Center Emphasized Visual Saliency and a Contrast-based Full Reference Image Quality Index

Md Abu Layek¹, Sanjida Afroz², TaeChoong Chung¹, Eui-Nam Huh¹

¹ Department of Computer Science and Engineering, Kyung Hee University Global Campus, Yongin-17104, Republic of Korea; ² Independent Researcher, MS in Psychology, University of Dhaka, Bangladesh

* abulayek@yahoo.com, layek@khu.ac.kr, sanjida.psy@gmail.com

abstract

Objective image quality assessment (IQA) is imperative in the current multimedia-intensive world, in order to assess the visual quality of an image at close to a human level of ability. Many parameters such as color intensity, structure, sharpness, contrast, presence of an object, etc., draw human attention to an image. Psychological vision research suggests that human vision is biased to the center area of an image and display screen. As a result, if the center part contains any visually salient information, it draws human attention even more and any distortion in that part will be better perceived than other parts. To the best of our knowledge, previous IQA methods have not considered this fact. In this paper, we propose a full reference image quality assessment (FR-IQA) approach using visual saliency and contrast; however, we give extra attention to the center by increasing the sensitivity of the similarity maps in that region. We evaluated our method on three large-scale popular benchmark databases used by most of the current IQA researchers (TID2008, CSIQ and LIVE), having a total of 3345 distorted images with 28 different kinds of distortions. Our method is compared with 13 state-of-the-art approaches. This comparison reveals the stronger correlation of our method with human-evaluated values. The prediction-of-quality score is consistent for distortion specific as well as distortion independent cases. Moreover, faster processing makes it applicable to any real-time application. The MATLAB code is publicly available to test the algorithm and can be found online.

Introduction

Computer-based automatic image quality assessment (IQA) has been sought after for decades because numerous image and video applications need this assessment to automate their quality maintenance. To date, IQA research has been significantly advanced, however, this is still an active area of research to bring the methods closer to human-level ability. In the literature, there are three principle IQA approaches. No-reference image quality assessment (NR-IQA) uses a single distorted image without any reference image, whereas in reduced-reference IQA (RR-IQA), partial information of a reference image is given. The third category is full-reference IQA (FR-IQA) where the complete reference image is given along with the distorted one. In this paper, we deal with FR-IQA.

Table presents a brief survey on several state-of-the-art IQA approaches, which we compared in this paper. We made a separate column for the pooling strategy because recent IQA research tends to combine multiple features where pooling plays an

¹http://layek.khu.ac.kr/CEQI
important role. Early pixel-based, faster IQA methods such as mean squared error (MSE) and peak signal-to-noise ratio (PSNR) consider neither the human visual system (HVS) nor any aspects of human perception. Thus, those approaches fail to achieve good correlation with human assessment [1][2]. Two images with the same PSNR or MSE may be perceived in totally different ways by a human observer. However, humans are the ultimate receiver of images; as a result, the search for methods that can achieve a closer correlation with humans is ongoing. Wang et al., in their revolutionary work on the structural similarity index or SSIM [3], argued that human visual perception is highly sensitive to structural information. The SSIM index incorporates luminance, contrast, and structural comparison information and achieves a very good correlation with the mean opinion scores (MOS) of human observers. Inspired by the success of SSIM, several extended versions, such as the multi-scale structural similarity for image quality assessment (MS-SSIM) [4] and Information content weighting for perceptual image quality assessment (IW-SSIM) [5], were proposed by the same research group. IW-SSIM utilizes an image pyramid to decompose the original and distorted images into versions of varying scales, and then computes the information content from the images. Finally, it finds the quality score using the information content as a weighting function.

Based on shared information between the reference and distorted images, Sheikh et al. proposed the information fidelity criteria (IFC) [6] and the visual information fidelity (VIF) [7]. The most apparent distortion (MAD) approach [8] separates images based on the distortion and applies either a detection-based strategy or an appearance-based strategy. Some of the methods, such as the noise quality measure (NQM) [9] and the visual signal-to-noise ratio (VSNR) [10], take into account the HVS by incorporating interactions among different visual signals. In contrast, other approaches, including the popular feature similarity index or FSIM [11], emphasize phase congruency [12][13][14]. FSIM uses the image gradient as a secondary feature and local quality maps are weighted by phase congruency to obtain the final score. The image gradient has been used effectively in a number of other works [15][16]. Xue et al., in their gradient magnitude standard deviation (GMSD) [17], used the gradient magnitude with a different pooling strategy, by applying the standard deviation, and Alaei et al. adopted a similar approach for assessing document images [18]. Both examples prove the effectiveness of standard deviation pooling, however, the authors of the GMSD approach showed that standard deviation (SD) pooling is not effective for all types of methods. Wang et al. proposed the multi-scale contrast similarity deviation (MCSD) method [19], which can be termed as a continuation of SSIM and MS-SSIM, since it also uses the root mean square (RMS) contrast similarity; however, they employed standard deviation pooling for the final score.

Meanwhile, inspired by vision-related psychological research, visual saliency (VS)-based IQA methods [20][21], which utilize different kinds of visual saliencies [22][23][24], have attracted researchers’ attention. In the visual saliency index (VSI) method [23], VS is used as both a quality map and the weighting function at the pooling stage. The spectral residual similarity index (SR-SIM) [21] uses the spectral residual saliency, which makes the approach very fast while maintaining a competitive correlation with the mean opinion score. Combining VS with other features has also become popular [25][26]. Li et al. proposed an approach that combines VS and FSIM while, recently, Jia et al. used contrast and spectral residual saliency as well as summation-based SD pooling.
| IQA Method | Principle Consideration | Pooling Used | Comments |
|------------|-------------------------|--------------|----------|
| PSNR       | Pixel-by-pixel error    | Average      | Primitive method, does not consider HVS, poor correlation with humans, low computation, widely used. |
| SSIM       | Luminance, contrast, structure | Average | Milestone method to consider structural information that better represents HVS. |
| MS-SSIM    | Multi-scale structure   | Weighted sum | An extension to SSIM that incorporates variations of viewing conditions and capable of multi-scale assessment. |
| IW-SSIM    | Information content extraction | Weighted by information content | Main emphasis is on information content extraction which is applicable to other methods as well, uses image pyramid. |
| MAD        | Most visible distortion | Weighted product | A state-of-the art IQA approach using phase congruency and the gradient magnitude weighted by phase congruency to calculate the final score. |
| FSIM       | Phase congruency, Gradient magnitude | Weighted average | Introduced visual saliency for IQA where VS is used for both the quality map and weighting function at the pooling stage. |
| GMSD       | Image gradient          | Standard deviation | A very fast assessment approach after PSNR showing competitive performance with other state-of-the-art approaches. |
| VSI        | Visual saliency (VS)    | Weighted average | Another fast method next to GMSD, but providing better results. |
| MCSD       | Multi-scale contrast    | Standard deviation | Quantifies the extracted reference information from a distorted image. |
| VIF        | Visual information extraction | Average | Considers HVS using distortion and noise, better than PSNR. |
| NQM        | Distortion and noise    | Squared sum  | Uses spectral residual saliency and image gradient, which shows competitive performance. |
| SR-SIM     | Saliency and gradient   | Weighted average | Uses deviation pooling in a combined method. |
| VSP        | Saliency and contrast   | Weighted sum of deviations | Gives special attention to the center area. |

PSNR: peak signal-to-noise ratio; SSIM: structural similarity index; MS-SSIM: multi-scale structural similarity; IW-SSIM: information content weighted structural similarity; MAD: most apparent distortion; FSIM: feature similarity index; GMSD: gradient magnitude standard deviation; VSI: visual saliency index; MCSD: multi-scale contrast similarity deviation; VIF: visual information fidelity; NQM: noise quality measure; SR-SIM: spectral residual similarity index; VSP: visual saliency plus; CEQI: center-emphasized quality index; HVS: human visual system.
In the context of the HVS, center bias in early eye movements is an established fact in psychological vision research [27, 28, 29, 30]. Bindemann found that eye movement is biased not only to the scene center but also to the screen center [31]. As a result, if a scene appears at the center of the screen, it will receive the most attention. For example, in Figure 1, the human eye will first move to the Block05 region and if that part has visually important information then it will attract even more attention. As a result, people will be more sensitive to the distortions in this region. To the best of our knowledge, there is no research in IQA considering this center bias for quality assessment.

In this paper, we propose a new method for IQA which accounts for the center emphasis in HVS. In the proposed method, we first obtain both the contrast and VS similarity maps for the entire image. To give center emphasis, we find the VS similarity map of the mid-region and apply element-wise multiplication in the mid-part to raise the similarity deviation there. However, for the contrast similarity, we apply element-wise squaring in the center part. Contrast is a local quality map, so we do not calculate the contrast of the mid-area separately. On the other hand, VS is a global quality map, and thus it is calculated differently in the mid-region. The final score is obtained by performing weighted summation of the standard deviations on both of the similarity maps; further details with mathematical equations are given in section 2.

Figure 1. The image 'Sailing.bmp' is split into nine blocks, where Block05 is the center area.

We evaluated our proposed method on three popular benchmark databases for IQA research and compared it with 13 other state-of-the-art approaches. The results in terms of the correlations with human evaluated scores show that the method proposed by us outperforms the other approaches, with a reasonable amount of processing time.
This paper is organized as follows. Section 1 describes some underlying theories and related techniques. Section 2 explains the proposed center-emphasized assessment approach, and the results with relevant discussions are presented in Section 3. Finally, the paper is concluded in Section 4.

1 Background

In this section, we briefly review the underlying theories on which the content of this paper relies, including the spectral residual visual saliency similarity, contrast similarity, standard deviation pooling, and relevant evaluation metrics.

1.1 Spectral Residual Visual Saliency Similarity

In the human visual system (HVS), some interesting or salient regions of an image receive more attention than other parts. Detection of image saliency itself is an active field in vision research. As such, the human is more sensitive to these salient parts and any distortion in these parts attracts more intense attention, which makes it an important feature in IQA methods. There are a lot of saliency detection techniques available [32], among which spectral residual saliency detection [22] is a very fast approach. We adopt the saliency map generator described in the SR-SIM [24] and the visual saliency plus contrast (VSP) [26] approaches.

For an image \( f(x,y) \), according to Reference [22], the spectral residual saliency (SRS) is computed as follows:

\[
M_f(u,v) = \text{abs}[\mathcal{F}\{f(x,y)\}(u,v)] \tag{1}
\]

\[
A_f(u,v) = \text{angle}[\mathcal{F}\{f(x,y)\}(u,v)] \tag{2}
\]

\[
L_M(u,v) = \log\{M_f(u,v)\} \tag{3}
\]

\[
R_f(u,v) = L_M(u,v) - h_n(u,v) * L_M(u,v) \tag{4}
\]

\[
SRS(x,y) = g(x,y) * [\mathcal{F}^{-1}\{\exp(R_f + jA_f)\}(x,y)]^2, \tag{5}
\]

where \( \mathcal{F} \) and \( \mathcal{F}^{-1} \) denote the Fourier and inverse Fourier transforms, respectively; \( \text{abs}(.) \) and \( \text{angle}(.) \) return the magnitude and argument of a complex number, respectively; \( h_n(u,v) \) is an \( n \times n \) mean filter; \( g(x,y) \) is a Gaussian filter; and * denotes the convolution operation.

In this way, we calculate the SRS for both the reference and distorted images denoted by \( SRS_r(x,y) \) and \( SRS_d(x,y) \), respectively. Then, the spectral residual visual saliency similarity (SRVSS\( (r,d) \)) is calculated as:

\[
SRVSS \tag{6}
\]

where \( \odot \) is the element-wise multiplication, \( ^2 \) is the element-wise squaring and \( c_1 \) is a positive constant used to increase the calculation stability.
1.2 Contrast Similarity

Contrast is a basic perceptual attribute of an image [33] which varies greatly over the image, and the contrast map (CM) contains the spatial distribution of those varying values. There are many ways devised for calculating CMs and in this paper, we adopted the RMS contrast from SSIM, since it achieves better performance for natural images. The RMS contrast map \( C_X \) of an image signal \( X \) is given by:

\[
C_X = \left[ \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_X)^2 \right]^{\frac{1}{2}},
\]

where \( N \) is the total number of pixels in the image, \( x_i \) is the intensity of pixel \( i \), and \( \mu_X \) is the mean intensity, defined as:

\[
\mu_X = \frac{1}{N} \sum_{i=1}^{N} x_i.
\]

Again, using Equation (7), we obtain the contrast maps \( C_r \) and \( C_d \) for both reference and distorted images, respectively, and find the contrast similarity \( CS(r,d) \) as follows:

\[
CS(r,d) = \frac{2C_r \odot C_d + c_2}{C_r^2 + C_d^2 + c_2},
\]

where similar to Equation (6), \( \odot \) is the element-wise multiplication, \(^2\) is the element-wise squaring and \( c_2 \) is a positive constant used to increase the calculation stability.

1.3 Standard Deviation Pooling

As discussed in the introduction, standard deviation (SD) pooling achieves very good performance in specific cases and is adopted by several successful methods. Jia et al. conducted an experiment with several other combinations of pooling and found that SD pooling provides the best correlation. The final quality score (\( QS \)) is calculated using the following equation:

\[
QS = \frac{1}{w_1 + w_2} \left[ w_1 \times \text{std} \{SRVSS(r,d)\} + w_2 \times \text{std} \{CS(r,d)\} \right],
\]

where \( w_1 \) and \( w_2 \) are weighting factors that specify the importance of VSS and CS, respectively. The standard deviations in the above equation are defined as:

\[
\text{std}(SRVSS(r,d)) = \left\{ \frac{1}{M} \sum_{i=1}^{M} (SRVSS_i - \mu_{SRVSS})^2 \right\}^{\frac{1}{2}}
\]

\[
\text{std}(CS(r,d)) = \left\{ \frac{1}{M} \sum_{i=1}^{M} (CS_i - \mu_{CS})^2 \right\}^{\frac{1}{2}},
\]

where \( M \) is the number of total elements in the similarity matrices; \( SRVSS_i \) and \( CS_i \) are the \( i \)th items; \( \mu_{SRVSS} \) and \( \mu_{CS} \) are the mean values of the \( SRVSS(r,d) \) and \( CS(r,d) \), respectively, and are given by:

\[
\mu_{SRVSS} = \frac{1}{M} \sum_{i=1}^{M} SRVSS_i
\]

\[
\mu_{CS} = \frac{1}{M} \sum_{i=1}^{M} CS_i
\]
1.4 Evaluation Metrics

The performance of any IQA method is usually measured by the mean squared error and several other correlations with the subjective scores that are human evaluated values usually in the form of MOS, or their differential DMOS. However, to apply linear correlation, the two compared values should be on the same scale and perfectly linearly correlated [34]. To ensure better fairness, before applying the linear correlation measurements, a logistic mapping function is used to convert the objective scores. We use the following nonlinear regression model as suggested by Sheikh [35].

\[ q' = \beta_1 \left( \frac{1}{2} - \frac{1}{1 + \exp(\beta_2(q - \beta_3))} \right) + \beta_4 q + \beta_5, \]  

(15)

where \( q \) is the objective score calculated by an IQA method, \( q' \) is the mapped value, and \( \beta_i, i = 1, 2, 3, 4, 5 \) are the parameters that are tuned based on the relationship between objective and subjective scores. We utilized the MATLAB function \textit{nlinfit} to find the optimal parameters. After the mapping is done, the subjective scores are then used with these mapped scores to find the correlation coefficients.

One of the widely adopted basic correlations is Pearson’s linear correlation coefficient (PLCC) which is defined as follows:

\[ \text{PLCC}(o, s) = \frac{\sum_{i=1}^{m} (o_i - \mu_o)(s_i - \mu_s)}{\left( \sum_{i=1}^{m} (o_i - \mu_o)^2 \right)^{1/2} \left( \sum_{i=1}^{m} (s_i - \mu_s)^2 \right)^{1/2}}, \]  

(16)

where \( m \) is the number of distorted images; \( o \) and \( s \) are vectors of the objective and subjective scores, respectively; and \( \mu_o \) and \( \mu_s \) are the mean scores, defined by:

\[ \mu_o = \frac{1}{m} \sum_{i=1}^{m} o_i \]  

(17)

\[ \mu_s = \frac{1}{m} \sum_{i=1}^{m} s_i \]  

(18)

In our case, the objective scores \( o \) are actually the mapped scores using Equation (15). If the nonlinear mapping in Equation (15) is to be avoided, then rank order coefficients can be used. The most popular Spearman’s rank-order correlation coefficient (SROCC) is defined as:

\[ \text{SROCC}(o, s) = \text{PLCC}(\text{rank}(o), \text{rank}(s)). \]  

(19)

The function \textit{rank()} of a vector returns a rank-vector, where the \( i \)-th entry contains the relative rank of the \( i \)-th item in the original vector.

Another popularly adopted rank order coefficient is Kendall’s rank-order correlation coefficient (KROCC), which is given as below:

\[ \text{KROCC}(o, s) = \frac{C - D}{m(m - 1)/2}, \]  

(20)

where \( C \) is the number of concordant pairs that are consistently correlated between objective and subjective scores; and \( D \) is the number of discordant pairs.

The root mean square error (RMSE) is also commonly adopted and is defined as:

\[ \text{RMSE}(o, s) = \left\{ \frac{1}{m} \sum_{i=1}^{m} (o_i - s_i)^2 \right\}^{1/2}. \]  

(21)

A larger value for PLCC, SROCC, and KROCC indicates that the corresponding method is better. On the other hand, a smaller value of the RMSE is a sign of a superior IQA.
2 Proposed Center-Emphasized Quality Assessment

The general flow diagram of our proposed method is presented in Figure 2. At first, the center parts of both the reference and distorted images are extracted. To do this, we split the image in $3 \times 3$ image blocks as shown in Figure 1 and the fifth block, which resides in the middle both horizontally and vertically, is taken as the center area. If the original image dimension is $(H \times W)$, then the corresponding dimension for the center block becomes $(H_{\text{mid}} \times W_{\text{mid}})$, where:

$$H_{\text{mid}} = \left\lceil \frac{H}{3} \right\rceil \quad \text{and} \quad W_{\text{mid}} = \left\lceil \frac{W}{3} \right\rceil .$$  (22)

The center block is defined as a rectangular area identified by two corner points $(x_{\text{min}}, y_{\text{min}})$ and $(x_{\text{max}}, y_{\text{max}})$, where:

$$x_{\text{min}} = \left\lceil \frac{H}{3} \right\rceil, \quad y_{\text{min}} = \left\lceil \frac{W}{3} \right\rceil, \quad x_{\text{max}} = \left\lceil \frac{H}{3} \right\rceil + H_{\text{mid}}, \quad \text{and} \quad y_{\text{max}} = \left\lceil \frac{W}{3} \right\rceil + W_{\text{mid}} .$$  (23)

First, the saliency similarity maps for the full images and middle images are found using Equations (1)–(6) and are denoted as $VSS$ and $VSS_{\text{mid}}$, respectively. Simultaneously, the contrast similarity map for the full-size, $CS$, is also obtained. As discussed before, we do not derive the CS map for middle images.

Then, we increase the sensitivity of the center area within both of the maps. Let $VSS(mid)$ and $CS(mid)$ be the center areas of $VSS$ and $CS$, respectively. The updated middle parts will be determined as follows:

$$VSS(mid) = VSS(mid) \odot VSS_{\text{mid}}$$  (24)

$$CS(mid) = CS(mid) \odot CS(mid),$$  (25)

where $\odot$ is the element-wise multiplication.

With the updated middle portion, we obtain the finalized maps $VSS_{\text{final}}$ and $CS_{\text{final}}$, and using Equation (10), we calculate the final quality score of the proposed method $CEQI$ as:

$$CEQI = \frac{1}{w_1 + w_2} \left[ w_1 \times \text{std}(VSS_{\text{final}}) + w_2 \times \text{std}(CS_{\text{final}}) \right]$$  (26)
Figure 2. Flow diagram of the proposed center-emphasized approach.

3 Results and Analysis

Experiments were carried out on three popular benchmark databases for IQA research—TID2008 [36], CSIQ [37] and LIVE [38]. Our approach was compared with 13 other state-of-the-art approaches as listed in Table 1. Basic information about the databases is given in Table 2 and the distortion information is recorded in Table 3.

Table 2. Basic information about the databases used for the experiments.

| Dataset | Reference Images | Distorted Images | Distortion Types | No. of Subjects |
|---------|------------------|------------------|------------------|----------------|
| TID2008 | 25               | 1700             | 17               | 838            |
| CSIQ    | 30               | 866              | 6                | 35             |
| LIVE    | 29               | 779              | 5                | 161            |

\[ CEQI = \frac{1}{w_1 + w_2} \left[ w_1 \times \text{std}(VSS_{\text{final}}) + w_2 \times \text{std}(CS_{\text{final}}) \right] \]
Table 3. Types of distortion used in each database.

| TID2008CSIQ | LIVE | Type of distortion | Abbreviation |
|-------------|------|-------------------|--------------|
| Y           | Y    | Y                 | Additive Gaussian noise | AGN |
| Y           | -    | -                 | Additive noise in color components | ANC |
| Y           | -    | -                 | Spatially correlated noise | SCN |
| Y           | -    | Masked noise      | MN |
| Y           | -    | High frequency noise | HFN |
| Y           | -    | Impulse noise     | IN |
| Y           | -    | Quantization noise | QN |
| Y           | Y    | Gaussian blur     | GB |
| Y           | -    | Image denoising   | IDN |
| Y           | Y    | JPEG compression  | JPEG |
| Y           | Y    | JPEG2000 compression | JP2K |
| Y           | -    | JPEG transmission errors | JGTE |
| Y           | -    | JPEG2000 transmission errors | J2TE |
| Y           | -    | Non-eccentricity pattern noise | NEPN |
| Y           | -    | Local block-wise distortions of different intensity | LBD |
| Y           | -    | Mean shift (intensity shift) | MS |
| Y           | -    | Contrast change   | CTC |
| -           | Y    | Fast fading Rayleigh | FF |
| -           | Y    | Additive pink Gaussian noise | AWPN |

For performance comparison, we use four commonly adopted metrics—Spearman’s rank-order correlation coefficient (SROCC), Kendall’s rank-order correlation coefficient (KROCC), Pearson’s linear correlation coefficient (PLCC), and the root mean square error (RMSE)—which we defined in Section 4.

Table 4 compares the four metrics among the different IQA models, for all of the three databases. The top three values for each metric are typed in boldfaced and the third highest value is colored black. However, in the case of RMSE, coloring is done in a reverse way, i.e., the lowest value is colored in blue and so on, since a lower error (RMSE)—which we defined in Section 1.4.

– For SROCC, KROCC and PLCC metrics, the higher the value, the better the method.
– For RMSE a lower score is better.
– For each row, the first, second and third-ranked performances are highlighted respectively in blue, red and black colors.
– For SROCC, KROCC and PLCC metrics, the higher the value, the better the method whereas for RMSE a lower score is better.

Table 4. Performance comparison of IQA methods on three databases.

| Dataset | Metric | PSNR | SSIM | MS-SSIM | IW-SSIM | MAD | FSIM | GMSD | VSI | MCSD | VIF | NQM | SRSIM | VSP | CEQI |
|---------|--------|------|------|---------|---------|-----|------|------|-----|------|-----|-----|-------|-----|------|
| SROCC   | 0.8759 | 0.7749 | 0.8542 | 0.8559 | 0.8490 | 0.8805 | 0.8927 | 0.8879 | 0.8911 | 0.7941 | 0.6241 | 0.8011 | 0.9061 | 0.9069 |
| KROCC   | 0.8696 | 0.5768 | 0.6558 | 0.6616 | 0.6445 | 0.6896 | 0.7022 | 0.7213 | 0.7133 | 0.5860 | 0.4688 | 0.7349 | 0.7225 | 0.7207 |
| PLCC    | 0.5109 | 0.3722 | 0.4851 | 0.4879 | 0.4836 | 0.4873 | 0.4878 | 0.4864 | 0.4884 | 0.4884 | 0.4884 | 0.4884 | 0.4884 | 0.4884 |
| RMSE    | 1.1527 | 0.8511 | 0.7373 | 0.6985 | 0.6763 | 0.6252 | 0.6498 | 0.6481 | 0.6253 | 0.5579 | 1.0649 | 0.6205 | 0.6958 | 0.5601 |
| SROCC   | 0.8108 | 0.3032 | 0.2122 | 0.1866 | 0.1822 | 0.1853 | 0.1947 | 0.1948 | 0.1948 | 0.1948 | 0.1948 | 0.1948 | 0.1948 | 0.1948 |
| KROCC   | 0.6531 | 0.6500 | 0.7386 | 0.7022 | 0.7063 | 0.7812 | 0.8129 | 0.8137 | 0.7352 | 0.5648 | 0.7718 | 0.8171 | 0.8138 | 0.8138 |
| PLCC    | 0.4176 | 0.8612 | 0.8991 | 0.9144 | 0.9502 | 0.9620 | 0.9941 | 0.9954 | 0.9927 | 0.7843 | 0.9370 | 0.9823 | 0.9823 | 0.9823 |
| RMSE    | 0.4174 | 0.1344 | 0.1149 | 0.1063 | 0.0814 | 0.1077 | 0.0766 | 0.0979 | 0.0770 | 0.0993 | 0.1756 | 0.0938 | 0.0794 | 0.0794 |
| SROCC   | 0.8785 | 0.4990 | 0.5512 | 0.5964 | 0.5975 | 0.5610 | 0.5346 | 0.5464 | 0.5961 | 0.5710 | 0.5854 | 0.5558 | 0.5953 | 0.5973 |
| KROCC   | 0.7012 | 0.8037 | 0.8358 | 0.8379 | 0.8379 | 0.8216 | 0.8080 | 0.8350 | 0.8350 | 0.8350 | 0.8350 | 0.8350 | 0.8350 | 0.8350 |
| PLCC    | 0.5132 | 0.5985 | 0.4568 | 0.5015 | 0.4901 | 0.5361 | 0.5111 | 0.9414 | 0.5940 | 0.9723 | 0.8773 | 0.9453 | 0.9523 | 0.9514 |
| RMSE    | 1.1517 | 0.9728 | 0.7580 | 0.7269 | 0.7276 | 0.7374 | 0.7518 | 8.4504 | 0.9329 | 5.4030 | 11.0941 | 7.5418 | 7.0506 | 9.3720 |
| SROCC   | 0.8786 | 0.8468 | 0.8354 | 0.8008 | 0.8094 | 0.8054 | 0.8594 | 0.9314 | 0.9246 | 0.9246 | 0.9246 | 0.9246 | 0.9246 | 0.9246 |
| KROCC   | 0.7012 | 0.8037 | 0.8358 | 0.8379 | 0.8379 | 0.8216 | 0.8080 | 0.8350 | 0.8350 | 0.8350 | 0.8350 | 0.8350 | 0.8350 | 0.8350 |
| PLCC    | 0.5132 | 0.5985 | 0.4568 | 0.5015 | 0.4901 | 0.5361 | 0.5111 | 0.9414 | 0.5940 | 0.9723 | 0.8773 | 0.9453 | 0.9523 | 0.9514 |
| RMSE    | 1.1517 | 0.9728 | 0.7580 | 0.7269 | 0.7276 | 0.7374 | 0.7518 | 8.4504 | 0.9329 | 5.4030 | 11.0941 | 7.5418 | 7.0506 | 9.3720 |

- For each row, the first, second and third-ranked performances are highlighted respectively in blue, red and black colors.
- For SROCC, KROCC and PLCC metrics, the higher the value, the better the method whereas for RMSE a lower score is better.
Table 5. Overall performance ranking of the compared IQA methods.

| IQA         | SROCC | KROCC | PLCC | RMSE |
|-------------|-------|-------|------|------|
| PSNR        | 14    | 14    | 14   | 13   |
| SSIM        | 12    | 12    | 12   | 12   |
| MS-SSIM     | 9     | 10    | 10   | 10   |
| IW-SSIM     | 8     | 8     | 8    | 6    |
| MAD         | 10    | 9     | 9    | 8    |
| FSIM        | 7     | 7     | 7    | 7    |
| GMSD        | 4     | 4     | 4    | 5    |
| VSI         | 5     | 6     | 6    | 11   |
| MCSRD       | 3     | 3     | 3    | 3    |
| VIP         | 11    | 11    | 11   | 1    |
| NQM         | 13    | 13    | 13   | 14   |
| SRSIM       | 6     | 5     | 5    | 9    |
| VSP         | 2     | 2     | 2    | 4    |
| CEQI (Proposed) | 1    | 1    | 1    | 2    |

– Number 1 indicates the best performing method and 14 is the worst.

Table 5 compares the SROCC performance for all distortion types; please refer to Table 3 for a description of the abbreviations. We see that different methods perform better for different distortions and performance even varies between databases. This is the case because images are not affected equally by a specific type of distortion—it depends on the color, salient regions, and perhaps a combination of many other factors. Still, distortion-wise comparison gives us a good understanding of whether an IQA method is biased to some noise type or not. It can be seen that the proposed CEQI performs consistently well for all types of distortion; it is not too biased to any specific type of distortion, while retaining an average performance within the top 3 methods.

Figure 3 shows scatter plots of the predicted scores for different IQA approaches with the MOS/DMOS values on the TID2008 database. These results show that CEQI’s prediction is consistent compared to other methods, while providing a better correlation. We do not include PSNR because its predictions are very inconsistent. NQM is also discarded for the same reason, although its performance is not as inconsistent as PSNR.

Although the prime consideration of an IQA model is the performance of its prediction, having a low computational cost is also a desirable feature, especially for a real-time system. We evaluated the various IQA models with MATLAB R2017b using a computer equipped with an Intel(R) Core(TM) i5-4670 CPU with a 3.40GHz processor and 16GB of RAM. The MATLAB codes provided by the authors were used and elapsed time was recorded using the traditional tic-toc function. The results of these tests are shown in Table 7. As expected, PSNR has the lowest computation time. Surprisingly, the gradient magnitude similarity division model can process 263.05 images per second with satisfactory performance (rank 4 as shown in Table 5). VIF shows very good performance for the LIVE database where it is the best-performing IQA, but it can only process 1.79 images per second on average, which makes it inappropriate for real-time systems or systems with low processing capability. On the other hand, CEQI takes 15.25 milliseconds to process an image, with the capability of processing 65.51 images per second. This frame rate meets the need for almost all kinds of real-time systems.
Table 6. Distortion-wise SROCC performance comparison of the IQA methods on three databases.

| Dataset | Dist. | PSNR | SSIM | MS-SSIM | IW-SSIM | MAD | FSIM | GMSD | VSI | MCSD | VIF | NQM | SRSIM | VSP | CEQI (Proposed) |
|---------|-------|------|------|---------|---------|-----|------|------|-----|------|-----|-----|-------|-----|-----------------|
| LIVE    | AVG   | 0.9068 | 0.9361 | 0.9307 | 0.8546 | 0.8949 | 0.9258 | 0.8952 | 0.8846 | 0.9528 | 0.8577 | 0.9252 | 0.9544 | 0.8777 | 0.9781 |
| CTC     | GB    | 0.9807 | 0.9762 | 0.9742 | 0.9744 | 0.9509 | 0.9787 | 0.9700 | 0.9684 | 0.9903 | 0.9730 | 0.9575 | 0.9575 | 0.9550 | 0.9746 |
| GB      | AVG   | 0.9904 | 0.9798 | 0.9841 | 0.9852 | 0.9503 | 0.9789 | 0.9700 | 0.9684 | 0.9903 | 0.9730 | 0.9575 | 0.9575 | 0.9550 | 0.9746 |

Table 7. Running time comparison of the IQA models.

| IQA  | Running Time (ms) | Images Per Second |
|------|-------------------|-------------------|
| PSNR | 2.47              | 268.49            |
| SSIM | 7.47              | 133.88            |
| MS-SSIM | 25.71          | 38.90             |
| IW-SSIM | 157.85        | 6.34              |
| MAD  | 696.00            | 1.44              |
| FSIM | 115.41            | 6.66              |
| GMSD | 3.80              | 263.05            |
| VSI  | 64.12             | 15.60             |
| MCSD | 6.34              | 157.70            |
| VIF  | 560.19            | 1.79              |
| NQM  | 106.59            | 9.38              |
| SRSIM| 10.29             | 97.20             |
| VSP  | 10.75             | 93.03             |
| CEQI (Proposed) | 15.27        | 65.51             |

- For each row, the first, second and third-ranked methods are highlighted respectivley in blue, red and black colors.
- The acronyms for distortions are defined in Table 6. AVG refers to average performance over all noises in a database.
Figure 3. Scatter plots of the mean opinion scores (MOS/DMOS) versus scores predicted by different methods on the TID2008 database. The black curves are obtained by the nonlinear fitting based on Equation 1.

4 Conclusions

In this paper, we considered the center bias of HVS and proposed a full-reference image quality assessment method, CEQI, combining visual saliency and contrast. We placed extra emphasis on the center part of the image so that any degradation within the center region results in more effects than other regions. The proposed approach was compared with other state-of-the-art IQA models and it outperforms other competing methods in most cases. Comparing individual distortion types, the proposed method gives consistent scores. Additionally, the running time is suitable for real-time applications. The center emphasis makes the method more balanced and robust. We believe that this center emphasis will enhance the performance of other existing IQA models, including no-reference and reduced-reference approaches. In our future work, we will investigate these possibilities.
acknowledgments

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (No.2017-0-00294, Service mobility support distributed cloud technology)

References

1. Girod, B. Psychovisual aspects of image processing: What’s wrong with mean squared error? In Proceedings of the Seventh Workshop on IEEE Multidimensional Signal Processing, Lake Placid, NY, USA, 23-25 September 1991; p. 2.

2. Wang, Z.; Bovik, A.C. Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE Signal Process. Mag. 2009, 26, 98–117.

3. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 2004, 13, 600–612.

4. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale structural similarity for image quality assessment. In Proceedings of the Thirtieth Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, 9–12 November 2003; IEEE: New York, NY, USA, 2003; Volume 2, pp. 1398–1402.

5. Wang, Z.; Li, Q. Information content weighting for perceptual image quality assessment. IEEE Trans. Image Process. 2011, 20, 1185–1198.

6. Sheikh, H.R.; Bovik, A.C.; De Veciana, G. An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Trans. Image Process. 2005, 14, 2117–2128.

7. Sheikh, H.R.; Bovik, A.C. A visual information fidelity approach to video quality assessment. In Proceedings of the First International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Doubletree Paradise Valley Resort, Scottsdale, AZ, USA, 23–25 January 2005; pp. 23–25.

8. Larson, E.C.; Chandler, D.M. Most apparent distortion: Full-reference image quality assessment and the role of strategy. J. Electron. Imaging 2010, 19, 011006.

9. Damera-Venkata, N.; Kite, T.D.; Geisler, W.S.; Evans, B.L.; Bovik, A.C. Image quality assessment based on a degradation model. IEEE Trans. Image Process. 2000, 9, 636–650.

10. Chandler, D.M.; Hemami, S.S. VSNR: A wavelet-based visual signal-to-noise ratio for natural images. IEEE Trans. Image Process. 2007, 16, 2284–2298.

11. Zhang, L.; Zhang, L.; Mou, X.; Zhang, D. FSIM: A feature similarity index for image quality assessment. IEEE Trans. Image Process. 2011, 20, 2378–2386.

12. Kovesi, P. Image features from phase congruency. Videre 1999, 1, 1–26.

13. Liu, Z.; Laganière, R. Phase congruence measurement for image similarity assessment. Pattern Recognit. Lett. 2007, 28, 166–172.
14. Saha, A.; Wu, Q.J. Perceptual image quality assessment using phase deviation sensitive energy features. *Signal Process.* **2013**, *93*, 3182–3191.

15. Chen, G.H.; Yang, C.L.; Xie, S.L. Gradient-based structural similarity for image quality assessment. In Proceedings of the 2006 IEEE International Conference on Image Processing, Atlanta, GA, USA, 8–11 October 2006; IEEE: New York, NY, USA, 2006; pp. 2929–2932.

16. Zhu, J.; Wang, N. Image quality assessment by visual gradient similarity. *IEEE Trans. Image Process.* **2012**, *21*, 919–933.

17. Xue, W.; Zhang, L.; Mou, X.; Bovik, A.C. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Trans. Image Process.* **2014**, *23*, 684–695.

18. Alaei, A.; Conte, D.; Raveaux, R. Document image quality assessment based on improved gradient magnitude similarity deviation. In Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR), Nancy, France, 23–26 August 2015; pp. 176–180.

19. Wang, T.; Zhang, L.; Jia, H.; Li, B.; Shu, H. Multiscale contrast similarity deviation: An effective and efficient index for perceptual image quality assessment. *Signal Process.: Image Commun.* **2016**, *43*, 1–9.

20. Zhang, L.; Shen, Y.; Li, H. VSI: A visual saliency-induced index for perceptual image quality assessment. *IEEE Trans. Image Process.* **2014**, *23*, 4270–4281.

21. Ma, Q.; Zhang, L. Saliency-based image quality assessment criterion. In *International Conference on Intelligent Computing*; Springer: Berlin, Heidelberg, Germany, 2008; pp. 1124–1133.

22. Hou, X.; Zhang, L. Saliency detection: A spectral residual approach. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2007 (CVPR’07), Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8.

23. Duan, L.; Wu, C.; Miao, J.; Qing, L.; Fu, Y. Visual saliency detection by spatially weighted dissimilarity. In Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20–25 June 2011; pp. 473–480.

24. Zhang, L.; Li, H. SR-SIM: A fast and high performance IQA index based on spectral residual. In Proceedings of the 2012 19th IEEE International Conference on Image Processing (ICIP), Orland, FL, USA, 30 September–3 October 2012; pp. 1473–1476.

25. Li, A.; She, X.; Sun, Q. Color image quality assessment combining saliency and fsim. In Proceedings of the Fifth International Conference on Digital Image Processing (ICDIP 2013), Beijing, China, 21–22 April 2013; International Society for Optics and Photonics: Bellingham WA 98225-6705 USA, 2013; Volume 8878, p. 88780F.

26. Jia, H.; Zhang, L.; Wang, T. Contrast and Visual Saliency Similarity-Induced Index for Assessing Image Quality. *IEEE Access* **2018**, *6*, 65885–65893.

27. Langford, R.C. How People Look at Pictures, A Study of the Psychology of Perception in Art. *Journal of Educational Psychology* 1936, *27*(5), 397-398.
28. Mannan, S.; Ruddock, K.; Wooding, D. Fixation sequences made during visual examination of briefly presented 2D images. *Spat. Vis.* **1997**, *11*, 157–178.

29. Parkhurst, D.; Law, K.; Niebur, E. Modeling the role of salience in the allocation of overt visual attention. *Vis. Res.* **2002**, *42*, 107–123.

30. Tatler, B.W. The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *J. Vis.* **2007**, *7*, 4.

31. Bindemann, M. Scene and screen center bias early eye movements in scene viewing. *Vis. Res.* **2010**, *50*, 2577–2587.

32. Cong, R.; Lei, J.; Fu, H.; Cheng, M.M.; Lin, W.; Huang, Q. Review of Visual Saliency Detection with Comprehensive Information. *arXiv 2018*, arXiv:1803.03391.

33. Peli, E. Contrast in complex images. *JOSA A* **1990**, *7*, 2032–2040.

34. Ding, Y. General Framework of Image Quality Assessment. In *Visual Quality Assessment for Natural and Medical Image*; Springer: Berlin, Heidelberg, Germany, 2018; pp. 45–62.

35. Sheikh, H.R.; Sabir, M.F.; Bovik, A.C. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on image processing* **2006**, *15*, 3440–3451.

36. Ponomarenko, N.; Lukin, V.; Zelensky, A.; Egiazarian, K.; Carli, M.; Battisti, F. TID2008—A database for evaluation of full-reference visual quality assessment metrics. *Adv. Modern Radioelectron.* **2009**, *10*, 30–45.

37. Larson, E.C.; Chandler, D. Categorical image quality (CSIQ) database, 2010. Available online: [http://vision.eng.shizuoka.ac.jp/](http://vision.eng.shizuoka.ac.jp/) (accessed on: 23 December 2018).

38. Sheikh, H. LIVE Image Quality Assessment Database Release 2, 2005. Available online: [http://live.ece.utexas.edu/research/quality](http://live.ece.utexas.edu/research/quality) (accessed on: 20 December 2018).