# Perceptual Reduced-Reference Visual Quality Assessment for Contrast Alteration

Min Liu, Ke Gu, Guangtao Zhai, *Member, IEEE,* Patrick Le Callet, *Senior Member, IEEE,* and Wenjun Zhang, *Fellow, IEEE*

***Abstract*—In image/video systems, contrast adjustment which manages to enhance visual quality is nowadays an important research topic. Yet very limited struggles have been made to the exploration of visual quality assessment for contrast adjustment. To tackle the issue, this paper proposes a novel reduced-reference (RR) quality metric with the integration of bottom-up and top-down strategies. The former one stems from the recently revealed free energy principle that tells that the human visual system seeks to comprehend an input image via uncertainty removal, while the latter one is towards using the symmetric K-L divergence to compare the histogram of the contrast-altered image with that of the pristine image. The bottom-up and top-down strategies are lastly incorporated to derive the Reduced-reference Contrast-altered Image Quality Measure (RCIQM). A comparison using numerous existing IQA models is carried out on five contrast related databases / subsets in CID2013, CCID2014, CSIQ, TID2008 and TID2013, and experimental results validate the superiority of the proposed technique.**

***Index Terms*—Contrast alteration, quality assessment (QA), reduced-reference (RR), hybrid parametric and non-parametric model (HPNP), bottom-up, top-down**

## I. Introduction

During recent years, the importance of visual media with ubiquitous applications have become obvious. Images and videos in most conditions are provided to human consumers. As the users' requirements and expectations for high-quality images / videos are increasingly rising, a reliable system to evaluate, control and improve the users' quality of experience (QoE) is urgently required, e.g., in some elegant technologies of compression [1], [2], enhancement [3], tone mapping [4], etc. This results in the demand of effective metrics of image quality assessment (IQA) for predicting the quality in accordance with human visual perception.

IQA can be classified into subjective assessment and objective assessment. The first captures human ratings of visual quality, namely, mean opinion scores (MOSs). But it easily suffers the drawbacks of being time-consuming, expensive and unpractical, and this thereby leads to the study of objective assessment, which aims to evaluate the image quality using mathematical models to estimate subjective ratings.

This work was supported in part by the National Science Foundation of China under Grant 61527804 and Grant 61305011. (Corresponding Author: Dr. Ke Gu).

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

K. Gu is with School of Computer Science and Engineering, Nanyang Technological University, Singapore, 639798 (email: guke.doctor@gmail.com).

Patrick Le Callet is with the LUNAM Université, Université de Nantes, France (email: patrick.lecallet@univ-nantes.fr).

Referring to the accessability of original visual signals to be compared with during the computation, objective IQA metrics can be further divided into three types: 1) *full-reference* (FR) IQA; 2) *reduced-reference* (RR) IQA; 3) *no-reference* (NR) IQA. This paper mainly discusses the former two. The FR peak signal-to-noise ratio (PSNR) has prevailed for years, in view of its computational simplicity and clear physical meaning. However they do not constantly well correlate with human judgments of quality [5]. As a result, the last few years have witnessed the emergence of numerous FR IQA metrics [6], [7], [9], [10], [11], [12], [13], [14], [15], [16]. The most famous one is perhaps the structural similarity index (SSIM) [6], which compares luminance, contrast and structural similarities of the original and corrupted images. Thereafter, many modified SSIM-type of metrics have been devised, e.g. multi-scale SSIM (MS-SSIM) [7], optimal scale selection-based SSIM (OSS-SSIM) [8], and analysis of distortion distribution-based SSIM (ADD-SSIM) [9].

During the recent decade, a great amount of FR IQA algorithms were devised in other tactics. For instance, the most apparent distortion (MAD) [11] utilizes the detection- and appearance-based model to assess the visual quality. The feature similarity index (FSIM) [12] and gradient similarity index (GSI) [13] consider the fact that the perception of an image by the human visual system (HVS) primarily relies on classical low-level features - magnitude and phase.

Assuming that the partial original image or some extracted features is used as side information, RR-IQA is applicable to a wider range of practical scenarios [17], [18], [19], [20], [21]. Referring to the recent discovery of free energy theory [22], the free energy based distortion metric (FEDM) [17] was explored by mimicking the brain generative model to characterize the input image signals. Some metrics, e.g. structural degradation model (SDM) [18], manage to improve FR SSIM to be valid RR techniques with few numbers.

Nonetheless, very limited efforts have been devoted to the field of IQA with contrast change [23], [24]. In [24], a novel patch based FR-IQA method for contrast quality evaluation was proposed. Moreover, existing IQA algorithms do not work well in this field. As a matter of fact, contrast is an important research topic [25], which has practical applications in image/video systems such as contrast enhancement technologies [26], [27], [28]. This motivates the design of a new specialized contrast-changed image database (CID2013) [29], including 400 contrast-changed images by mean shifting and four kinds of transfer mappings, and its advanced version (CCID2014) [30].

In this paper we further explore the issue of contrast-adjusted IQA, and develop a new RR IQA model with the combination of bottom-up and top-down strategies. Relative to the frequently seen distortion types, e.g. JPEG / JPEG2000 compressions, the human visual sensation of image contrast (mainly including brightness and contrast alteration) inclines to the measurement in visual and psychological fields. A recently revealed free energy principle indicates that the HVS attempts to understand a visual signal through reducing the uncertainty and measures the psychovisual quality to be the agreement of an input image and its explanations derived by the internal generative model. With this, we evaluate the visual quality of contrast-altered images in the bottom-up model combining the generative model, which is constructed by the non-parametric autoregressive (AR) model via perceptual information for weighting.

On the other hand, as argued in several existing contrast enhancement methods [26], [27], the histogram modification can lead to the contrast adjustment and largely influence users' experiences. The top-down strategy is to compare two distances between histograms; one is between the contrast-adjusted image and its original version, and the other is between the contrast-altered image and the one created from the original image through histogram equalization. The Kullback-Leibler (K-L) divergence [31], one of the most popular information-theoretic "distances" comparing two probability distributions, is naturally taken into account. But the K-L divergence is non-symmetric and brings unstable results in calculation. So we use the symmetrized and smoothed Jensen-Shannon (JS) divergence [32] to compute the two distances stated above. Finally, the bottom-up and top-down strategies are combined to develop the Reduced-reference Contrast-changed Image Quality Measure (RCIQM), whose superiority is verified over existing visual quality evaluators.

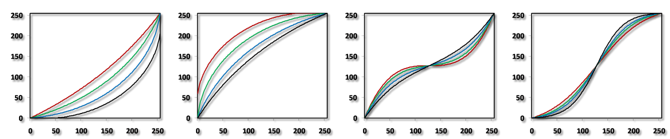
The remainder of this paper is arranged below. Section II first reviews existing contrast relevant image databases. In Section III, we construct the bottom-up and top-down models followed by combining them to put forward the RCIQM metric. Section IV conducts comparative studies of our measure with numerous existing FR- and RR-IQA methods on CID2013 [29], CDID2014 [30], CSIQ [11], TID2008 [33] and TID2013 [34] databases, and then reports and discusses the results of experiments. Section V concludes this paper.

(a) Mean-shifted images in TID2008.

(b) Contrast-changed images in TID2008.

Fig. 1: Sample images in the TID2008 database [33]: (a) average shifting; (b) contrast change.

Fig. 2: Typical contrast-changed images in the CSIQ database [11].



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

Fig. 3: Transfer curves in the CID2013 database [29]: (a) concave arcs; (b) convex arcs; (c) cubic functions; (d) logistic functions.

## II. Contrast Related Image Databases

The explorations of modern visual quality evaluation date from the beginning of this century, yet they mainly focused their attentions on the commonly seen compression, Gaussian blur and white noise, until the TID2008 database released. In that database, contrast related image subsets (mean shifting and contrast change) and associated MOS values were first open to the public. For a direct and clear understanding, we show a group of classical mean-shifted and contrast-changed images in Fig. 1. These images are obviously different from the aforesaid four distortion types. Soon after the emergence of TID2008, the CSIQ database consisting of contrast-altered images was also released, as given in Fig. 2. Instead of using the classical natural images in TID2008, which come from the Kodak database [35], CSIQ applies 30 new source images spanning a wider range of contents and scenes.

In practice, contrast-changed IQA metrics can be employed to instruct and optimize contrast enhancement schemes, which are of extreme significance in many cases for the purpose of increasing the image contrast and thus improving the image quality, even superior to the pristine version. None of existing IQA models, however, has satisfactory performance, as given in the experimental results later. Furthermore, contrast related images in TID2008, CSIQ and TID2013 are not numerous enough. To that end, we recently introduced a particular and challenging CID2013 database [29], which is composed of fifteen natural images from the Kodak database and 400 contrast-altered images and corresponding subjective opinion scores from inexperienced observers with different majors. Those images can be separated into two classes. The first class is produced with mean shifting the original image $I_o$ with positive or negative number ($+\triangle I$ or $-\triangle I$). The shift $\triangle I$ has six degrees of $\{20, 40, 60, 80, 100, 120\}$. The second class of contrast-changed images is generated using transfer mappings, including concave arc, convex arc, cubic function and logistic function. We present the transfer curves in Fig. 3 and several examples in Fig. 4. Later, we have also extended the CID2013 to a larger-scale CCID2014 database with 655 contrast-altered images [30].

(a) Mean-shifted images.

(b) Images transferred by concave arcs.

(c) Images transferred by convex arcs.

(d) Images transferred by cubic functions.

(e) Images transferred by logistic functions.

Fig. 4: Sample images in the CID2013 database [29].

## III. Proposed RCIQM Metric

We first list important notations and abbreviations in Table I for readers' conveniences to follow the subsequent context easily.

### A. Bottom-Up Strategy

It is widely argued that people would prefer a visual signal with balanced lighting and proper contrast. In comparison to typical distortion types, e.g. image / video coding, the HVS perception to image contrast, which is affected by luminance and contrast variation, is supposed to highly correlate with the visual and psychological measurement. We first establish the bottom-up model based on the free energy principle, which generates an approximating estimation of the psychovisual quality [17].

To specify, Friston recently revealed that free energy can explain and unify some existing brain principles in physical and biological sciences regarding learning, action and perception. The primary supposition behind it lies in that the perception process is managed by an internal generative model in brain, akin to the Bayesian-based brain assumption [36]. Using this model, the brain is able to construct a manner to actively infer the valuable information from input image signals and reduce the uncertain residual. This manner can be regarded as a probabilistic model, and we can decompose it into a likelihood term and a prior term. Inverting the first term, the HVS can deduce the posterior possibility of the given image. A gap naturally exists between the real external scene and the brain's estimation, due to the non-universal internal generative model. We reasonably suppose the difference of the external given image and its output generative-model-explainable part to be highly connected to the psychovisual quality, and even used to assess contrast-changed images.

TABLE I: Important notations and abbreviations.

| | |
|---|---|
| $x$ | Pixel index |
| $y$ | Pixel value |
| $\alpha$ | Parameter of AR model |
| $\beta$ | Parameter of Bi-lateral filtering |
| $\mathcal{Y}^k(y_i)$ | $k$ member neighborhood vector of $y_i$ |
| $\varepsilon_i, \varepsilon_i'$ | Error term |
| $\mu$ | Local mean |
| $\sigma$ | Local variance |
| $h_x$ | Spatial Euclidean distance |
| $h_y$ | Photometric distance |
| $l$ | Lumimance information |
| $c$ | Contrast information |
| $s$ | Structural information |
| $w, s, t$ | Weighting paramter |
| $H$ | Entropy of error |
| $p_0, p_1$ | probability density |
| $\mathcal{D}_{KL}$ | K-L divergence |
| $\mathcal{D}_{JS}$ | J-S divergence |
| $Q_{bu}$ | Quality measure of bottom-up strategy |
| $Q_{td}$ | Quality measure of top-down strategy |

Note that the free energy is an error difference map of the input image and its resulting optimal explanation estimated by the brain generative model. In this error difference map, higher value pixels represent hard explained areas, while lower value regions indicate easily described pixels with the generative model. The error difference map is acquired through making free energy minimized. On the basis of the analysis in [37], the free-energy minimization has a strong connection to the predictive coding. So we are capable of approximating it to be the entropy of the residuals of the input visual signal and its reconstructed one.

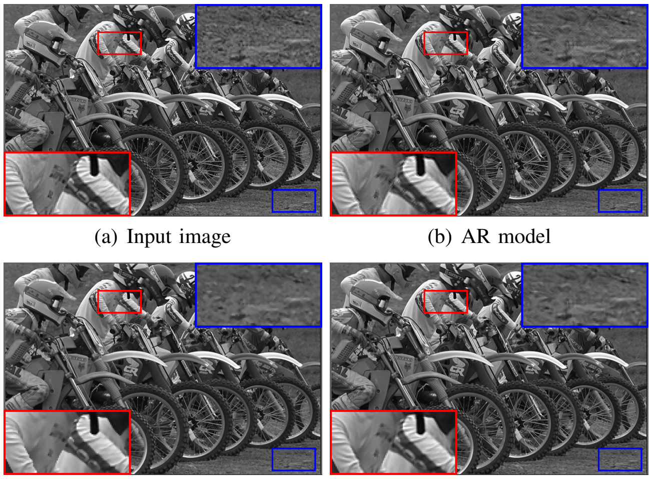
The internal generative model is defined to be a new hybrid parametric and non-parametric (HPNP) model, which fuses the linear AR model with the bi-lateral filtering. The first AR model is simple and it can simulate a broad range of natural scenes by varying its parameters [38], [39]. Particularly, the AR model is expressed by

$$y_i = \mathcal{Y}^k(y_i)\boldsymbol{\alpha} + \varepsilon_i \tag{1}$$

where $y_i$ indicates the pixel value located at $x_i$, $\mathcal{Y}^k(y_i)$ provides the vector for $k$ member neighborhood of $y_i$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_k)^T$ stands for an AR coefficients' vector, and $\varepsilon_i$ represents a difference between the truthes and estimations. To decide $\boldsymbol{\alpha}$, the linear system can represent in matrix form as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{Y}\boldsymbol{\alpha}\|_2 \tag{2}$$

with $\mathbf{y} = (y_1, y_2, ..., y_k)^T$ and $\mathbf{Y}(i,:) = \mathcal{Y}^k(y_i)$. A simple way to solve this linear system is using the least square method and derive the approximate solution as $\hat{\boldsymbol{\alpha}} = (\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T\mathbf{y}$.



(a) Input image (b) AR model

(c) Bi-lateral filtering (d) HPNP model

Fig. 5: The sample image "motocross bikes" and the associated filtered images using AR model, bilateral filtering, and HPNP model.

We provide an example in Fig. 5 to straightforwardly illustrate the visualized effect of the AR model. One can see that, the AR model performs well at the texture regions (indicated by blue rectangle), while it may lead to instability at edges (indicated by red rectangle). So we further take advantage of the bi-lateral filtering [40], which is a non-linear filtering of good ability to preserve edges and calculate simply [41]. Also, the bi-lateral filtering has only two variables $(h_x, h_y)$, rendering it convenient to control. We define this filtering by

$$y_i = \mathcal{Y}^k(y_i)\boldsymbol{\beta} + \varepsilon'_i \tag{3}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_k)^T$ indicates the coefficients' vector of bi-lateral filtering, and $\varepsilon'_i$ provides the error difference. The $\boldsymbol{\beta}$ is manipulated by the spatial Euclidean distance between $x_i$ and $x_j$ as well as the photometric distance between $y_i$ and $y_j$. It can be estimated by

$$\beta_j = \exp\Big(\frac{-\|x_i - x_j\|^2}{2h_x^2}\Big) \exp\Big(\frac{-(y_i - y_j)^2}{2h_y^2}\Big) \tag{4}$$

where $h_x$ and $h_y$ are assigned as 3 and 0.1 (default values) in each local $3 \times 3$ part, to alter the relative importance of the Euclidean and photometric distances. As given in Figs. 5(b)-(c), the bi-lateral filtering usually performs better than the AR model at image edges.

In the following, the HPNP model combines the merits of both parametric AR model and non-parametric bi-lateral filtering, and thus generate the estimation of $\bar{y}_i$ to be

$$\bar{y}_i = \gamma \cdot \mathcal{Y}^k(y_i)\hat{\boldsymbol{\alpha}} + (1-\gamma) \cdot \mathcal{Y}^k(y_i)\boldsymbol{\beta} \tag{5}$$

where $\gamma$ is used for adjusting the relative contribution of the AR model and the bi-lateral filtering. The value of $\gamma$ is determined based on the criterion making texture and edge regions to be preserved as well as possible, and is assigned to be 0.3 in this research, as displayed in Fig. 5(d).

In general, salient regions attract much attention and thus highly influence the visual quality. Some technologies can be used to detect visual saliency, for example, saliency detection models [42] and phase congruency [43], [44]. But for images with mean shifting or contrast change, luminance and contrast information should be more decisive factors. Apart from these two, the structural information is also considered since it may destroy the completeness of objects and thus heavily affects the perceptual quality of the HVS to a given scene. As thus, this paper incorporates the luminance, contrast and structural information for weighting. To be more concretely, for the pixel located at $x_i$ having the value $y_i$, we measure the luminance, contrast and structural information by

$$l(y_i, \bar{y}_i) = \frac{2\mu\bar{\mu} + c_1}{\mu^2 + \bar{\mu}^2 + c_1} \tag{6}$$

$$c(y_i, \bar{y}_i) = \frac{2\sigma\bar{\sigma} + c_2}{\sigma^2 + \bar{\sigma}^2 + c_2} \tag{7}$$

$$s(y_i, \bar{y}_i) = \frac{\tilde{\sigma} + c_3}{\sigma\bar{\sigma} + c_3} \tag{8}$$

where $c_1, c_2$ and $c_3$ are low-value fixed numbers for alleviating instability when denominators are close to zero. A $11 \times 11$ circular-symmetric Gaussian weighting function $\mathbf{v} = \{v_i | i = 1, 2, ..., N\}$ is used with 1.5 standard deviation and normalized to unit sum ($\sum_{i=1}^N v_i = 1$). The statistics $\mu$, $\bar{\mu}$, $\sigma^2$, $\bar{\sigma}^2$ and $\tilde{\sigma}$ are estimated using the same way in [6]. So the weight is defined to be

$$w_i = l(y_i, \bar{y}_i) \cdot c(y_i, \bar{y}_i) \cdot s(y_i, \bar{y}_i), \tag{9}$$

and the estimation error of the difference of the truth scene and brain's estimations for the local pixel at $x_i$ is evaluated by

$$\bar{e}_i = w_i(y_i - \bar{y}_i). \tag{10}$$

I have also tried to use saliency and region of interest as weighting, e.g., some state-of-the-art models [45], [46], [47]. The results indicate that these models contribute no greater performance than the weighting model mentioned above, but always introduce more computational cost.

For the original image $I_o$, the point-wise error $\bar{e}_i$ can be computed using Eqs. (1)-(10) to get the error map $E_o$. The free energy of this error map is measured by entropy:

$$H(E_o) = -\sum p_i(E_o) \log p_i(E_o) \tag{11}$$

where $p_i(E_o)$ is the probability density of grayscale $i$ in the error map $E_o$. With the same manner, we measure entropy of $H(E_c)$ for the contrast-changed image $I_c$. The psychovisual quality of $I_c$ compared to $I_o$ within the bottom-up strategy is finally defined as their difference:

$$Q_{bu} = H(E_o) - H(E_c). \tag{12}$$

In practice, the kernel of the bottom-up strategy lies in the HPNP model for approximating the internal generative model in human brain. Images with high contrast and visual quality generally have an abundant number of valuable details. Our HPNP model is of different description abilities between low- and high-complexity visual signals. For a fixed input image with its free energy $H(E_o)$, the positive contrast change will increase the visual quality by revealing undiscernible details. This renders the designed HPNP model inefficient to characterize the contrast-altered image, and thus makes its free energy $H(E_c)$ higher than $H(E_o)$ and $Q_{bu}$ lower than zero. On the contrary, the negative contrast change will decrease the visual quality by concealing details, which leads to the associated free energy $H(E_c)$ smaller than $H(E_o)$ and $Q_{bu}$ larger than zero.

### B. Top-Down Strategy

One of important applications related to contrast alteration is the familiar contrast enhancement technology, which can be regarded as the positive contrast change for validly advancing the contrast and raising the visual quality of an input image. Broadly speaking, contrast enhancement targets to create a more informative or visually-pleasing image or both. Observers often regard enhanced images as removing a curtain of fog from a picture [26]. For example, as shown in Fig. 4(e), the four images processed by logistic transfers seem to thin the fog from the associated source image and attain a certain improvement of visual quality.

In existing contrast enhancement technologies, histogram equalization (HE) is possibly the easiest and widely used solution, which aims to increase the image informativeness. Its fundamental idea is to redistribute values of pixels in terms of the probability distribution of gray levels of the input image, for the purpose of flattening and stretching the dynamic range of image histogram and give an global contrast improvement. Yet the HE is very likely to introduce artifacts, cause a considerable change on the mean brightness, and produce undesirable visual deterioration. So this technique is presently considered to be far less than the ideal contrast enhancement algorithm.

In the last several years, some solutions were designed to modify HE, in order to overcome those drawbacks [26], [27]. In [26], the authors suggested that the enhanced image should be not far from its original image, and they provided a compromise scheme. Instead of using the uniformly distributed histogram $\mathbf{h_u}$ as the target histogram, their goal is to find a modified histogram $\tilde{\mathbf{h}}$ that is near to $\mathbf{h_u}$ as desired, but also not far from the original image histogram $\mathbf{h_o}$. This is a bi-criteria optimization problem, and can be formulated as a weighted sum of the two objectives:

$$\tilde{\mathbf{h}} = \arg\min_{\mathbf{h}} \|\mathbf{h} - \mathbf{h_o}\| + \phi\|\mathbf{h} - \mathbf{h_u}\| \tag{13}$$

where $\tilde{\mathbf{h}}$, $\mathbf{h}$, $\mathbf{h_o}$, $\mathbf{h_u} \in \mathbb{R}^{256\times1}$, and $\phi$ is a control parameter varying over $[0, \infty)$.

Inspired by Eq. (13), we in the top-down strategy concentrate on measuring two distances; one is between the histogram $\mathbf{h_c}$ of the contrast-changed image and $\mathbf{h_o}$, and the other is between $\mathbf{h_c}$ and $\mathbf{h_u}$. Nonetheless, the construction of the top-down model is not straightforward. Firstly, we find that $\mathbf{h_u}$ is not a good choice, because the histograms of most images cannot be distributed uniformly after HE due to various image contents or scenes. Instead, this paper applies the equalized histogram $\mathbf{h_e}$ that is produced from $\mathbf{h_o}$ using HE. Secondly, it is important to note that the free energy in the bottom-up strategy is measured by entropy, so we had better evaluate the aforesaid two distances with the same dimension for the combination of bottom-up and top-down models to predict the visual quality score of the contrast-adjusted image. The K-L divergence [31], probably the most frequently used "distance" that compares the distinction between two probability distributions in probability theory and information theory, is of the expected dimension. Given two probability densities $p_0$ and $p_1$, the K-L divergence is defined as

$$\mathcal{D}_{KL}(p_1\|p_0) = \int p_1(x)\log\frac{p_1(x)}{p_0(x)}dx. \tag{14}$$

This K-L divergence is however non-symmetric and easy to bring some troubles in real applications. Simple examples illustrate that the ordering of the arguments in the K-L distance might yield substantially different results. We resort to the symmetric K-L divergence accordingly. In [32], the authors have summarized many symmetric forms of K-L divergence, e.g. algebraic mean and geometric mean. Here we consider using the symmetrized and smoothed Jensen-Shannon (JS) divergence as follows:

$$\mathcal{D}_{JS}(p_0, p_1) = \frac{1}{2}\mathcal{D}_{KL}(p_0\|\bar{p}) + \frac{1}{2}\mathcal{D}_{KL}(p_1\|\bar{p}) \tag{15}$$

with

$$\bar{p} = \frac{1}{2}(p_0 + p_1). \tag{16}$$

Except for the J-S divergence, there are many commonly used alternative methods, including Earth Mover's Distance [48], histogram intersections [49], [50], and $L$ norms ($L = 1, 2, \infty$). It was found by performance comparisons that the J-S divergence performs the best in this application scenario, Earth Mover's Distance and histogram intersections perform fairly, and three typical $L$ norms perform poorly. As thus, we finally determined to employ the J-S divergence.

As a consequence, given three probability densities $p_o$, $p_e$ and $p_c$ for an original image and its HE and contrast-altered counterparts, the quality of $I_c$ compared to $I_o$ within the top-down part is determined by

$$Q_{td} = \mathcal{D}_{JS}(p_c, p_o) + s\mathcal{D}_{JS}(p_c, p_e) \tag{17}$$

where $s$ is a fixed weighting parameter for altering the relative importance between the above two distances. The value of $s$ is empirically assigned as 2. More discussion regarding its sensitivity will be provided in the next section. The analyses in the histogram modification method point out that proper-contrast images should be a good tradeoff between the original image histogram and the uniformly distributed one. Our top-down model is properly developed for this, and it can thereby judge the quality levels of contrast-changed images.

### C. The Combination Stage

Popular contrast enhancement technologies are devoted to highlighting undiscernible details [28] or redistributing image histogram [26], [27]. Given an image, the former bottom-up model aims to estimate how much detailed information is contained, while the latter top-down model is to measure whether the histogram is properly distributed. From the viewpoint of working, these two models play complementary roles. As thus, we fuse bottom-up and top-down strategies to approximate the HVS perception to the contrast-altered image quality. Since

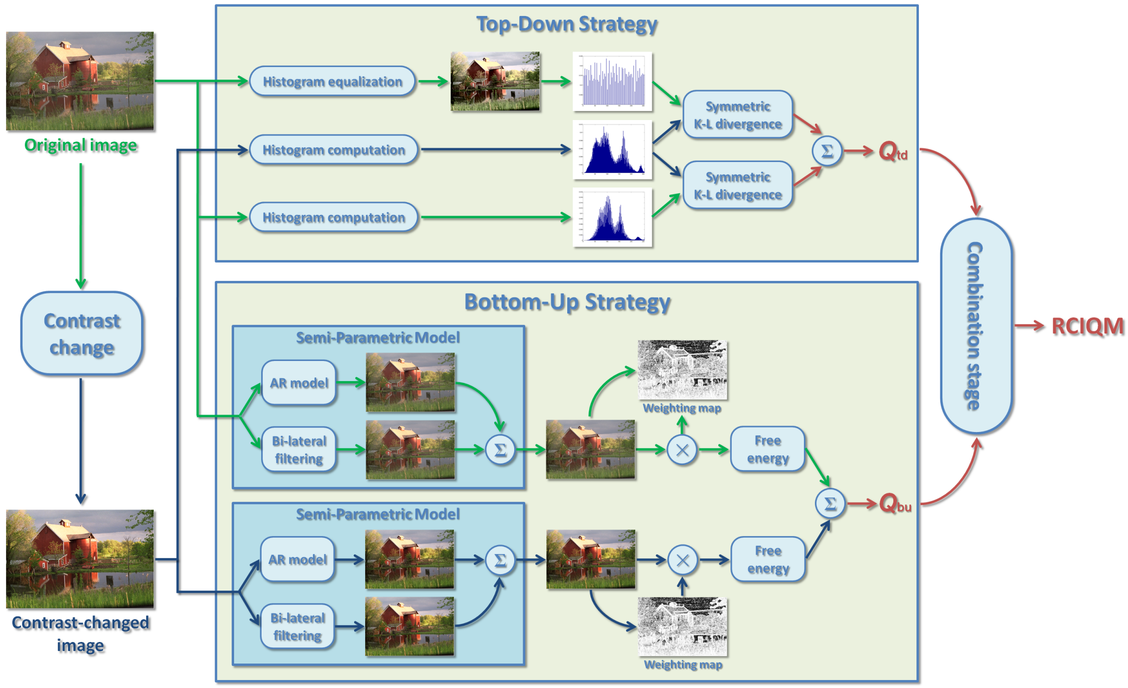



Fig. 6: The flowchart of the proposed RCIQM algorithm.

the quality measures based on the two models are of the same dimension (i.e. entropy) in our research, they can be directly integrated. The RCIQM is finally defined to be a simply linear function combining the two quality predictions in bottom-up and top-down parts:

$$\text{RCIQM} = Q_{bu} + tQ_{td} \tag{18}$$

where $t$ is a constant weight that is used to control the relative significance between the bottom-up and top-down strategies. The value of $t$ is empirically assigned to be 0.3. We will analyze its sensitivity in the next section. The parameters $s$ and $t$ are determined by making the proposed RCIQM metric have the optimal correlation with human judgements of quality using the CCID2014 database. All the parameters used in the proposed RCIQM model have fixed values. We present the flowchart in Fig. 6 for helping readers to readily understand how the RCIQM metric works.

Furthermore, we want to discuss why the proposed RCIQM is a RR IQA metric. In the bottom-up model, the RR feature only includes one single number of the free energy $H(E_o)$, and two histograms $\mathbf{h_o}$ and $\mathbf{h_e}$ are required to transmitted as the ancillary information in the top-down part. In reality, $\mathbf{h_e}$ is the output of the equalized $\mathbf{h_o}$. So the RR information used in RCIQM just includes $H(E_o)$ and $\mathbf{h_o}$ (totally 257 numbers), far less than the size of the original image. Besides, a supplementary specification is that, according to the convention, this paper utilizes different signs (e.g. $p_o$ and $\mathbf{h_o}$, $p_e$ and $\mathbf{h_e}$) but with the same meaning.

## IV. Experimental Results and Discussion

### A. Testing Metrics and Databases

In this paper, we validate the proposed RCIQM algorithm and compare with an enormous number of classical and state-of-the-art IQA metrics: 1) Classical FR IQA: PSNR, SSIM [6], and MS-SSIM [7], which mainly focus on the visual quality evaluation for frequently encountered JPEG / JPEG2000 compression, Gaussian blur, and white noise; 2) State-of-the-art FR IQA: FSIM [12], GSI [13], local-tuned-global model (LTG) [14], and visual saliency induced index (VSI) [15], which were recently designed to cope with a broad range of distortion types that, apart from the above typical types, also include quantization noise, non eccentricity pattern noise, and etc; 3) Recently devised RR IQA: FEDM [17], SDM [18] and RIQMC [30], which assume that partial original references can be made available as side information to help to predict quality of the distorted image.

Currently, subjective image quality databases that have been released to public mainly involve compression, blur and noise distortions. To our knowledge, there only exist five contrast related image databases / subsets (CID2013, CCID2014, TID2008, TID2013 and CSIQ), which were selected as the testing bed in this work. The CID2013 database [29], which was proposed in our recent work. It has totally 400 images produced from 15 reference images with mean shifting and four kinds of transfer mappings. The MOS of each image is available and ranges from 1.4 to 4.2. The CCID2014 database [30], which supplements three new distortion types – positive and negative gamma transfers and compound functions with mean-shiftings followed by logistic functions – to the CID2013 database. This database encompasses 655 contrast-altered images, whose MOS values range from 1.4 to 4.4.

The TID2008 database [33], which involves 200 mean-shifted and contrast-adjusted images derived from 25 reference images (24 natural images and one artificial image) at four levels of distortions. The MOS of each image is from 3.4 to 7.7. The TID2013 database [34], which extends the original

TABLE II: Performance indices on TID2008, CSIQ, TID2013 and CID2013. We bold the top two models.

| Metrics | Type | CID2013 database (400 images) [29] | | | | | CCID2014 database (655 images) [30] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PLCC | SRCC | KRCC | AAE | RMS | PLCC | SRCC | KRCC | AAE | RMS |
| PSNR | FR | 0.6503 | 0.6649 | 0.4847 | 0.3787 | 0.4734 | 0.6832 | 0.6743 | 0.4834 | 0.3610 | 0.4775 |
| SSIM [6] | FR | 0.8119 | 0.8132 | 0.6140 | 0.2743 | 0.3638 | 0.8256 | 0.8136 | 0.6063 | 0.2900 | 0.3689 |
| MS-SSIM [7] | FR | 0.8543 | 0.8554 | 0.6593 | 0.2462 | 0.3239 | **0.8458** | **0.8271** | **0.6236** | **0.2757** | **0.3488** |
| FSIM [12] | FR | 0.8574 | 0.8486 | 0.6663 | 0.2460 | 0.3207 | 0.8183 | 0.7654 | 0.5705 | 0.3023 | 0.3758 |
| GSI [13] | FR | 0.8353 | 0.8372 | 0.6371 | 0.2677 | 0.3426 | 0.8073 | 0.7768 | 0.5711 | 0.3093 | 0.3859 |
| LTG [14] | FR | **0.8656** | **0.8605** | **0.6723** | **0.2432** | **0.3120** | 0.8384 | 0.7901 | 0.5938 | 0.2893 | 0.3564 |
| VSI [15] | FR | 0.8571 | 0.8501 | 0.6567 | 0.2518 | 0.3210 | 0.8209 | 0.7734 | 0.5735 | 0.3013 | 0.3734 |
| FEDM [17] | RR | 0.7533 | 0.7271 | 0.5604 | 0.3100 | 0.4098 | 0.6717 | 0.5729 | 0.4073 | 0.3973 | 0.4844 |
| SDM [18] | RR | 0.7158 | 0.6145 | 0.4363 | 0.3584 | 0.4352 | 0.7360 | 0.6733 | 0.4862 | 0.3643 | 0.4426 |
| RIQMC [30] | RR | **0.8995** | **0.9005** | **0.7162** | **0.2211** | **0.2723** | **0.8726** | **0.8465** | **0.6507** | **0.2610** | **0.3194** |
| RCIQM (Pro.) | RR | **0.9187** | **0.9203** | **0.7543** | **0.1949** | **0.2461** | **0.8845** | **0.8565** | **0.6695** | **0.2455** | **0.3051** |

| Metrics | Type | TID2008 database (200 images) [33] | | | | | TID2013 database (250 images) [34] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PLCC | SRCC | KRCC | AAE | RMS | PLCC | SRCC | KRCC | AAE | RMS |
| PSNR | FR | 0.5131 | 0.5207 | 0.3640 | 0.6637 | 0.8258 | 0.5071 | 0.5425 | 0.3630 | 0.6710 | 0.8454 |
| SSIM [6] | FR | 0.5057 | 0.4877 | 0.3402 | 0.6704 | 0.8300 | 0.5658 | 0.4905 | 0.3432 | 0.6387 | 0.8087 |
| MS-SSIM [7] | FR | 0.6654 | 0.5877 | 0.4303 | 0.5842 | 0.7182 | 0.6476 | **0.5450** | 0.4012 | 0.5899 | 0.7474 |
| FSIM [12] | FR | 0.6458 | 0.4388 | 0.3331 | 0.5989 | 0.7346 | 0.6578 | 0.4398 | 0.3572 | 0.5649 | 0.7388 |
| GSI [13] | FR | 0.6739 | 0.5126 | 0.3946 | 0.5804 | 0.7108 | 0.6665 | 0.4985 | **0.4024** | **0.5630** | 0.7312 |
| LTG [14] | FR | 0.6795 | 0.4655 | 0.3285 | 0.5759 | 0.7059 | 0.6749 | 0.4639 | 0.3458 | 0.5769 | 0.7237 |
| VSI [15] | FR | 0.6312 | 0.4571 | 0.3450 | 0.6066 | 0.7462 | **0.6785** | 0.4643 | 0.3705 | 0.5734 | **0.7205** |
| FEDM [17] | RR | 0.6594 | 0.3228 | 0.2057 | 0.5899 | 0.7233 | 0.6504 | 0.3217 | 0.2373 | 0.5954 | 0.7451 |
| SDM [18] | RR | **0.7817** | **0.7378** | **0.5456** | **0.4761** | **0.6001** | 0.5831 | 0.3482 | 0.2389 | 0.6331 | 0.7968 |
| RIQMC [30] | RR | **0.8585** | **0.8095** | **0.6224** | **0.3807** | **0.4933** | **0.8651** | **0.8044** | **0.6178** | **0.3746** | **0.4920** |
| RCIQM (Pro.) | RR | **0.8807** | **0.8578** | **0.6705** | **0.3617** | **0.4556** | **0.8866** | **0.8541** | **0.6675** | **0.3560** | **0.4537** |

| Metrics | Type | CSIQ database (116 images) [11] | | | | | Database size-weighted average | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PLCC | SRCC | KRCC | AAE | RMS | PLCC | SRCC | KRCC | AAE | RMS |
| PSNR | FR | 0.9002 | 0.8621 | 0.6449 | 0.0548 | 0.0733 | 0.6425 | 0.6462 | 0.4620 | 0.4286 | 0.5473 |
| SSIM [6] | FR | 0.7450 | 0.7397 | 0.5323 | 0.0855 | 0.1124 | 0.7369 | 0.7182 | 0.5295 | 0.3722 | 0.4740 |
| MS-SSIM [7] | FR | 0.8959 | 0.8833 | 0.6899 | 0.0592 | 0.0748 | 0.7987 | **0.7651** | **0.5790** | **0.3395** | 0.4301 |
| FSIM [12] | FR | 0.9435 | 0.9421 | 0.7889 | 0.0434 | 0.0558 | 0.7909 | 0.7081 | 0.5476 | 0.3470 | 0.4395 |
| GSI [13] | FR | 0.9325 | 0.9354 | 0.7721 | 0.0462 | 0.0608 | 0.7850 | 0.7275 | 0.5540 | 0.3528 | 0.4453 |
| LTG [14] | FR | 0.9560 | 0.9414 | 0.7880 | 0.0392 | 0.0494 | **0.8087** | 0.7279 | 0.5561 | 0.3397 | **0.4232** |
| VSI [15] | FR | 0.9533 | 0.9504 | 0.8096 | 0.0398 | 0.0509 | 0.7939 | 0.7183 | 0.5514 | 0.3500 | 0.4369 |
| FEDM [17] | RR | **0.9617** | **0.9550** | **0.8189** | **0.0359** | **0.0462** | 0.7078 | 0.5687 | 0.4234 | 0.4042 | 0.5043 |
| SDM [18] | RR | 0.9175 | 0.9141 | 0.7445 | 0.0521 | 0.0670 | 0.7261 | 0.6338 | 0.4616 | 0.3958 | 0.4880 |
| RIQMC [30] | RR | **0.9652** | **0.9579** | **0.8279** | **0.0343** | **0.0441** | **0.8829** | **0.8567** | **0.6710** | **0.2672** | **0.3362** |
| RCIQM (Pro.) | RR | **0.9645** | **0.9569** | **0.8198** | **0.0353** | **0.0445** | **0.8985** | **0.8792** | **0.7010** | **0.2494** | **0.3134** |

four distortion levels in TID2008 to five, generates a total number of 250 mean-shifted and contrast-altered versions. The MOS of each image ranges from 2.6 to 7.2. The CSIQ database [11], which consists of 116 images of contrast change that are created from 30 source images at three to four degradation levels. The DMOS of each image is available and ranges from 0 to 0.7.

### *B. Evaluation Protocols*

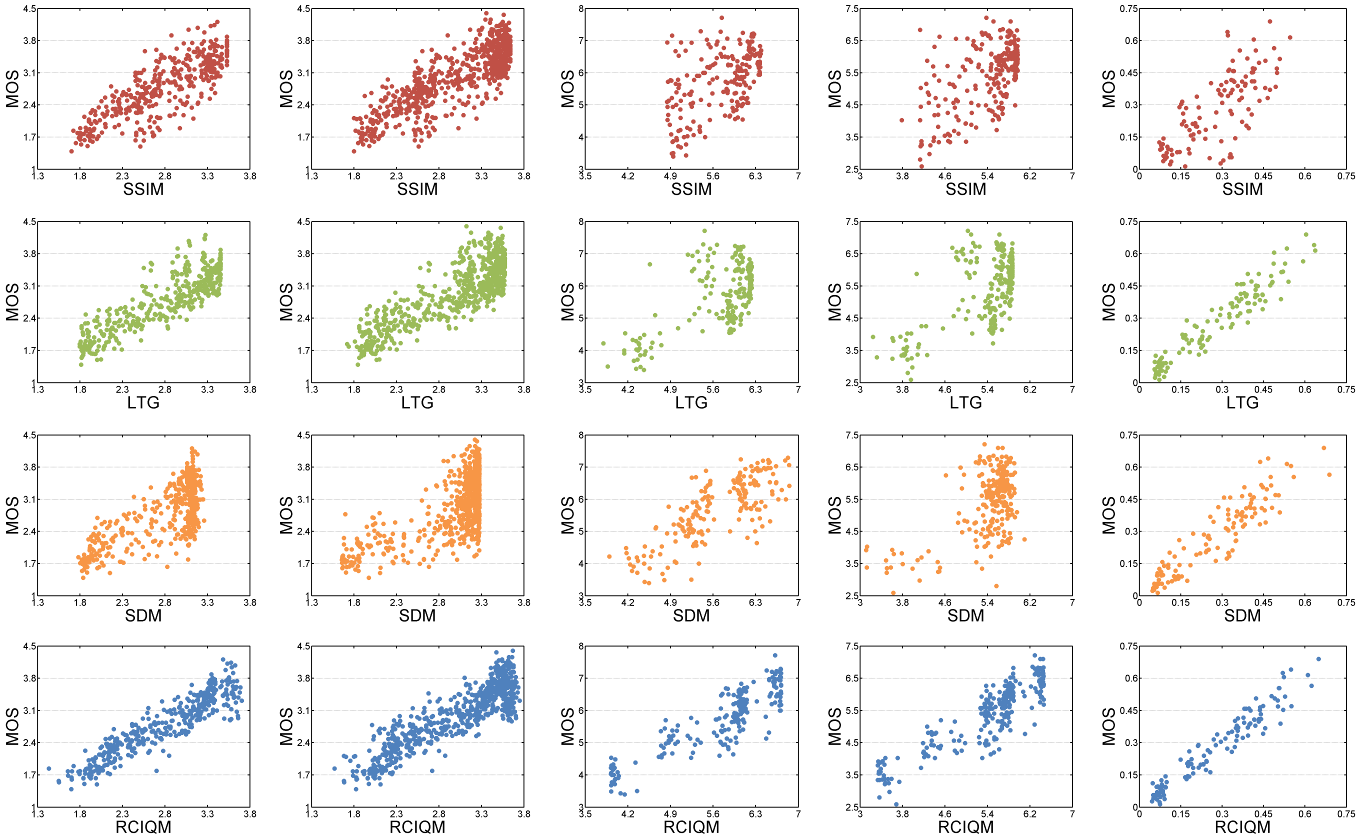
Using the five image databases / subsets, the objective prediction estimations of each competing IQA models is computed after conducting the nonlinear regression to map quality scores to human ratings with the five-parameter logistic regression function [51]:

$$q(\varepsilon) = \phi_1\left(\frac{1}{2} - \frac{1}{1 + e^{\phi_2(\varepsilon-\phi_3)}}\right) + \phi_4\varepsilon + \phi_5 \qquad (19)$$

where $\varepsilon$ and $q(\varepsilon)$ respectively indicate the raw and mapped scores, and $\phi_j$ $(j = 1, \ldots, 5)$ are free parameters to be ascertained. Five typical performance indices, according to suggestions given by the Video Quality Experts Group (VQEG) [51], are used to evaluate and compare the proposed RCIQM with the competing IQA models tested in this study. The first index Pearson linear correlation coefficient (PLCC) is measured between MOS scores and objective evaluations after nonlinear regression by Eq. (19), in order to measure the prediction accuracy. The second and third indices Spearman rank-order correlation coefficient (SRCC) and Kendall's rank-order correlation coefficient (KRCC) are to compute the monotonicity by ignoring the relative distance between the data, which are independent of any monotonic nonlinear mapping between subjective and objective quality scores. The last two indices average absolute prediction error (AAE) and root mean-squared (RMS) error are to quantify the difference of the subjective quality scores and converted objective IQA predictions after the nonlinear mapping of Eq. (19). Among these evaluations, a value close to 1 for PLCC, SRCC and KRCC, and near to 0 for AAE and RMS means superior correlation with subjective opinions.

### *C. Performance Evaluation*

Table II presents the performance indices of the proposed RCIQM and other ten IQA methods on CID2013, CCID2014, TID2008, TID2013 and CSIQ databases. For comprehensive comparisons, we further calculate the average results across




(a) CID2013 (b) CCID2014 (c) TID2008 (d) TID2013 (e) CSIQ

Fig. 7: Scatter plots of MOS / DMOS versus classical FR SSIM, recent FR LTG, RR SDM, and our RR RCIQM models on five databases.

the five databases above, which is defined by

$$\bar{\delta} = \frac{\sum_i \delta_i \cdot \pi_i}{\sum_i \pi_i} \tag{20}$$

where $\delta_i$ $(i = 1, 2, 3, 4, 5)$ stands for the correlation measure for each database. For the database size-weighted average, $\pi_i$ are set as the number of images in each database, i.e. 400 for CID2013, 655 for CCID2014, 200 for TID2008, 250 for TID2013, and 116 for CSIQ. Table II also reports the database size-weighted average results.

Referring to the nature of our proposed metric and the results listed in Table II, we give three conclusions. First, it is apparent that our metric achieves very promising result on each database and the weighted average. We note that only the proposed RCIQM technique has acquired the SRCC values larger than 0.92 on the CID2013 database, and greater than 0.85 on the large-scale CCID2014, TID2008 and TID2013 databases. Although a few IQA models (e.g. FEDM and RIQMC) work well in the CSIQ database, our RCIQM is also of the highest performance, even higher than 0.95 in linearity (Pearson) and monotonic measure (Spearman).

Second, as compared to FR- and RR-IQA algorithms tested in this paper, it can be readily viewed that the proposed RCIQM is of the optimal performance on average, clearly better than the second-place RIQMC and the third-place LTG methods. In fact, almost all FR and RR IQA methods assume that the reference image is prefect. But there exist contrast-changed images produced by the positive contrast alteration have better quality than their original ones, and this leads to serious deterioration in the performance of FR and RR IQA techniques when assessing contrast-altered images.

TABLE III: Statistical significance comparison with the F-test.

| | PSNR | SSIM | MS-SSIM | FSIM | GSI |
|---|---|---|---|---|---|
| CID2013 | +1 | +1 | +1 | +1 | +1 |
| CCID2014 | +1 | +1 | +1 | +1 | +1 |
| TID2008 | +1 | +1 | +1 | +1 | +1 |
| CSIQ | +1 | +1 | +1 | +1 | +1 |
| TID2013 | +1 | +1 | +1 | +1 | +1 |

| | LTG | VSI | FEDM | SDM | RIQMC |
|---|---|---|---|---|---|
| CID2013 | +1 | +1 | +1 | +1 | +1 |
| CCID2014 | +1 | +1 | +1 | +1 | +1 |
| TID2008 | +1 | +1 | +1 | +1 | +1 |
| CSIQ | +1 | +1 | 0 | +1 | 0 |
| TID2013 | +1 | +1 | +1 | +1 | +1 |

Third, our algorithm has many advantages over those in existing literature. For example, our RCIQM is insensitive to small translations and rotations that hardly influence the image quality or semantics. This phenomenon is mainly because the components of our metric, including the global image histogram and free energy entropy, are almost unchanged during small translations and rotations. Conversely, most existing IQA metrics predict severe visual quality degradation under these situations, since they compare the original and distorted images in a point-wise or block-based manner.

TABLE IV: Performance comparison of various amounts of RR information. We bold the top two metrics.

| Metrics | Info. number | CID2013 database (400 images) [29] | | | | | CCID2014 database (655 images) [30] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PLCC | SRCC | KRCC | AAE | RMS | PLCC | SRCC | KRCC | AAE | RMS |
| LTG [14] | whole | 0.8656 | 0.8605 | 0.6723 | 0.2432 | 0.3120 | 0.8384 | 0.7901 | 0.5938 | 0.2893 | 0.3564 |
| $RCIQM_1$ | 257 | **0.9187** | **0.9203** | **0.7543** | **0.1949** | **0.2461** | **0.8845** | **0.8565** | **0.6695** | **0.2455** | **0.3051** |
| $RCIQM_2$ | 129 | **0.9181** | **0.9195** | **0.7533** | **0.1951** | **0.2470** | **0.8839** | **0.8562** | **0.6690** | **0.2458** | **0.3058** |
| $RCIQM_4$ | 65 | 0.9157 | 0.9163 | 0.7477 | 0.1977 | 0.2505 | 0.8825 | 0.8552 | 0.6675 | 0.2478 | 0.3076 |
| $RCIQM_8$ | 33 | 0.9173 | 0.9183 | 0.7511 | 0.1965 | 0.2482 | 0.8828 | 0.8556 | 0.6683 | 0.2470 | 0.3072 |
| $RCIQM_{16}$ | 17 | 0.9172 | 0.9186 | 0.7519 | 0.1964 | 0.2483 | 0.8829 | 0.8557 | 0.6686 | 0.2468 | 0.3070 |
| $RCIQM_{32}$ | 9 | 0.9165 | 0.9177 | 0.7504 | 0.1975 | 0.2493 | 0.8823 | 0.8548 | 0.6672 | 0.2476 | 0.3078 |
| $RCIQM_{64}$ | 5 | 0.9162 | 0.9180 | 0.7505 | 0.1978 | 0.2497 | 0.8813 | 0.8543 | 0.6666 | 0.2487 | 0.3090 |

| Metrics | Info. number | TID2008 database (200 images) [33] | | | | | TID2013 database (250 images) [34] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PLCC | SRCC | KRCC | AAE | RMS | PLCC | SRCC | KRCC | AAE | RMS |
| LTG [14] | whole | 0.6795 | 0.4655 | 0.3285 | 0.5759 | 0.7059 | 0.6749 | 0.4639 | 0.3458 | 0.5769 | 0.7237 |
| $RCIQM_1$ | 257 | 0.8807 | **0.8578** | **0.6705** | 0.3617 | 0.4556 | **0.8866** | **0.8541** | **0.6675** | **0.3560** | **0.4537** |
| $RCIQM_2$ | 129 | **0.8811** | 0.8450 | 0.6651 | **0.3501** | **0.4549** | **0.8873** | 0.8509 | 0.6619 | **0.3586** | **0.4524** |
| $RCIQM_4$ | 65 | **0.8828** | **0.8629** | **0.6759** | **0.3612** | **0.4519** | 0.8864 | **0.8520** | **0.6633** | 0.3624 | 0.4540 |
| $RCIQM_8$ | 33 | 0.8795 | 0.8497 | 0.6586 | 0.3667 | 0.4579 | 0.8839 | 0.8431 | 0.6540 | 0.3641 | 0.4586 |
| $RCIQM_{16}$ | 17 | 0.8739 | 0.8366 | 0.6550 | 0.3620 | 0.4677 | 0.8852 | 0.8450 | 0.6551 | 0.3589 | 0.4563 |
| $RCIQM_{32}$ | 9 | 0.8700 | 0.8268 | 0.6423 | 0.3691 | 0.4744 | 0.8791 | 0.8324 | 0.6428 | 0.3682 | 0.4675 |
| $RCIQM_{64}$ | 5 | 0.8629 | 0.8179 | 0.6283 | 0.3867 | 0.4862 | 0.8699 | 0.8215 | 0.6317 | 0.3814 | 0.4837 |

| Metrics | Info. number | CSIQ database (116 images) [11] | | | | | Database size-weighted average | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PLCC | SRCC | KRCC | AAE | RMS | PLCC | SRCC | KRCC | AAE | RMS |
| LTG [14] | whole | 0.9560 | 0.9414 | 0.7880 | 0.0392 | 0.0494 | 0.8130 | 0.7390 | 0.5684 | 0.3325 | 0.4163 |
| $RCIQM_1$ | 257 | **0.9645** | **0.9569** | **0.8198** | **0.0353** | **0.0445** | **0.8985** | **0.8792** | **0.7010** | **0.2494** | **0.3134** |
| $RCIQM_2$ | 129 | 0.9614 | 0.9508 | 0.8039 | 0.0373 | 0.0463 | **0.8981** | 0.8764 | **0.6979** | **0.2486** | **0.3137** |
| $RCIQM_4$ | 65 | 0.9558 | 0.9458 | 0.7928 | 0.0399 | 0.0495 | 0.8966 | **0.8772** | 0.6967 | 0.2522 | 0.3154 |
| $RCIQM_8$ | 33 | 0.9590 | 0.9487 | 0.8003 | 0.0386 | 0.0477 | 0.8965 | 0.8751 | 0.6948 | 0.2524 | 0.3160 |
| $RCIQM_{16}$ | 17 | 0.9611 | 0.9493 | 0.8039 | 0.0377 | 0.0465 | 0.8962 | 0.8739 | 0.6951 | 0.2509 | 0.3167 |
| $RCIQM_{32}$ | 9 | 0.9624 | 0.9485 | 0.8027 | 0.0367 | 0.0458 | 0.8944 | 0.8701 | 0.6906 | 0.2538 | 0.3198 |
| $RCIQM_{64}$ | 5 | **0.9639** | **0.9549** | **0.8165** | **0.0360** | **0.0448** | 0.8918 | 0.8677 | 0.6879 | 0.2584 | 0.3243 |

### D. Visualized Comparison

Furthermore, we show scatter plots of classical FR SSIM, state-of-the-art FR LTG and RR SDM, and our RCIQM metrics on CID2013, CCDI2014, TID2008, TID2013 and CSIQ databases in Fig. 7. It is easy to find that our technique has acquired the impressive linearity and monotonicity, which also confirms the substantially high performance of our metric on the visual quality evaluation of contrast-changed images.

### E. Statistical Significance

Additionally, the f-test is adopted to compute the statistical significance of the presented algorithm. The f-test measures the prediction residuals of the converted objective quality scores (after the five-parameter logistic nonlinear regression function) and subjective MOS/DMOS values. Suppose $F$ be the ratio of two residual variances and $F_c$ (decided by the residuals' number and confidence level) be the judgement threshold, the performance distinction of two testing IQA metrics is regarded as significant in case of $F > F_c$. We show the statistical significance between the proposed RCIQM and other testing IQA metrics in Table III, in which the symbol "+1", "0" or "-1" stands for that our model is statistically superior, indistinguishable, or inferior to the associated metric. Clearly, our RCIQM operates very well. On the four large-scale CID2013, CCID2014, TID2008 and TID2013 databases, the proposed model outperforms nearly all testing IQA metrics. Our method is superior to most IQA models but is only comparable to FEDM and RIQMC on the CSIQ database. Overall, our technique is currently of the best performance for the IQA of contrast adjustment.

### F. Analysis on RR Information

For a RR IQA metric, besides the correlation performance, one important index is the amount of RR information used. The RR information in our RCIQM technique consists of one single number of free energy entropy and a global histogram with 256 bins extracted from the pristine image. A simple but valid way for reducing the RR information is to combine the neighboring bins in the global histogram together. Here we fuse $n$ adjacent bins to one, where $n$ is assigned to be 1, 2, 4, 8, 16, 32 and 64, respectively. We dub these metrics as $RCIQM_1$, $RCIQM_2$, $RCIQM_4$, $RCIQM_8$, $RCIQM_{16}$, $RCIQM_{32}$ and $RCIQM_{64}$, which all adopt the same parameters used in RCIQM.

Notice that the $RCIQM_1$ is itself our original proposed RCIQM algorithm with 257 (=256+1) numbers as RR features, while the rest six metrics require 129 (=128+1), 65 (=64+1), 33 (=32+1), 17 (=16+1), 9 (=8+1) and 5 (=4+1) numbers in sequence. We present the performance evaluations of the above seven approaches and the effective LTG model on five testing image databases and the weighted average in Table IV, where the amount of information used in each quality metric is also provided for comparison. It can be readily found that as the number of RR features reduces, the performance of our metric decreases in most cases but with a very small amount. For example, the $RCIQM_{64}$ with only 5 numbers as RR information has also attained better prediction accuracy than the recent LTG method (the best in testing FR-IQA methods) on average. That is to say, in addition to an effective RR IQA approach, this paper provides more choices between the correlation performance and the amount of RR information as well.

## V. Conclusion

In this paper, we have investigated into the problem of IQA with contrast change, and introduce a new Reduced-reference Contrast-changed image Quality Index (RCIQM) by combining bottom-up and top-down strategies. Considering that the visual quality of the contrast-altered image is highly connected to the psychovisual mechanism in the human brain, the bottom-up strategy applies the new HPNP model with luminance, contrast and structural information for weighting. In the top-down strategy we compare the histogram of the contrast-changed image with those of the original and histogram equalized versions using symmetric K-L divergence. Results of experiments on five contrast related databases (namely, CID2013, CCID2014, TID2008, CSIQ and TID2013) verify the superiority of our proposed RCIQM over up to ten classical / state-of-the-art FR- and RR-IQA models in quantitative performance measures and statistical significance comparison. Furthermore, we want to stress that: 1) no matter how large the image size is, such as 4K HD (high definition), the proposed method merely needs a single number of free energy (in the bottom-up strategy) and fixed numbers of global histogram of the original image (in the top-down strategy); 2) our algorithm is insensitive to small translations and rotations which exert very little influence on visual quality in comparison with most existing IQA approaches, which makes RCIQM effective in a wide range of environments.

## References


[1] 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.

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

[3] 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.

[4] 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.

[5] 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.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[7] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," *in Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, pp. 1398-1402, Nov. 2003.

[8] K. Gu, M. Liu, G. Zhai, X. Yang, and W. Zhang, "Quality assessment considering viewing distance and image resolution," *IEEE Trans. Broadcasting*, vol. 61, no. 3, pp. 520-531, Sep. 2015.

[9] K. Gu, S. Wang, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Analysis of distortion distribution for pooling in image quality prediction," *IEEE Trans. Broadcasting*, vol. 62, no. 2, pp. 446-456, Jun. 2016.

[10] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284-2298, Sep. 2007.

[11] 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

[12] 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.

[13] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500-1512, Apr. 2012.

[14] 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.

[15] 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.

[16] 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.

[17] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41-52, Jan. 2012.

[18] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," *in Proc. IEEE Int. Symp. Circuits and Syst.*, pp. 1095-1098, May 2013.

[19] L. Ma, S. Li, F. Zhang, and K. N. Ngan, "Reduced-reference image quality assessment using reorganized DCT-based image representation," *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 824-829, Aug. 2011.

[20] M. Liu, G. Zhai, Z. Zhang, S. Tan, K. Gu, and X. Yang, "Using image signature for effective and efficient reduced-reference image quality assessment," *in Proc. IEEE Int. Conf. Multimedia and Expo Workshops*, pp. 1-6, Jul. 2014.

[21] L. Ma, S. Li, and K. N. Ngan, "Reduced-reference image quality assessment in reorganized DCT domain," *Signal Process.: Image Commun.*, vol. 28, no. 8, pp. 884-902, Sep. 2013.

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

[23] W. Lin, L. Dong, and P. Xue, "Visual distortion gauge based on discrimination of noticeable contrast changes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 900-909, Jul. 2005.

[24] 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.

[25] E. Peli, "Contrast in complex images," *Journal of Optical Society of America*, vol. 7, pp. 2032-2040, Oct. 1990.

[26] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921-1935, Sept. 2009.

[27] 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, Sep. 2015.

[28] X. Wu, "A linear programming approach for optimal contrast-tone mapping," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1262-1272, May 2011.

[29] 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, Sep. 2013.

[30] 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.

[31] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Stat.*, vol. 22, pp. 79-86, 1951.

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

[33] 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.

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

[35] Kodak Lossless True Color Image Suite: http://r0k.us/graphics/kodak/

[36] D. C. Knill and A. Pouget, "The Bayesian brain: The role of uncertainty in neural coding and computation," *Trends Neurosci.*, vol. 27, no. 12, pp. 712-719, 2004.

[37] H. Attias, "A variational bayesian framework for graphical models," *in Proc. Adv. Neural Inf. Process. Syst.*, vol. 12, pp. 209-215, 2000.

[38] 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.

[39] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Trans. Broadcasting*, vol. 60, no. 3, pp. 555-567, Sep. 2014.

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

[41] P. Milanfar, "A tour of modern image filtering," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 106-128, Jan. 2013.

[42] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254-1259, Nov. 1998.

[43] M. C. Morrone, J. Ross, D. C. Burr, and R. Owens, "Mach bands are phase dependent," *Nature*, vol. 324, pp. 250-253, Nov. 1986.

[44] P. Kovesi, "Image features from phase congruency," *Videre: J. Comp. Vis. Res.*, vol. 69, no. 3, pp. 1-26, 1999.

[45] Y. Fang, W. Lin, B.-S. Lee, C.-T. Lau, Z. Chen, and C.-W. Lin, "Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum," *IEEE Trans. Multimedia*, vol. 14, no. 1, pp. 187-198, Feb. 2012.

[46] X. Min, G. Zhai, Z. Gao, and K. Gu, "Visual attention data for image quality assessment databases," *in Proc. IEEE Int. Symp. Circuits and Systems*, pp. 894-897, Jun. 2014.

[47] 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.

[48] O. Pele and M. Werman, "Fast and robust earth mover's distances," *in Proc. IEEE Int. Conf. Computer Vision*, pp. 460-467, Sept. 2009.

[49] B. Schauerte and G. A. Fink, "Web-based learning of naturalized color models for human-machine interaction," *in Proc. IEEE Int. Conf. Digital Image Computing: Techniques and Applications*, pp. 498-503, Dec. 2010.

[50] B. Schauerte and R. Stiefelhagen, "Learning robust color name models from web images," *in Proc. IEEE Int. Conf. Pattern Recognition*, pp. 3598-3601, Nov. 2012.

[51] 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/.


PLACE PHOTO HERE

**Min Liu** received the B.E. degree in electronic engineering from Xidian University, Xi'an, China, in 2012. Her research interests include image/video quality assessment and perceptual signal processing. She is visiting at the Department of Electrical Engineering, Technical Institute of the University of Nantes, Nantes, France.

PLACE PHOTO HERE

**Ke Gu** received the B.S. and PhD degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015. He is the reviewer for IEEE T-IP, T-MM, T-CSVT, T-CYB, T-BC, J-STSP, SPL, Information Sciences, Neurocomputing, SPIC, JVCI, DSP, etc. He has reviewed more than 30 journal papers each year. From Jul. to Nov. 2014, he was visiting at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. From Dec. 2014 to Jan. 2015, he was visiting at the School of Computer Engineering, Nanyang Technological University, Singapore. From Feb. to Mar. and from Nov. to Dec. 2015, he was visiting at the Department of Computer Science and Technology, Peking University, Beijing, China. His research interests include quality assessment, contrast enhancement and visual saliency detection. Dr. Gu received the Best Paper Award at the IEEE International Conference on Multimedia and Expo (ICME) 2016. He is a Special Session Organizer for VCIP 2016.

PLACE PHOTO HERE

**Guangtao Zhai** (M'10) received the B.E. and M.E. degrees from Shandong University, Shandong, China, in 2001 and 2004, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2009, where he is currently a Research Professor with the Institute of Image Communication and Information Processing. From 2008 to 2009, he was a Visiting Student with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he was a Post-Doctoral Fellow from 2010 to 2012. From 2012 to 2013, he was a Humboldt Research Fellow with the Institute of Multimedia Communication and Signal Processing, Friedrich Alexander University of Erlangen-Nuremberg, Germany. He received the Award of National Excellent Ph.D. Thesis from the Ministry of Education of China in 2012. His research interests include multimedia signal processing and perceptual signal processing.

PLACE PHOTO HERE

**Patrick Le Callet** (SM'14) received the M.Sc. and Ph.D. degrees in image processing from the Ecole Polytechnique de l'Universite de Nantes, Nantes, France. He was with the Department of Electrical Engineering, Technical Institute of the University of Nantes, Nantes, as an Assistant Professor from 1997 to 1999 and a full-time Lecturer from 1999 to 2003. Since 2003, he teaches with the Department of Electrical Engineering and the Department of Computer Science, École Polytechnique de l'Universite de Nantes (Engineering School), where he is currently a Full Professor. Since 2006, he has been the Head of the Image and Video Communication Laboratory with CNRS IRCCyN, a group of more than 35 researchers. He is currently involved in research dealing with the application of human vision modeling in image and video processing. He has co-authored over 200 publications and communications and holds 13 international patents in his research topics. His current research interests include 3-D image and video quality assessment, watermarking techniques, and visual attention modeling and applications. Prof. Le Callet served as an Associate Editor of several journals, such as the IEEE Transactions on Image Processing (T-IP), the IEEE Transactions on Circuits System and Video Technology, the EURASIP Journal on Image and Video Processing, and SPIE Electronic Imaging.

PLACE PHOTO HERE

**Wenjun Zhang** (F'11) received the B.S., M.S. and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 1984, 1987 and 1989, respectively. From 1990 to 1993, He worked as a post-doctoral fellow at Philips Kommunikation Industrie AG in Nuremberg, Germany, where he was actively involved in developing HD-MAC system. He joined the Faculty of Shanghai Jiao Tong University in 1993 and became a full professor in the Department of Electronic Engineering in 1995. As the national HDTV TEEG project leader, he successfully developed the first Chinese HDTV prototype system in 1998. He was one of the main contributors to the Chinese Digital Television Terrestrial Broadcasting Standard issued in 2006 and is leading team in designing the next generation of broadcast television system in China from 2011. He holds more than 40 patents and published more than 90 papers in international journals and conferences. Prof. Zhang's main research interests include digital video coding and transmission, multimedia semantic processing and intelligent video surveillance. He is a Chief Scientist of the Chinese National Engineering Research Centre of Digital Television (NERC-DTV), an industry/government consortium in DTV technology research and standardization and the Chair of Future of Broadcast Television Initiative (FOBTV) Technical Committee.