No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization

Ke Gu, Weisi Lin, Guangtao Zhai, Xiaokang Yang, Wenjun Zhang, and Chang Wen Chen

***Abstract*—The general purpose of seeing a picture is to attain information as much as possible. With it, we in this paper devise a new no-reference/blind metric for image quality assessment (IQA) of contrast distortion. For local details, we first roughly remove predicted regions in an image since unpredicted remains are of much information. We then compute entropy of particular unpredicted areas of maximum information via visual saliency. From global perspective, we compare the image histogram with the uniformly distributed histogram of maximum information via the symmetric K-L divergence. The proposed blind IQA method generates an overall quality estimation of a contrast-distorted image by properly combining local and global considerations. Thorough experiments on five databases/subsets demonstrate the superiority of our training-free blind technique over state-of-the-art full- and no-reference IQA methods. Furthermore, the proposed model is also applied to amend the performance of general-purpose blind quality metrics to a sizable margin.**

***Index Terms*—Image quality assessment, no-reference/blind, contrast distortion, information maximization, saliency**

I. Introduction

A PICTURE is worth a thousand words. Currently, images are playing an increasingly important role in recording information, communicating thought and expressing emotion in our daily lives. Facing huge amount of visual data being created, stored, transmitted and consumed every moment, it is impossible to monitor incessantly by manual labor. In such condition, a real-time system towards accurately evaluating and controlling visual data is eagerly required. This promotes the development of objective image quality assessment (IQA) metrics, which predict the visual quality using mathematical models to simulate subjective opinion scores [1].

There are generally three types of objective IQA methods. The first type is the so-called *full-reference (FR)* IQA. In literature, the primary principle behind most FR-IQA models is that the human visual system (HVS) is strongly sensitive to the degradations in image structures [2-9]. The second type in the IQA study is *reduced-reference (RR)* IQA, which supposes partial information or few features extracted from the original image are available [10-11]. The third type of objective IQA methods are currently popular *blind/no-reference (NR)* IQA models. A majority of this type of methods depend on statistical regulations, e.g. [12-15].

Ke Gu and Weisi Lin are with the School of Computer Science and Engineering, Nanyang Technological University, Singapore, 639798 (e-mail: guke.doctor@ntu.edu.sg; wslin@ntu.edu.sg).

G. Zhai, X. Yang, W. Zhang are with Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, 200240, China (email: zhaiguangtao/xkyang/zhangwenjun@sjtu.edu.cn).

C. W. Chen is with the State University of New York at Buffalo, Buffalo, NY 14260 USA (email: chencw@buffalo.edu).

Lately, the IQA of contrast distortion has attracted broad attentions [16-20], since this study can be used for image contrast enhancement [21-22]. Nonetheless, a vast majority of existing IQA models were found to not well correlate with the human judgment of quality for contrast-altered images [20]. To cope with this task, the reduced-reference image quality metric for contrast change (RIQMC) was devised based on phase congruency and information statistics of the histogram, acquiring superior performance beyond existing models and managing to enhance original natural images [20].

Despite the successfulness of the RIQMC in the contrast-distorted IQA and contrast enhancement, it inevitably needs a single number, the phase congruency based entropy of the original image, as the reduced reference, and is thus highly limited to contrast enhancement for original natural images. It is obvious that an IQA method with the ability to blindly assess a wider scope of input images, such as underexposed or overexposed images, is much more desirable. To this aim, we are towards developing a new no-reference image quality metric for contrast distortion (NIQMC) based on the concept of information maximization.

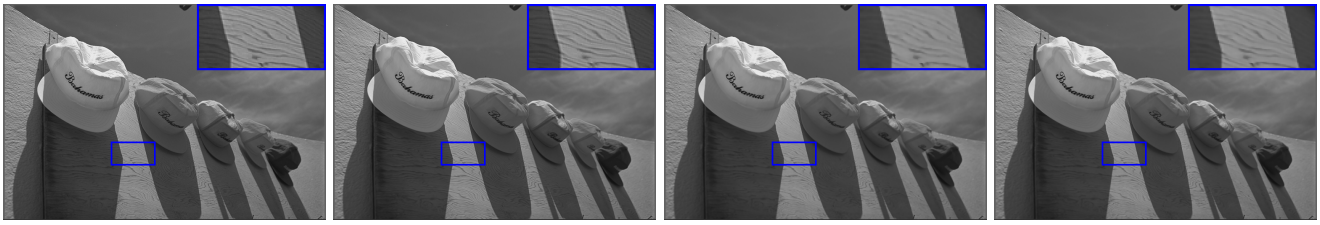
In particular, the proposed NIQMC metric performs a local and global manner. Given a visual signal, we first concentrate on its local details. Generally speaking, high contrast images communicate much information. An image signal contains a great amount of unpredictable information which is of much value, so we first remove predictable components from the image and leave unpredictable remains. Then entropy of the remains compared to that of the whole pixels of the input image will better represent the amount of valid information. We might also account for this viewpoint from the angle of free energy based brain principle, as discussed in Section II. Considering the fact that human eyes are attracted more by the regions of greater amount of information, visual saliency detection technique is also used to search for the optimal areas which have maximum information, and compute the entropy of the selected regions to be the local quality measure.

The second consideration for the given image comes from the global perspective. Likewise, in terms of the information maximization, it is reasonable to assume that the high-quality image is of the histogram towards uniform distribution. We in this research compare the symmetric Kullback-Leibler (K-L) divergence between the input image histogram and the uniformly distributed one to be the global quality measure. Finally, we calculate the linearly weighted mean of the two measures stated above, yielding the overall quality score of the input contrast-distorted image.

The remainder of this paper is arranged below. In Section

(a) Input image (b) AR model (c) Bilateral filtering (d) Semi-parametric model

Fig. 1: The sample image "buildings" and the associated filtered images with AR model, bilateral filtering and semi-parametric model.



(a) Input image (b) AR model (c) Bilateral filtering (d) Semi-parametric model

Fig. 2: The sample image "five hats" and the associated filtered images with AR model, bilateral filtering and semi-parametric model.

II, we present the devised blind NIQMC model in detail. In Section III, a thorough comparison across our algorithm and a large set of IQA methods is conducted on the contrast related databases [16, 20, 23-25], and through trials the proposed algorithm was found to improve general-purpose NR metrics. We lastly conclude this paper in Section IV.

II. Blind NIQMC Metric

In most conditions, human beings can boost the efficiency of information acquirement through particular mechanisms, such as visual saliency, which was proven to have a close tie with a neural circuit in the primate visual cortex [26]. On the basis of the essential behavior of human beings for achieving information as much as possible, this paper designs the blind NIQMC metric via information maximization. In other words, the main principle behind our NR-IQA model lies in that an image with more valuable information has better quality. We suppose that the HVS combines the local and global strategies to perceive a visual signal, judging its quality score and salient regions. Based on this, our blind NIQMC model attempts to predict the visual quality of contrast-altered images.

A. Local Quality Measure

The first consideration of our approach is devoted to the measurement of local details. In our common sense, an image with large contrast represents much meaningful information. Nevertheless, there is a great amount of residue information contained in most images, such as large areas of blue sky or green grassland in the background, which generally provide little information. With this, we first discard the predictable components from the image, which are estimated by using a semi-parametric model based on the autoregressive (AR) model and the bilateral filtering.

The AR model is simple and effective in simulating a broad range of natural scenes by adjusting its parameters [27]. Its parameters were shown to invariant to object transformations. In this paper we leverage the AR model following the same strategy used in [13]. The AR model is however oftentimes not stable at image edges. We take an example to illustrate this problem. As can be seen from Figs. 1(a)-(b), the text regions have been heavily devastated by the AR model, ushering the ringing artifacts. Akin results can be found at the buildings' edges as well. So we further take advantage of the bilateral filtering [28], which is a non-linear filtering with good edge-preserving ability and easy to establish and calculate.

As illustrated in Fig. 1, in comparison to the input image, one can see that the bilateral filtering is able to protect edges better than the AR model, without introducing any ringing artifact. Yet another example shown in Fig. 2 indicates that the bilateral filtering falls in processing texture regions, which leads to a great amount of spatial frequency reduction. On the contrary, the AR model is applicable to texture synthesis and thus preserves texture parts well. These two models are good at processing smooth areas. So it is natural to combine the merits of the AR model and the bilateral filtering, to acquire better results on edge, texture and smooth regions. By this guidance, we propose the semi-parametric model via a linear fusion, deriving the estimation of predictable data ($\mathbf{y}^p$):

$$y_i^p = \frac{\mathcal{Y}^k(y_i)\hat{\mathbf{a}} + w\mathcal{Y}^k(y_i)\mathbf{b}}{1+w} \tag{1}$$

where $\hat{\mathbf{a}}$ and $\mathbf{b}$ are separately estimated parameters using AR model and bilateral filtering; $\mathcal{Y}^k(y_i)$ defines a vector consisting of k member neighborhood of y_i; w is assigned to be 4 for emphasizing the strength of the edge-preserving effect that comes from the bilateral filtering. We show the results based on the semi-parametric model in Figs. 1-2(d). It is noted that the adaptive weighting scheme should be better but may cause much computational cost. Here we adopt constant weights in (1) and the future work will be devoted to the exploration of adaptive weighting strategy.

The filtered image can be treated as an approximation of the predictable information, which can be explained from the

angle of free energy principle [29]. To specify, the recently revealed free energy based brain theory unifies some existing brain theories in biological and physical sciences regarding human action, perception, thinking and learning. The basic premise is based on a supposition that the cognitive process is governed by an internal generative model in human brain. With this model, the human brain can separate an input image into the orderly (predicted) and disorderly (unpredicted) parts[1]. According to the analysis in [13], the authors have confirmed that the internal generative model can be approximated with the AR model. In comparison, by bringing about the bilateral filtering, this paper develops a more reliable semi-parametric model that has well performance at edge, texture and smooth areas. We can use parameter vectors $\hat{\mathbf{a}}$ and $\mathbf{b}$ to control the internal generative model. Thus, the unpredictable component is obtained by discarding the predictable data from the input image:

$$y_i^u = y_i - y_i^p \tag{2}$$

and then entropy of the error map $\mathbf{y}^u$ is computed by

$$E(p) = -\int p(t)\log p(t)dt \tag{3}$$

where $p(t)$ indicates the probability density of grayscale t.

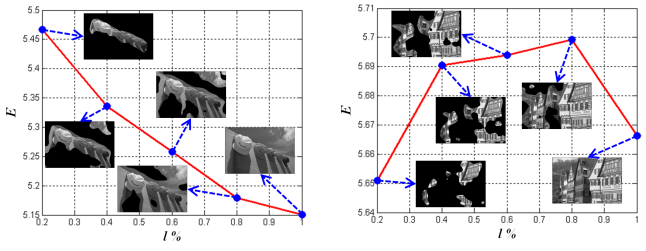
In addition, there still exists an important problem about the selection of suitable regions. This consideration is from an ordinary problem. If you have seen the famous portrait "Mona Lisa", do you remember what the foreground is? For the vast majority, an elegant lady with a mysterious smile will emerge in their minds. But if you ask what the background is, most people might remember nothing. That is, we human beings will focus on some "significant" regions although we have adequate time to see the whole image.

On the basis of information maximization, we assume that human beings wish to select maximum-information areas to be perceived. In order to maintain the semantic information, we restrict the chosen areas not less than one fifth of the size of the image. In our study, visual saliency has been used for optimal areas selection, because on one hand salient regions are generally what we easily remember, and on the other hand the aforesaid semi-parametric model can be also managed to detect visual saliency. It should be noted that visual saliency is a different concept from what we apply in this work, and instead it just provides several candidate regions which are likely to be of maximum information.

More concretely, we take into account the lately developed free energy inspired saliency detection technique (FES) [30]. The FES model performs at a small scale through resizing an image to a coarse 63×47 pixel representation. In light of the semi-parametric model similar to (1), the FES algorithm estimates the error map and computes its local entropy map in each color channel, and then combines three filtered and normalized local entropy maps in different color channels to produce the final saliency map.

In our NIQMC metric, we do not directly pick the salient areas for weighting, akin to the strategy adopted in many IQA approaches [2, 8, 31, 32, 33], because it has been presented

[1]Interested readers can be directed to [29] for more relevant information.



(a) (b)

Fig. 3: The trend of entropy values (E) with different percentages ($l\%$) on the two different types of images shown in Figs. 1-2.

above that the maximum-information region is required and the saliency map is merely used for assistance. Considering that if an image contains obvious foreground and background it will have comparatively centralized salient regions (i.e. the foreground) which communicate the valuable (unpredictable) information, or else salient areas may be distributed dispersedly and thus the valuable (unpredictable) information will be represented almost using the entire image.

As thus, after obtaining the saliency map, we sort all the pixels in $\mathbf{y}^p$ to form the vector $\mathbf{y}^s$ according to the salient importance. Next we calculate entropy values of the former $\{l_1, l_2, ..., l_n\}\%$ pixels in $\mathbf{y}^s$, denoted as $\{E_{l_1}, E_{l_2}, ..., E_{l_n}\}$. Finding a good tradeoff between efficacy and efficiency, we set $\{l_1, l_2, l_3, l_4, l_5\}$ as $\{20, 40, 60, 80, 100\}$ in this work. We indicate how the optimal areas selection works in Fig. 3. As we can see, the "five hats" image has definite foreground (i.e. five hats) and background, and in such condition the most 20% salient pixels are generally considered as the maximum-information region. In contrast, the "buildings" image cannot be divided into the foreground and background clearly, or in other words, its foreground (i.e. six buildings) nearly occupies the entire image, and thereby we think of the majority of the most 80% salient pixels to be the maximum-information area. Finally, we define the local quality measure by

$$Q_L = \max\{E_{l_1}, E_{l_2}, ..., E_{l_5}\}. \tag{4}$$

B. Global Quality Measure

The second consideration of our approach comes from the perspective of global information measure. A chief concept related to the image contrast is entropy, which ignores the influence of pixels' locations but rather merely considers the distribution of pixel values. More precisely, entropy supposes that the uniformly distributed histogram $\mathbf{u}$ corresponds to the maximum information, and thus an image whose histogram $\mathbf{h}$ is alike to $\mathbf{u}$ has the large global information. That is to say, the smaller difference between $E(\mathbf{u})$ and $E(\mathbf{h})$, namely $\Delta E = E(\mathbf{u}) - E(\mathbf{h}) = \int u(t)\log u(t) - h(t)\log h(t)dt \geq 0$, the higher contrast the image has. Whereas, we notice that this metric overlooks the interaction between $\mathbf{u}$ and $\mathbf{h}$. So the advanced K-L divergence, one of the most typical *distance* comparing two probability distributions in information theory, is exploited. Given two probability densities $\mathbf{h}_0$ and $\mathbf{h}_1$, the

K-L divergence is defined as

$$D_{\mathrm{KL}}(\mathbf{h}_1\|\mathbf{h}_0) = -\int h_1(t)\log h_0(t)dt + \int h_1(t)\log h_1(t)dt$$
$$= H(\mathbf{h}_1,\mathbf{h}_0) - E(\mathbf{h}_1) \quad (5)$$

where $H(\mathbf{h}_1,\mathbf{h}_0)$ is the cross entropy of $\mathbf{h}_1$ and $\mathbf{h}_0$. Via the usage of the K-L divergence, the interaction between $\mathbf{h}_1$ and $\mathbf{h}_0$ has been included.

The K-L distance is however non-symmetric and easy to cause some troubles in real applications. In [34], the authors present simple examples to illustrate that the ordering of the arguments in the K-L divergence might produce substantially different results. We accordingly resort to the symmetric K-L divergence. So far, many symmetric forms have been devised [34], e.g. arithmetical mean, geometric mean, and harmonic mean. Apart from the three functions stated above, there still exists a symmetrized and smoothed format, dubbed as the Jensen-Shannon (J-S) divergence:

$$D_{\mathrm{JS}}(\mathbf{h}_0,\mathbf{h}_1) = \frac{D_{\mathrm{KL}}(\mathbf{h}_0\|\mathbf{h}_\Delta) + D_{\mathrm{KL}}(\mathbf{h}_1\|\mathbf{h}_\Delta)}{2} \quad (6)$$

where $\mathbf{h}_\Delta = \frac{1}{2}(\mathbf{h}_0+\mathbf{h}_1)$.

It was observed by tests that, as opposed to the symmetric forms based on arithmetical, geometric and harmonic means, the use of the J-S divergence and 128-bin histogram results in around 2% performance gain. Therefore, given the histogram $\mathbf{h}$ and $\mathbf{u}$ of pixel values, the global quality measure is defined by

$$Q_G = D_{\mathrm{JS}}(\mathbf{h},\mathbf{u}). \quad (7)$$

Notice that the developed local and global quality measures have reversed meanings; that is, the higher local Q_L (or the smaller global Q_G) indicates that the image has the greater contrast and better quality.

C. Combined Quality Measure

According to the concept of information maximization, we have proposed two quality measures. By the predictable data removal and the optimal region selection, the former one quantifies the valuable information from the aspect of local details. Enlightened from the underlying concept of the practical HE technique, the latter part in our blind NIQMC metric uses the symmetrized and smoothed J-S divergence to measure whether the input histogram is properly distributed compared to the uniform distribution. From the viewpoint of working, the two parts above play complementary roles. As thus, they are integrated to approximate the HVS perception to the visual quality of contrast-altered images. Since these two measures are of the same dimension (i.e. entropy), we can straightforwardly combine them together. The NIQMC is therefore defined as a simple linear fusion of the two quality measures:

$$\mathrm{NIQMC} = \frac{Q_L + \gamma Q_G}{1+\gamma} \quad (8)$$

where γ is a constant weight that is used for controlling the relative significance between the local and global strategies. In this study, we set γ as -2.2 by making the proposed NIQMC metric have the best correlation performance measure on the TID2013 database. Also, it is easy to find that the semi-parametric model does not work efficiently due to the use of pixel-wise AR model. So this paper adopts the sampling method for decreasing the implementation time [35]; that is, the semi-parametric model are computed once every m pixels in both horizontal and vertical directions. Here we let $m = 7$, because it was found by tests that this choice degrades the performance to a small extent but rather largely reduces the computational complexity of the proposed NIQMC metric.

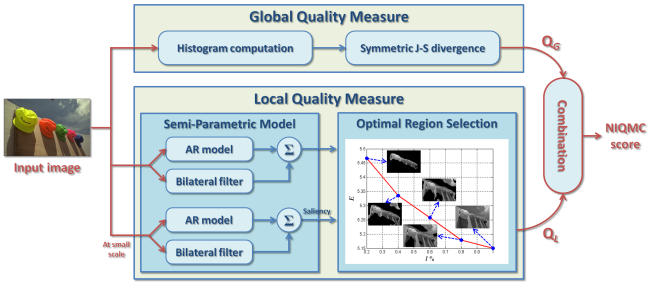


Fig. 4: The basic framework of our blind NIQMC metric.

Finally, we illustrate the basic framework of the proposed blind NIQMC approach in Fig. 4, in order for helping readers to effortlessly understand how to deploy the metric. Given an image signal, we first compute its global quality measure through using the symmetric and smoothed J-S divergence. Subsequently, using the semi-parametric model to process the image at the original and small scales respectively, we conduct the optimal region selection to predict the local quality measure. The NIQMC score is lastly derived via the combination of the global and local estimations. The implementation code will be available to the public soon at https://sites.google.com/site/guke198701/publications.

III. Experimental Results

In this section, we validate our training-free blind NIQMC model and compare with a large body of classical and state-of-the-art IQA algorithms. FR-IQA metrics involve state-of-the-art FSIM [2], IGM [5], LTG [7] and VSI [8]. NR-IQA methods are made up of prevailing BRISQUE [12], NFERM [13] and IL-NIQE [14]. To the best of our knowledge, there exist five databases (CID2013, CCID2014, CSIQ, TID2008 and TID2013) concerning image contrast, which were chosen as the testing bed in this work.

The performance of an IQA method is typically evaluated from two aspects from the angle of prediction power [36], namely prediction accuracy and prediction monotonicity. The computation of the correlation performance usually needs a five-parameter regression procedure to remove the nonlinearity of predicted scores. Thereafter, we calculate three performance indices [36], including the Spearman rank order correlation coefficient (SRCC) and the Kendall's rank order correlation coefficient (KRCC) for evaluating the prediction monotonicity, and the Pearson linear correlation coefficient (PLCC) for the prediction accuracy. Note that a value close to one for PLCC, SRCC and KRCC represents the superior correlation in line with subjective human ratings.

TABLE I: Performance indices on contrast related datasets and average results. We bold the top two models.

Quality Metrics	Type	CID2013 (400 images) [16]			CCID2014 (655 images) [20]			CSIQ (116 images) [23]		
		PLCC	SRCC	KRCC	PLCC	SRCC	KRCC	PLCC	SRCC	KRCC
FSIM [2]	FR	0.8574	0.8486	0.6663	0.8201	0.7658	0.5707	0.9378	0.9420	0.7883
IGM [5]	FR	0.8467	0.8246	0.6470	0.7992	0.7246	0.5356	0.9492	**0.9547**	**0.8174**
LTG [7]	FR	**0.8656**	**0.8605**	**0.6723**	**0.8384**	**0.7901**	**0.5938**	**0.9560**	0.9414	0.7880
VSI [8]	FR	0.8571	0.8506	0.6579	0.8209	0.7734	0.5736	**0.9532**	**0.9504**	**0.8096**
BRISQUE [12]	NR	0.3351	0.2552	0.1745	0.3575	0.2123	0.1445	0.1471	0.0473	0.0365
NFERM [13]	NR	0.4074	0.3497	0.2385	0.4181	0.3616	0.2470	0.4831	0.3742	0.2667
IL-NIQE [14]	NR	0.5682	0.5273	0.3708	0.5764	0.5121	0.3590	0.5468	0.5005	0.3510
NIQMC (Pro.)	NR	**0.8691**	**0.8668**	**0.6690**	**0.8438**	**0.8113**	**0.6052**	0.8747	0.8533	0.6689

Quality Metrics	Type	TID2008 (200 images) [24]			TID2013 (250 images) [25]			Average (1621 images)		
		PLCC	SRCC	KRCC	PLCC	SRCC	KRCC	PLCC	SRCC	KRCC
FSIM [2]	FR	0.6880	0.4403	0.3348	0.6819	0.4413	0.3588	0.8001	0.7086	0.5481
IGM [5]	FR	**0.6950**	0.3630	0.2690	**0.6891**	0.3717	0.2935	0.7918	0.6667	0.5130
LTG [7]	FR	0.6795	**0.4655**	0.3285	0.6749	0.4639	0.3458	**0.8087**	**0.7279**	**0.5561**
VSI [8]	FR	0.6819	0.4571	**0.3450**	0.6785	**0.4643**	**0.3705**	0.8002	0.7184	0.5518
BRISQUE [12]	NR	0.0786	0.1181	0.0787	0.1429	0.0551	0.0359	0.2694	0.1752	0.1193
NFERM [13]	NR	0.2705	0.2162	0.1472	0.2423	0.1956	0.1320	0.3748	0.3160	0.2163
IL-NIQE [14]	NR	0.2244	0.1833	0.1223	0.2275	0.1517	0.1030	0.4750	0.4189	0.2927
NIQMC (Pro.)	NR	**0.7767**	**0.7324**	**0.5419**	**0.7225**	**0.6458**	**0.4687**	**0.8253**	**0.7927**	**0.5967**

As given in Table I, we list the performance results of our blind NIQMC technique and seven computing IQA methods. Image contrast distortion is quite difficult to evaluate even though state-of-the-art FR-IQA approaches can resort to the help of the whole original image [20]. Despite this, in the recently constructed CID2013 and CCID2014 databases that are dedicated to the distortion type of contrast adjustment, the NIQMC has attained encouraging results, beyond 0.8 in accordance with SRCC[2]. Compared with other metrics, one can see that our approach is clearly superior to all the blind quality metrics and the majority of FR-IQA models that are usually hard to be matched by blind IQA algorithms because of the use of the overall original image, and moreover, it is even on par with the lately designed FR LTG method.

As for the CSIQ database, it can be easily found that the blind NIQMC model outperforms all the NR-IQA algorithms included in this paper, but is somewhat inferior to the state-of-the-art FR quality metrics. From the results of correlation performance on the TID2008 and TID2013 databases, our NIQMC approach performs better than the FR- and NR-IQA models considered.

A comprehensive comparison of each competing quality metric is further conducted via the weighted average referring to the number of images in each database. We illustrate the associated results in Table I as well. Apparently, the NIQMC has constantly achieved the optimal prediction monotonicity (using SRCC and KRCC) and prediction accuracy (using PLCC), higher than the second-place LTG method to a large extent, about 9% in terms of the important index SRCC.

A good IQA metric should be simultaneously of efficiency and efficacy. We therefore compute the mean implementation time of all the 655 contrast-adjusted images with the size of 512×768. The working platform uses the MATLAB2010a on a computer with Intel i7-2600 processor at 3.40GHz and 4GB memory. Table II tabulates the average computational cost of each testing quality metric. With a series computing, the devised NR NIQMC model takes less than three seconds to assess a 512×768 color image. Considering that the local and global quality measures are independent of each other, and the computation for each image pixel is also independent in the AR model or the bilateral filtering, we are very likely to conduct the parallel computing to reduce the implementation time to a large extent in the practical application.

[2]One important application of IQA methods is to tell which one has the better (or best) quality across two (or many) pictures. SRCC measuring the prediction monotonicity exactly illustrates this, and thus it is usually treated as one of the most important indices in the IQA [2, 7, 20, 35].

TABLE II: Average computational time on the CCID2014 database.

Metrics	FSIM	IGM	LTG	VSI
Time (second/image)	0.6746	18.326	0.0452	0.2106

Metrics	BRISQUE	NFERM	IL-NIQE	NIQMC
Time (second/image)	0.2760	41.832	3.0638	2.9001

TABLE III: The number and ratio of image pairs in accordance with subjective human ratings in each IQA method. We highlight the best two performed metrics with boldface.

Quality Metrics	Group 1		Group 2	
	Pairs	Ratio	Pairs	Ratio
FSIM [2]	11,232	80.05%	563	87.97%
IGM [5]	10,946	78.01%	564	88.12%
LTG [7]	**11,382**	**81.11%**	**570**	**89.06%**
VSI [8]	11,319	80.67%	561	87.66%
BRISQUE [12]	6075	43.29%	165	25.78%
NFERM [13]	8198	58.42%	499	77.97%
IL-NIQE [14]	3843	27.39%	71	11.09%
NIQMC (Pro.)	**11,670**	**83.17%**	**578**	**90.31%**

An important role of IQA methods is to judge which one is better between two images, especially one distorted image and its original counterpart. Therefore, a relevant experiment is carried out on the CCID2014 database. In this database, there are totally 14,032 image pairs, each of which associates to the same source image. For each metric, we tabulate the number of image pairs (and the ratio) whose orders are the same with those of subjective MOS values, as provided in "Group 1" of Table III. One can see that the proposed blind NIQMC metric has achieved the optimal result, far beyond state-of-the-art NR-IQA algorithms. Moreover, we also carry

TABLE IV: Sensitivity test (SRCC) of the parameter γ in (8).

Dataset \ γ	-2	-2.1	**-2.2**	-2.3	-2.4
CID2013	0.8662	0.8667	**0.8668**	0.8672	0.8673
CCID2014	0.8106	0.8110	**0.8113**	0.8117	0.8118
CSIQ	0.8520	0.8519	**0.8533**	0.8525	0.8510
TID2008	0.7328	0.7325	**0.7325**	0.7309	0.7305
TID2013	0.6453	0.6453	**0.6458**	0.6443	0.6443

TABLE V: Performance on TID2013 and cross-validation on LIVE, CSIQ and LIVEMD databases. We bold the best performed one.

Quality Metrics	Number of Features	TID2013 (3000 images) [25]		
		PLCC	SRCC	KRCC
BRISQUE [12]	36	0.5481	0.5287	0.3741
NFERM [13]	23	0.6808	0.6233	0.4516
NFERM-II (Pro.)	24	**0.6915**	**0.6298**	**0.4578**

Quality Metrics	Number of Features	*SRCC*		
		LIVE	CSIQ	LIVEMD
BRISQUE [12]	36	0.5567	0.4083	0.1930
NFERM [13]	23	0.7960	0.5643	0.1885
NFERM-II (Pro.)	24	**0.8062**	**0.6279**	**0.1947**

out a similar experiment about a distorted image compared to its original version. In such condition, 640 image pairs are included. From "Group 2" of Table III, we are able to derive the same conclusion validating the superiority of our blind NIQMC method.

The sensitivity of parameter γ in (8) is also compared. We enumerate a small range of values around the chosen value and compute the corresponding performance indices on five testing image databases, as reported in Table IV. One can see that the parameter γ is of good immunity and robustness to the changing of values. The proposed blind NIQMC algorithm is made up of two components, namely the local quality measure and the global one, and thus the contribution of each component deserves a quantified performance comparison. Results show that both two components provide significant contributions, but the their combination, NIQMC, has attained the optimal performance among the three.

Apart from predicting the visual quality of contrast-altered images, the proposed NIQMC approach is also applicable to another important application, advancing the performance of general-purpose NR-IQA measures. We in this paper choose to improve the state-of-the-art NFERM metric [13] with the NIQMC model, since on one hand this metric is of high performance and few features compared with other blind quality measures, and on the other hand this metric and the NIQMC have several components in common, e.g. preprocessing the distorted image via the AR model, and we might elaborately combine them together with only a few components.

We improve the NFERM metric through introducing the NIQMC score as a new feature, generating the NFERM-II with 24 features. The popular and state-of-the-art BRISQUE and NFERM are also included for comparison. On the TID2013 database, we randomly separate the overall 3,000 images into two groups. The first group consists of 2,400 images associated to 80% reference images, and the second group are composed of 600 images associated to the rest 20% reference images. To make sure that the IQA metrics are robust across image contents and not biased by the particular train-test split, this random 80% train - 20% test procedure is repeated with 1000 times. We report the median result of the 1,000 performance indices in Table V. It is easy to observe that introducing the new feature bring a remarkable improvement on the original NFERM metric, and moreover, it has delivered the optimal performance among the three.

A cross-validation is further conducted using the TID2013 database for training and using popular LIVE [37], CSIQ [23] and LIVEMD [38] databases for testing. According to the results, two main conclusions can be drawn. Firstly, the new feature, namely the NIQMC index, is proved once more of good ability to improve the NFERM metric with only one feature introduced. Secondly, despite the performance gain, the changes on LIVE and LIVEMD databases are quite small, which is possibly because no contrast-related distortions are included in these two databases. In addition, it can be also found that the correlation performance is not high, especially in the LIVEMD database. The problem, in my point of view, might be caused by the insufficiency of training samples. That is to say, the usage of only 24 image scenes in the TID2013 database for learning is easily leading to under-fitting and biased the results on other image scenes. To compensate for this deficiency, a collection of more than 1,000 various scenes will reduce this bias and provide a good-fitting.

IV. Conclusion

In this paper, we have investigated the problem of visual quality assessment of contrast-distorted images. Three main contributions have been made in this research. First, we have proposed a training-free blind quality method based on the concept of information maximization, which is much better than the classical FR-IQA models and lately devised NR-IQA metrics and equivalent to the state-of-the-art FR approaches. Second, compared with the majority of existing IQA models, our blind NIQMC technique can accurately judge which has larger contrast and better quality between two images. Third, the score of our NIQMC metric can be treated as an effective feature for improving the performance of the general-purpose blind NFERM method to a wide margin, even outperforming existing dominant and recently proposed general-purpose no-reference quality measures.

There might still be some room for improvement in performance. In the future work we will further study the IQA of contrast change from four aspects. The first one is to devise a nonlinear strategy to adaptively weight the local quality measure and the global one. Due to the fact that our NIQMC and the general-purpose NFERM are individually developed, the direct combination of them used in this paper cannot be the best way, and thus the second work is towards devising a more reasonable fusion manner. The third work is to consider the influence of the display devises on which the images are scored, because the gamma function could largely affect the image quality with different transfer curves. The last one is to better insert the important chromatic component into our blind NIQMC metric, akin to recent LTG and IL-NIQE.

REFERENCES

[1] W. Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image Represent.*, vol. 22, no. 4, pp. 297-312, May 2011.

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

[3] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-motivated rate distortion optimization for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 4, pp. 516-529, Apr. 2012.

[4] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-inspired divisive normalization for perceptual video coding," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1418-1429, Apr. 2013.

[5] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 43-54, Jan. 2013.

[6] Y. Fang, K. Zeng, Z. Wang, W. Lin, Z. Fang, and C.-W. Lin, "Objective quality assessment for image retargeting based on structural similarity," *IEEE J. Emerg. Sel. T. Circuits Syst.*, vol. 4, no. 1, pp. 95-105, Mar. 2014.

[7] K. Gu, G. Zhai, X. Yang, and W. Zhang, "An efficient color image quality metric with local-tuned-global model," *in Proc. IEEE Int. Conf. Image Process.*, pp. 506-510, Oct. 2014.

[8] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270-4281, Oct. 2014.

[9] L. Li, W. Xia, Y. Fang, K. Gu, J. Wu, W. Lin, and J. Qian, "Color image quality assessment based on sparse representation and reconstruction residual," *J. Vis. Commun. Image Represent.*, vol. 38, pp. 550-560, Jul. 2016.

[10] D. Tao, X. Li, W. Lu, and X. Gao, "Reduced-reference IQA in contourlet domain," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 6, pp. 1623-1726, Dec. 2009.

[11] X. Gao, W. Lu, D. Tao, and X. Li, "Image quality assessment based on multiscale geometric analysis," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1409-1423, Jul. 2009.

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

[13] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50-63, Jan. 2015.

[14] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. on Image Process.*, vol. 24, no. 8, pp. 2579-2591, Aug. 2015.

[15] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, W. Lin, W. Zhang, and W. Gao, "Blind quality assessment of tone-mapped images via analysis of information, naturalness and structure," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 432-443, Mar. 2016.

[16] K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, "Subjective and objective quality assessment for images with contrast change," *in Proc. IEEE Int. Conf. Image Process.*, pp. 383-387, Sept. 2013.

[17] K. Gu, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "Automatic contrast enhancement technology with saliency preservation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 9, pp. 1480-1494, Sept. 2015.

[18] Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai, "No-reference quality assessment of contrast-distorted images based on natural scene statistics," *IEEE Sig. Process. Lett.*, vol. 22, no. 7, pp. 838-842, Jul. 2015.

[19] S. Wang, K. Ma, H. Yeganeh, Z. Wang and W. Lin, "A patch-structure representation method for quality assessment of contrast changed images," *IEEE Sig. Process. Lett.*, vol. 22, no. 12, pp. 2387-2390, Dec. 2015.

[20] K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Trans. Cybernetics*, vol. 46, no. 1, pp. 284-297, Jan. 2016.

[21] K. Panetta, S. S. Agaian, Y. Zhou, and E. J. Wharton, "Parameterized logarithmic framework for image enhancement," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 2, pp. 460-473, Apr. 2011.

[22] S. Wang, K. Gu, S. Ma, W. Lin, X. Liu, and W. Gao, "Guided image contrast enhancement based on retrieved images in cloud," *IEEE Trans. Multimedia*, vol. 18, no. 2, pp. 219-232, Feb. 2016.

[23] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, Mar. 2010. [Online], Available: http://vision.okstate.edu/csiq

[24] N. Ponomarenko *et al.*, "TID2008-A database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, pp. 30-45, 2009.

[25] N. Ponomarenko *et al.*, "Image database TID2013: Peculiarities, results and perspectives," *Sig. Process.: Image Commun.*, vol. 30, pp. 57-55, Jan. 2015.

[26] N. Bruce and J. Tsotsos, "Saliency based on information maximization," *in Proc. Adv. Neural Inf. Process. Syst.*, pp. 155-162, 2005.

[27] X. Wu, G. Zhai, X. Yang, and W. Zhang, "Adaptive sequential prediction of multidimensional signals with applications to lossless image coding," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 36-42, Jan. 2011.

[28] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *in Proc. IEEE Int. Conf. Comput. Vision*, pp. 836-846, Jan. 1998.

[29] K. Friston, "The free-energy principle: A unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, pp. 127-138, 2010.

[30] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Visual saliency detection with free energy theory," *IEEE Sig. Process. Lett.*, vol. 22, no. 10, pp. 1552-1555, Oct. 2015.

[31] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot, "No-reference image blur assessment based on discrete orthogonal moments," *IEEE Trans. Cybernetics*, vol. 46, no. 1, pp. 39-50, Jan. 2016.

[32] L. Li, Y. Zhou, W. Lin, J. Wu, X. Zhang, and B. Chen, "No-reference quality assessment of deblocked images," *Neurocomputing*, vol. 177, pp. 572-584, Feb. 2016.

[33] K. Gu, S. Wang, H. Yang, W. Lin, G. Zhai, X. Yang, and W. Zhang, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1-13, Jun. 2016.

[34] D. H. Johnson and S. Sinanović, "Symmetrizing the Kullback-Leibler distance," *IEEE Trans. Information Theory*, 2001.

[35] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218-3231, Oct. 2015.

[36] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Mar. 2000, http://www.vqeg.org/.

[37] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment Database Release 2," [Online]. Available: http://live.ece.utexas.edu/research/quality

[38] D. Jayaraman, A. Mittal, A. K. Moorthy, and A. C. Bovik, "Objective quality assessment of multiply distorted images," *in Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, pp. 1693-1697, Nov. 2012.