality index based on visual saliency guided sampling and gabor filtering," *Proc. IEEE Int'l. Conf. Image Process.* (IEEE, Piscataway, NJ, 2013), pp. 186–190.
[26] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE Image Quality Assessment Database Release 2", http://live.ece.utexas.edu/research/quality, accessed November 2014.
[27] A. Borji, D. Sihite, and L. Itti, "Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study," IEEE Trans. Image Process. **22**, 55–69 (2013).
[28] X. Hou and L. Zhang, "Saliency detection: a spectral residual approach," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* (IEEE, Piscataway, NJ, 2007), pp. 1–8.
[29] R. Achanta and S. Susstrunk, "Saliency detection using maximum symmetric surround," *Proc. IEEE Int'l. Conf. Image Process.* (IEEE, Piscataway, NJ, 2010), pp. 2653–2656.
[30] U. Engelke and H.-J. Zepernick, "Psychophysical assessment of perceived interest in natural images: The ROI-D database," *Proc. IEEE Int'l. Conf. Vis. Commun. Image Process.* (IEEE, Piscataway, NJ, 2011), pp. 1–4.
[31] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," J. Electron. Imaging **19**, 011006 (2010).
[32] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," IEEE Trans. Image Process. **20**, 1185–1198 (2011).