Jagannath University Journal of Science
Volume 07, Number II, June 2021, pp. 7–21
https://jnu.ac.bd/journal/portal/archives/science.jsp
ISSN 2224-1698

JnUJSci

Center Emphasized Structural Contrast and a Saliency-induced Image Quality Index

Research Article

Md. Abu Layek[1*], Md. Zulfiker Mahmud[1], Md Rafiqul Islam[1], Mst Sirazam Monira[2], A F M Shahab Uddin[2], Sanjida Afroz[3]

[1]*Department of Computer Science and Engineering, Jagannath University, Dhaka-1100, Bangladesh;*
[2]*Department of Computer Science and Engineering, Kyung Hee University Global Campus, Yongin-17104, Republic of Korea;*
[3]*Mental Health Counselor, Labaid, Dhaka, Bangladesh.*

Received: 01 December 2020
Accepted: 15 February 2021

Abstract: Image quality assessment (IQA) has caught researchers' attention for decades due to its inevitable importance to assess an image's visual quality close to the human ability. In the course of time, several methods have been devised which uses different features such as luminance, contrast, structure or saliency whereas some recent approaches combine one or more features to get better performance. Psychological research advocates that the human visual system (HVS) is biased to the center part of a scene and display screen. Any kind of distortions occupying the center area is perceived intensely by human observer than other areas, especially, if the center area contains any visually important information. However, current state-of-the-art IQA methods do not consider this center bias. In this paper, at first, we derive a full reference image quality assessment method 'Center emphasized Structural Contrast-induced image Quality Index (CSCQI)' by modifying only the center part of the structural contrast map. Then, we obtain the 'Saliency and Structural Contrast-induced image Quality Index (SSCQI)' combining spectral residual visual saliency with the structural contrast, and finally, we propose the 'Center emphasized Saliency and Structural Contrast-induced image Quality Index (CSSCQI)' using structural contrast with visual saliency and modifying center areas for both of the similarity maps to increase the distortion sensitivities there. For the latter two methods, the final score is calculated using a novel mixed-mode pooling approach 'summation of weighted mean and standard deviation'. Evaluations on four large-scale benchmark databases (TID2013, TID2008, CSIQ and LIVE) and comparison with 13 state-of-the-art methods reveal the competitiveness of the proposed approaches. The MATLAB code is publicly available online to test the algorithms and can be found at this Link: http://layek.khu.ac.kr/CSSCQI.

Keywords: *Image Quality Assessment ● Structural Contrast ● Visual Saliency ● Full-Reference ● Center Emphasized*

1. Introduction

With the growing graph of digital vision technologies, the Image Quality Assessment (IQA) becomes a crucial part of many image processing applications, such as image restoration, compression, transmission, super-resolution and the like. In real life applications, an image has to pass through a pipeline of processing stages and as a result, the original image can get distorted easily and exhibits a certain level of annoyance. So, IQA comes into the scene and plays an important role to measure the quality of the degraded

* Corresponding author: Md. Abu Layek
E-mail: layek@khu.ac.kr

images. As humans are the ultimate viewer of the images, the subjective evaluation by humans is the precise way to quantify the visual quality of images, but it is very expensive, cumbersome and time-consuming. For a very long time, many researchers are focused to develop a metric which can automatically predict the perceived quality of an image, known as objective image quality assessment metric. The IQA methods are being developed by considering that, the results of these approaches should be consistent statistically with the perceived quality of human observers.

Full-reference (FR), reduced-reference (RR) and no-reference (NR) are the three existing approaches of the objective image quality assessment. If the complete reference image is available, then the approach is full-reference image quality assessment (FR-IQA), and if the reference image is partially available or some extracted features are available as an information then it is called reduced-reference image quality assessment (RR-IQA) and finally in many cases the reference image is not available then it is no-reference image quality assessment (NR-IQA). The center of attention of this paper is FR-IQA.

The traditional and simplest full reference image quality metric is the peak signal to noise ratio (PSNR) or mean square error (MSE) (Wang and Bovik, 2009). They are very simple to calculate but do not correlate well with human perception. As a result, the search for better and close to human level IQA is ever expected, to this date, many sophisticated IQA methods have already been proposed. Emphasizing the sensitivity of HVS's to different visual signals such as contrast, luminance, frequency content and the interaction between them, some methods were proposed such as the visual signal-to-noise ratio index (VSNR) (Chandler and Hemami, 2007) and the noise quality measure index (NQM) (Damera-Venkata *et al*., 2000). But those error visibility methods ignored the important characteristics of HVS and resulted in a poor correlation. Wang *et al.* (Wang *et al*., 2004) proposed the structural similarity index (SSIM), which is a remarkable turning in the IQA research. The main motivation of SSIM is that the HVS is highly adapted to extract the structural information from the visual field. So, the structural similarity measurement can be a good estimation of the perceptual image quality. In their later work, Wang et al. proposed the multi-scale extension of SSIM (MS-SSIM) (Wang *et al*., 2003) which produced a better result than SSIM. Wang and Li proposed the IW-SSIM (Wang and Li, 2011) by introducing information content extraction and information content weighting based pooling strategy. Sheikh et al. proposed the information fidelity criterion (IFC) (Sheikh *et al*., 2005) by quantifying the information shared between the reference and distorted images. An extended version of IFC is the visual information fidelity index (VIF) (Sheikh and Bovik, 2006). IFC and VIF treat HVS as a communication channel and these methods decompose an image into different sub-bands that have different weights at the pooling stage. The salient low-level features fetch vital information to interpret the scene, based on this consideration Zhang et al. proposed the feature similarity index (FSIM) (Zhang *et al*., 2011). FSIM employs two features, phase congruency and gradient magnitude to compute the local similarity map then phase congruency map is again used as the weighting function. The image gradients are sensitive to image distortions, based on this criterion Xue et al. proposed the gradient magnitude similarity deviation (GMSD) (Xue *et al*., 2014) by introducing a novel standard-deviation based pooling strategy. The success of gradient magnitude and standard deviation pooling inspired Nafchi et al. to propose Mean Deviation Similarity Index (MDSI) (Nafchi *et al*., 2016), however, they modified the gradient similarity map through a fusion technique. The multi-scale contrast similarity deviation (MCSD) (Wang *et al*., 2016) is proposed by Wang et al. also uses the root mean square (RMS) contrast similar to SSIM but employing standard deviation pooling for the final score.

Visual saliency detection finds out the most attractive regions in an image which is a similar task as IQA, as a result, by incorporating visual saliency (VS) with IQA methods can improve the performance (Hou and Zhang, 2007; Ma and Zhang, 2008; Duan *et al*., 2011; Zhang *et al*., 2012). The Spectral Residual-based Similarity Index (SR-SIM) (Zhang and Li, 2012) combined the spectral residual visual saliency with image gradient to model HVS in a better way whereas Zhang et al. proposed the visual saliency index (VSI) (Zhang, Shen and Li, 2014) combining visual saliency (VS) with the gradient magnitude and weighted by VS at the pooling stage. Bae and Kim proposed the structural contrast quality index (SCQI) (Bae and Kim, 2016) using multilevel contrast, structure and chrominance information which can characterize both the local and global perceptual visual qualities. Wang et al. proposed a local linear model (LLM) (Wang *et al*., 2017) based integrated IQA in combination with a distortion-specific compensation strategy using a convolutional neural network. Wavelet-based IQA approaches also common, Reisenhofer et al. proposed a Haar wavelet-based perceptual similarity index (HPSI) (Reisenhofer *et al*., 2018) utilizing the coefficients obtained from a Haar wavelet decomposition to assess local similarities between two images. Combining two or more features and finding quality score through a final pooling stage also become popular (Li, She and Sun, 2013; Jia *et al*., 2018), Li et al. proposed an approach by combining VS and FSIM while

Jia et al. use contrast and spectral residual saliency as well as standard deviation pooling (Jia *et al*., 2018).

To this end, in designing an IQA, no one considered the center bias in early eye movements which is already known from psychological vision research (Langford 1936; Mannan *et al*., 1997; Parkhurst *et al*., 1997; Tatler, 2007). The experiment of Bindemann reveals that eye movement is biased to the scene center as well as to the center of the display screen. Hence, a scene appearing at the center of screen gets the most attention and presence of any distortion in that area caught by the human eye more intensely. Figure 1 shows the 'Monacrh.bmp' image file from the LIVE database where the right part is showing the extracted center area. The human eye will first move to center-area and in this example, the most salient region has also involved the butterfly, as a result, people will find any kind of distortion easily there. Recently, we proposed the center emphasized quality index (CEQI) (Layek *et al*., 2019), where we combined contrast and spectral residual saliency, and finally increase the distortion sensitivity at center region. However, in that work, we did not consider color information. In this paper, we proposed three IQA approaches by combining structural contrast with visual saliency and considering center importance. We adopted the structural contrast quality assessment as the base method because it considers luminance, chrominance as well as the structural contrast. Also, the clarity of their implementation enables us to implement our ideas on top of it. Significant improvement, as well as competitive-ness with state-of-the methods, indicates the effectiveness of our proposed methods. We can list the contributions of this paper as below, details about the implementations are given in section 2:

(i) First, we propose the 'Center emphasized Saliency and Structural Contrast-induced image Quality Index (CSCQI)' by modifying only the center part of the structural contrast map. Similarity maps are usually a down-scaled matrix relative to the reference and distorted images where a value 1 represents exactly similar, 0 as totally dissimilar, a value between 0 and 1 represents the degree of similarity. When a person assesses a distorted image with respect to a reference, he/she actually search-for the dissimilar areas in the distorted image. To incorporate our center-emphasize idea, we apply a simple element-wise square to represent the dissimilar areas as more contrasting.

(ii) The second contribution of this paper is the successful merging of Spectral Residual Visual Saliency(SRS) with SCQI to derive a new and improved IQA i.e. 'Saliency and Structural Contrast-induced image Quality Index (SSCQI)'. After obtaining both structural contrast and spectral residual saliency similarity maps, the final score is calculated using a novel mixed-mode pooling approach. In this case, we do not give any special importance to the center part.

(iii) Thirdly, we propose the 'Center emphasized Saliency and Structural Contrast-induced image Quality Index (CSSCQI)' using structural contrast with visual saliency and modifying center areas of both of the similarity maps to increase the distortion sensitivities there. Visual saliency detection is dependent on the current view of the image thus we compute the visual saliencies and the similarity map of the full and center image separately.

(iv) Finally, we propose a novel mixed-mode pooling approach which we utilize in the second and third methods (SSCQI, CSSCQI). The study in GMSD unveiled that standard deviation is not suitable for SSIM, MS-SSIM or FSIM, also the authors of SCQI used weighted average and achieved quite good results. On the other hand, SD pooling has successfully employed by several papers with the saliency maps. As a result, we apply a weighted average in SCQI map and SD pooling for saliency map and finally compute the weighted sum to obtain the final score.

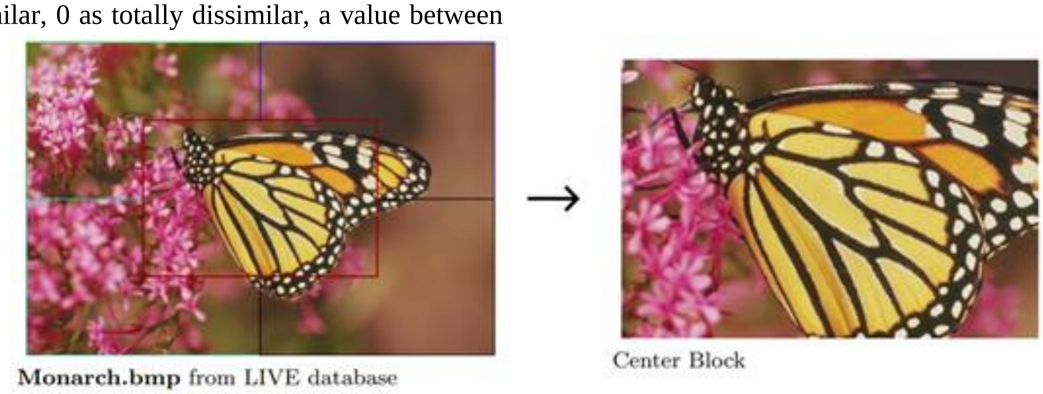


Figure 1. The image Monarch.bmp and its' center area

We evaluated our proposed methods on four large-scale popular benchmark databases for IQA research and compared with 13 other state-of-the-art methods. Results show that the techniques proposed by us outperform other comparing approaches. Adding visual saliency improved the correlation of predicted score with the human evaluated values and center emphasis boosts-up the performance with or without saliency.

The paper is organized as follows. Section 1 describes theories and related techniques. Section 2 explains the proposed IQA approaches along with the novel mixed-mode pooling strategy, and the results with relevant discussions are presented in Section 3. Finally, we conclude the paper in Section 4.

1. Background

In this section, we briefly review the related theories on which the content of this paper relies; the structural contrast, spectral residual visual saliency, and the evaluation metrics.

1.1 Structural Contrast Quality Index

Contrast is a basic perceptual attribute of an image which varies greatly over the image [29]. Bae et al. [18] proposed structural contrast quality index (SCQI) by adopting the structural contrast index (SCI) [30], that can estimate the perceptual complexity of image texture patterns as the ratio of structureness and contrast intensity. Here, we briefly describe the SCQI in an independent manner, more details are given in section 2 and in Figure 2, further details are in the reference SCQI article [18].

SCQI gives a quality index value which is calculated as

$$SCQI = \sum \left[\frac{1}{W} \sum_{i=1}^{K} w_i \odot SCQI_{Map} \right], \qquad (1)$$

where K is the number of items in the $SCQI_{Map}$,w_i is the local weight, W is the summation of all weights, and $SCQI_{Map}$ is the similarity map calculated by element- wise multiplication of six similarity maps sm_i,i=1,...,6as

$$SCQI_{Map} = \prod_{i=1}^{6} sm_i. \qquad (2)$$

The six similarity maps are based on six features; two chrominance, three contrast sensitivity function (CSF) on contrast energy frequencies (low, mid and high), and the structural contrast. However, SCQI uses the inverse value of structural contrast index than proposed in [30] $(SCI = \tau^{-1})$ so that more distortion-sensitive image texture regions have higher SCI values that are more important regions to HVS.

All of the similarity maps are measured in the same way. If the reference image is R, the distorted image in the question is D, and corresponding feature matrices are $fm_r(i)$ and f md (i) respectively, then the similarity measure sm_i is given as

$$sm_i(R, D) = \frac{(2fm_r(i).fm_d(i) + c_i)}{(fm_r(i)^2 + fm_d(i)^2 + c_i)}, \qquad (3)$$

where c_i are six corresponding positive constants to increase calculation stability.

Now, let's take a look at how the feature matrices are obtained.

At first, the images are converted to LMN color space from RGB to separate luminance (L) and two chromniance (M, N) components using following relationship

$$\begin{bmatrix} L \\ M \\ N \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \qquad (4)$$

Using the above equation, we get the *M and N* Chroma matrices for both reference and distorted images and compute the Chroma maps using equation 3.

The *SCI* and three *CSF* features are derived from the luminance. *CSF* matrices are calculated as below

$$csf(k) = \sum_{(u,v)\epsilon R_k} p(u, v), \qquad (5)$$

where, $p(u, v)$ is the normalized magnitude of a DCT coefficient at (u, v), k is one of the low, mid or high and R_k are the corresponding ranges.

The structural contrast τ is given as,

$$\tau = CI^{\alpha}/TP^{\beta}, \qquad (6)$$

where α and β are model parameters, CI is the contrast intensity, and TP is the structureness reflecting the randomness of texture patterns given by

$$TP = m_4/(m_2)^2, \qquad (7)$$

where m_k is the k-th moment of normalized DCT AC coefficients, and is defined by

$$m_k = \sum_{\omega \in B, \omega \neq 0} \omega^k p(\omega), \qquad (8)$$

where ω is a spatial frequency value in cycles per degree (cpd) for *(u,v)-th* DCT coefficient and is calculated by

$$\omega = \delta.\sqrt{u^2 + v^2}, \tag{9}$$

where δ is a const ant and $p(\omega)$ is the magnitude of a normalized DCT coefficient at ω and is defined as

$$p(\omega) = (\epsilon + |c(\omega)|^{\lambda})/Z, \tag{10}$$

where $c(\omega)$ is the DCT coefficient value at ω, ϵ is a very small constant value to avoid unstable results when the denominator is close to zero, λ is an adjustment parameter to fit the measured experimental results, and Z is a normalization factor given by

$$Z = \sum_{\omega \in B, \omega \neq 0} (\epsilon + |c(\omega)|^{\lambda}). \tag{11}$$

The contrast intensity in 6 is defined as

$$CI = m_0/N^2, \tag{12}$$

where N is the height or width of a $N \times N$ DCT block.

We adopt the same model parameters as it is used in *SCQI* which are set to ϵ=0.25, λ=1, α=1, and β=1. With these values, *SCI* is simplified as below

$$SCI = \frac{\sum_{(u,v)\in B}\{(u^2+v^2)^2 \odot (\epsilon + |c(u,v)|)\}}{\sum_{(u,v)\in B}\{(u^2+v^2)^2 \odot (\epsilon + |c(u,v)|)\}^2}, \tag{13}$$

where (u,v) is the (u,v)-th DCT coefficient value.

1.2 Spectral Residual Visual Saliency Similarity

In any image, salient regions caught more attention than other areas and detection of those parts is called saliency detection. Therefore, the human is more sensitive to the distortions in those interesting and salient regions than other parts. Any distortion in these parts attract intense attention, which makes it an important feature for IQA, as a result, many IQA researchers utilize visual saliency as an important feature. Spectral residual saliency detection [13] is a very fast approach among the various detection techniques [31]. We adopt the image saliency map generator as described in the SR-SIM [16].

For an image f (x, y), the spectral residual saliency (*SRS*) is computed as follows:

$$M(u,v) = abs[\mathscr{F}\{f(x,y)\}(u,v)] \tag{14}$$

$$A(u,v) = angle[\mathscr{F}\{f(x,y)\}(u,v)] \tag{15}$$

$$L(u,v) = \log\{M(u,v)\} \tag{16}$$

$$R(u,v) = L(u,v) - h_n(u,v) * L(u,v) \tag{17}$$

$$SRS(x,y) = g(x,y) * [\mathscr{F}^{-1}\{\exp(R + jA)\}(x,y)]^2, \tag{18}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ are the Fourier transform and the inverse; *abs*(.) return the magnitude and *angle*(.) returns the argument of a complex number. *(x,y)* is a Gaussian filter; $h_n(u,v)$ is an $n \times n$ mean filter; and $*$ denotes the convolution operation.

Using equations 14–18, we calculate the spectral residual saliencies for both the reference and distorted images denoted by $SRS_r(x,y)$ and $SRS_d(x,y)$, respectively. Then, the saliency similarity map $SRS_{Map}(r,d)$ is calculated as:

$$SRS_{Map}(r,d) = \frac{2SRS_r(x,y) \odot SRS_d(x,y) + c_1}{SRS_r(x,y)^2 + SRS_d(x,y)^2 + c_1}, \tag{19}$$

where ^{2} is the element-wise squaring, $\odot$ refers to element-wise multiplication, and c_1 is a positive constant for calculation stability.

1.3 Evaluation Metrics

We can measure the performance of a IQA method using some correlation measurements with respect to the human evaluated subjective scores, the root mean square error (RMSE) is also used. Before applying the linear correlation, the two compared values should be on the same scale and perfectly linearly correlated [32]. For this purpose, a logistic mapping function is used to convert the objective scores before applying the linear correlation measurements. We adopt the following nonlinear regression model as suggested by Sheikh [33].

$$q' = \beta_1\left\{\frac{1}{2} - \frac{1}{1+\exp(\beta_2(q-\beta_3))}\right\} + \beta_4 q + \beta_5, \tag{20}$$

where q is the objective score, q', is the mapped value, and β_i are the 5 parameters that are tuned based on the relationship between objective and subjective scores. To find the optimal parameters, we utilized the *nlinfit* function that is already built within MATLAB. The subjective scores are then used with these mapped scores to find the following correlation coefficient.

The Pearson's linear correlation coefficient (PLCC) is defined as follows:

$$PLCC(o,s) = \frac{\sum_{i=1}^{m}(o_i - \mu_o)(s_i - \mu_s)}{\left\{\sum_{i=1}^{m}(o_i - \mu_o)^2\right\}^{\frac{1}{2}}\left\{\sum_{i=1}^{m}(s_i - \mu_s)^2\right\}^{\frac{1}{2}}}, \tag{21}$$

where o and s are vectors of the objective and subjective scores, respectively; μ_o and μ_s are their mean scores; and m is the number of distorted images. The objective scores of o are actually the mapped scores using Equation (20). If we want to avoid the nonlinear mapping in Equation (20), rank order coefficients can be used. The popular

Spearman's rank-order correlation coefficient (SROCC) is given as:

$$SROCC(o,s) = PLCC(rank(o), rank(s)). \quad (22)$$

Applying the *rank()* function on a score vector returns a rank-vector where the i-th entry contains the relative rank of the i-th item in the score vector. Another popular rank order metric is the kendall's rank-order correlation coefficient (KROCC), which is given as below:

$$KROCC(o,s) = \frac{C_p - D_p}{m(m-1)/2}, \quad (23)$$

Where C_p and D_p are the number of concordant and discordant pairs.

The root mean square error (RMSE) is defined as:

$$RMSE(o,s) = \left\{ \frac{1}{m} \sum_{i=1}^{m} (o_i - s_i)^2 \right\}^{\frac{1}{2}} \quad (24)$$

A larger value of PLCC, SROCC, and KROCC is an indicator of superior method whereas in the case of RMSE smaller is better IQA. Again, SROCC is treated as the most important correlation measurement among these metrics.

2. Proposed Assessment Methods

The complete flow diagram of our proposed 'Center emphasized Saliency and Structural Contrast-induced image Quality Index (CSSCQI)' is presented in Figure 2. Before going detail to the final system we discuss the other two methods with the help of the same figure.

1.4 Center emphasized Structural Contrast induced image Quality Index (CSCQI)

To derive this method, we just modify the center part of the $SCQI_{Map}$ which we get by applying equations (1)– (13). We define the center region for both images and similarity maps as follows:

1.5 Saliency and Structural Contrast induced image Quality Index (SSCQI)

For the original dimension of *(H × W)*, the corresponding dimension of the center block becomes *(H_{mid} × W_{mid})*, where:

$$H_{mid} = \left\lceil \frac{H}{2} \right\rceil \quad \text{and} \quad W_{mid} = \left\lceil \frac{W}{2} \right\rceil. \quad (25)$$

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

$$x_{min} = \left\lceil \frac{H}{4} \right\rceil, \; y_{min} = \left\lceil \frac{W}{4} \right\rceil,$$

$$x_{max} = \left\lceil \frac{H}{4} \right\rceil + H_{mid}, \quad \text{and} \quad y_{max} = \left\lceil \frac{W}{4} \right\rceil + W_{mid} \quad (26)$$

If we denote the center part of the $SCQI_{Map}$ as $SCQI_{Map}(mid)$, then the updated center part $CSCQI_{Map}(mid)$ is defined as

$$CSCQI_{Map}(mid) = SCQI_{Map}(mid) \odot SCQI_{Map}(mid), \quad (27)$$

where ʘ is the element-wise multiplication.

With the updated center region, we get the center emphasized $SCQI_{Map}$ denoted as $CSCQI_{Map}$. After that, using equation 1we compute the final $CSCQI$ value as

$$CSCQI = \sum \left[\frac{1}{W} \sum_{i=1}^{K} w_i \odot CSCQI_{Map} \right]. \quad (28)$$

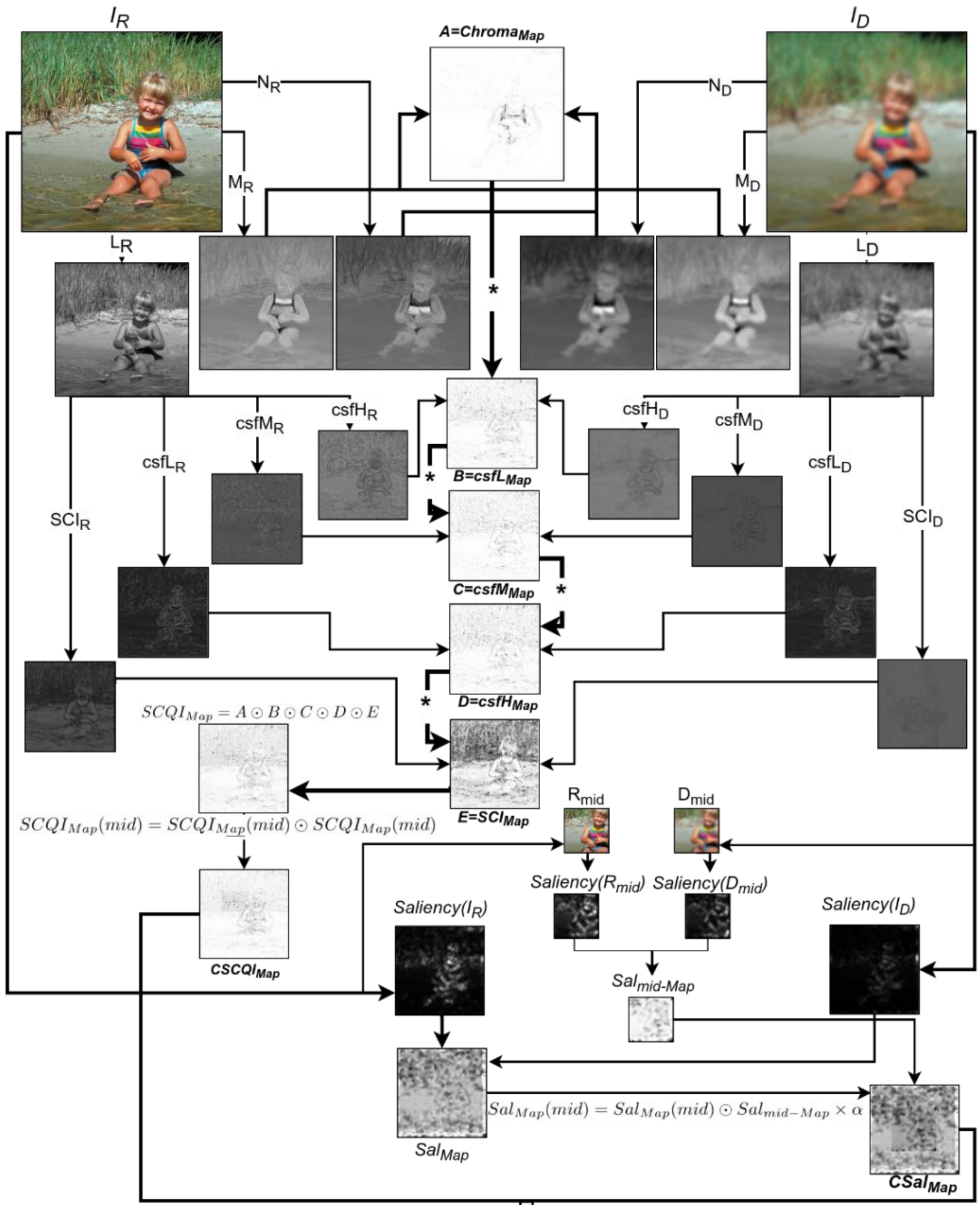
Neither SCQI nor the CSCQI discussed above include visual saliency for quality assessment. As discussed in the introduction, visual saliency is directly related to HVS and several other works used visual saliency with a success. In our recent paper, we combined spectral residual saliency with RMS contrast which exhibits very good performance [28]. As a result, in this work, we incorporate the spectral residual visual saliency with SCQI which again gives satisfactory improvement compared to SCQI as shown in section 3.

Here, we do not incorporate our center-emphasized idea. The $SCQI_{Map}$ is computed using equation 2 and the saliency similarity map Sal_{Map} is computed using equations 14 – 19.

Finally, we apply a mixed-mode pooling strategy which is discussed later in section 2.4, on both $SCQI_{Map}$ and Sal_{Map} using Equations 29, which gives us the final quality score $SSCQI$ *as*:

$$SSCQI = W_1 \times \sum \left[\frac{1}{W} \sum_{i=1}^{K} w_i \odot SCQI_{Map} \right] + W_2 \times \{1 - stdev(Sal_{Map})\} \quad (29)$$

where K is the number of items in the $SCQI_{Map}$,w_i is the local weight, W is the summation of all w_i,W_1 and W_2 are positive weighting factors (W_1+W_2=1) which specify the importance between saliency and structural contrast.



$$CSSCQI = W_1 \times \sum \left[\frac{1}{W} \sum_{i=1}^{K} w_i \odot CSCQI_{Map} \right] + W_2 \times \{1 - stdev(CSal_{Map})\}$$

Figure 2. Flow diagram of quality index calculation in the proposed center emphasized approach.

It is to be noted that, in the second part of the equation, the standard deviation of saliency map *stdev*($CSal_{Map}$) is subtracted from 1 because a higher *SCQI* score refers to better similarity whereas the value of *stdev* ($CSal_{Map}$) bears the opposite meaning.

1.6 Center emphasized Saliency and Structural Contrast induced image Quality Index (CSCQI)

This is the final proposed method where we combine visual saliency and SCQI emphasizing the center part for both of the similarity maps. The full process is described as a flowchart in Figure 2.

First, the saliency similarity maps for the full images and middle images are found using Equations (14)–(19) and are denoted as Sal_{Map} and Sal_{mid}-Map, respectively. Then, we increase the sensitivity of the center area within the Sal_{Map}. If the center area of the full saliency map is $Sal_{Map}(mid)$, then it is updated to $CSal_{Map}(mid)$:

$$CSal_{Map}(mid) = Sal_{Map}(mid)) \odot Sal_{mid-Map} \quad (30)$$

With the updated center part, we obtain the center-emphasized saliency map denoted by $CSal_{Map}$. Simultaneously, the $CSCQI_{Map}$ is also calculated using equation 28 in the same way as in section 2.1. Sal_{Map} computes the relative importance within the whole image, as a result, deriving saliency from middle image separately gives us fine-grained feature. In contrast, structural contrast is a local feature and thus we do not derive the $SCQI_{Map}$ for middle images.

Finally, in this case, we again apply the mixed-mode pooling strategy as discussed in section 2.4 on both $CSCQI_{Map}$ and $CSal_{Map}$ using Equation 31, which gives us the final quality score $CSSCQ$I as:

$$CSSCQI = W_1 \times \sum \left[\frac{1}{W} \sum_{i=1}^{K} w_i \odot CSCQI_{Map} \right] + W_2 \times \{1 - stdev(CSal_{Map})\} \quad (31)$$

where the parameters bear similar meaning as in equation 29.

1.7 The Mixed-mode Pooling Strategy

As we have mentioned 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 with the spectral residual saliency [22]. However, with SCQI, as we also found in our experiments, SD pooling is not giving satisfactory performance. As a result, in this paper, we propose a novel mixed-mode pooling strategy the 'summation of weighted mean and standard deviation' combining both weighted average and standard deviation and final quality score is obtained by summing-up them. If two feature similarity maps are *FSM1* and *FSM2* where *FSM1* is weighted average pooling friendly and *FSM2* performs better with SD pooling, then using mixed-mode pooling we can compute the quality score as:

$$QS = W_1 \times \sum \left[\frac{1}{W} \sum_{i=1}^{K} w_i \odot FSM1 \right] + W_2 \times \{stdev(FSM2)\} \quad (32)$$

Where K is the number of items in the FSM1, w_i is the local weight, W is the summation of all w_i, W_1 and W_2 are positive weighting factors (W_1+W_2=1) which specify the importance between saliency and structural contrast. The standard deviation in the above equation is defined as:

$$std(FSM2) = \left\{ \frac{1}{M} \sum_{i=1}^{M} (FSM2_i - \mu_{FSM2})^2 \right\}^{\frac{1}{2}} \quad (33)$$

where M is the number of total elements in the similarity matrix; $FSM2_i$ is the *ith* item; $\mu FSM2$ is the mean value of the $FSM2$ and is given by:

$$\mu_{FSM2} = \frac{1}{M} \sum_{i=1}^{M} FSM2_i \quad (34)$$

3. Results and Analysis

Experiments were carried out on four popular benchmark databases for IQA research TID2013[34], TID2008[35], CSIQ[36] and LIVE[37]. Our approach was compared with 13 other state-of-the-art recent IQA methods namely SSIM[4], MS-SSIM[5], IW-SSIM[6], MAD[38], FSIMc[9], GMSD[10], MCSD[12], VIF[38],

VSI[15], HPSIe[20], MDSI[11],LLM[19] and SCQI[18]. CEQI[28], SCQI and the proposed approaches are again compared in Table 4 to clearly illustrate the effectiveness of center emphasizing and the mixed-mode pooling strategy. Basic information about the databases is given in Table 1 and the distortion information is recorded in Table 2.

Dataset	Reference Images	Distorted Images	Distortion Types	No. of Subjects
TID2013	25	3000	24	971
TID2008	25	1700	17	838
CSIQ	30	866	6	35
LIVE	29	779	5	161

Table 1. Basic information about the databases used for experiments.

TID2013	TID2008	CSIQ	LIVE	Type of distortion	Abbreviation
Y	Y	Y	Y	Additive Gaussian noise	AGN
Y	Y	-	-	Additive noise in color components	ANC
Y	Y	-	-	Spatially correlated noise	SCN
Y	Y	-	-	Masked noise	MN
Y	Y	-	-	High frequency noise	HFN
Y	Y	-	-	Impulse noise	IN
Y	Y	-	-	Quantization noise	QN
Y	Y	Y	Y	Gaussian blur	GB
Y	Y	-	-	Image denoising	IDN
Y	Y	Y	Y	JPEG compression	JPEG
Y	Y	Y	Y	JPEG2000 compression	JP2K
Y	Y	-	-	JPEG transmission errors	JGTE
Y	Y	-	-	JPEG2000 transmission errors	J2TE
Y	Y	-	-	Non-eccentricity pattern noise	NEPN
Y	Y	-	-	Local block-wise distortions of different intensity	LBD
Y	Y	-	-	Mean shift (intensity shift)	MS
Y	Y	Y	-	Contrast change	CTC
Y	-	-	-	Change of color saturation	CCS
Y	-	-	-	Multiplicative Gaussian noise	MGN
Y	-	-	-	Comfort noise	CMN
Y	-	-	-	Lossy compression of noisy images	LCN
Y	-	-	-	Image color quantization with dither	ICQ
Y	-	-	-	Chromatic aberrations	CA
Y	-	-	-	Sparse sampling and reconstruction	SS
-	-	-	Y	Fast fading Rayleigh	FF
-	-	Y	-	Additive pink Gaussian noise	AWPN

Table 2. Description databases diving into the types of distortions used

Dataset	Metric	SSIM	MS-SSIM	IW-SSIM	MAD	FSIMc	GMSD	MCSD	VIF	VSI	HPSIe	MDSI	LLM	SCQI	CSCQI	SSCQI	CSSCQI
TID 2013	SROCC	0.7417	0.7859	0.7779	0.8086	0.8510	0.8044	0.8089	0.6679	0.8965	0.8732	0.8899	0.9037	**0.9052**	**0.9058**	0.9045	**0.9057**
	KROCC	0.5588	0.6047	0.5977	0.6236	0.6665	0.6339	0.6385	0.5067	0.7183	0.6923	0.7123	0.7209	**0.7327**	**0.7337**	0.7323	**0.7346**
	PLCC	0.7895	0.8329	0.8319	0.8267	0.8769	0.8590	0.8648	0.7720	0.9000	0.8935	**0.9085**	0.9068	0.9071	0.9075	**0.9082**	**0.9087**
	RMSE	0.7608	0.6861	0.6880	0.6975	0.5959	0.6346	0.6225	0.7880	0.5404	0.5568	**0.5181**	0.5277	0.5219	0.5207	**0.5187**	**0.5175**
TID 2008	SROCC	0.7749	0.8542	0.8559	0.8340	0.8840	0.8907	0.8911	0.7491	0.8979	**0.9104**	**0.9207**	0.9077	0.9051	0.9072	0.9076	**0.9088**
	KROCC	0.5768	0.6568	0.6636	0.6445	0.6991	0.7092	0.7133	0.5860	0.7123	**0.7373**	**0.7513**	**0.7368**	0.7294	0.7320	0.7329	0.7357
	PLCC	0.7732	0.8451	0.8579	0.8306	0.8762	0.8788	0.8844	0.8084	0.8762	**0.9067**	**0.9160**	**0.8971**	0.8899	0.8917	0.8933	0.8941
	RMSE	0.8511	0.7173	0.6895	0.7473	0.6468	0.6404	0.6263	0.7899	0.6466	**0.5661**	**0.5383**	**0.5982**	0.6120	0.6073	0.6030	0.6010
CSIQ	SROCC	0.8755	0.9132	0.9212	0.9466	0.9309	**0.9570**	**0.9592**	0.9194	0.9422	**0.9603**	0.9568	0.9050	0.9434	0.9442	0.9497	0.9473
	KROCC	0.6900	0.7386	0.7522	0.7963	0.7684	**0.8122**	**0.8171**	0.7532	0.7850	**0.8234**	**0.8122**	0.7238	0.7863	0.7883	0.7974	0.7937
	PLCC	0.8612	0.8991	0.9144	0.9502	0.9191	**0.9541**	**0.9560**	0.9257	0.9279	**0.9580**	0.9531	0.9000	0.9268	0.9278	0.9373	0.9338
	RMSE	0.1334	0.1149	0.1063	0.0818	0.1034	**0.0786**	**0.0770**	0.0993	0.0979	**0.0753**	0.0795	0.1232	0.0986	0.0979	0.0915	0.0939
LIVE	SROCC	0.9460	0.9512	**0.9604**	0.9567	0.9599	0.9546	0.9603	**0.9719**	0.9464	0.9585	0.9577	**0.9608**	0.9438	0.9440	0.9460	0.9455
	KROCC	0.8057	0.8181	**0.8379**	0.8290	**0.8366**	0.8236	0.8350	**0.8571**	0.8000	0.8242	0.8194	0.8230	0.7929	0.7932	0.7968	0.7961
	PLCC	0.9385	0.9468	0.9515	0.9493	0.9503	0.9511	0.9540	**0.9723**	0.9431	**0.9601**	**0.9687**	0.9578	0.9373	0.9373	0.9386	0.9379
	RMSE	7.9838	7.4380	7.1116	7.2690	7.2002	7.1374	6.9329	**5.4030**	7.6856	**6.4675**	**5.7347**	7.7678	8.0593	8.0567	7.9758	8.0204
OVERALL	SROCC	0.7987	0.8453	0.8445	0.8557	0.8865	0.8695	0.8728	0.7678	0.9104	0.9072	0.9169	0.9135	0.9160	**0.9170**	**0.9175**	**0.9180**
	KROCC	0.6179	0.6679	0.6713	0.6827	0.7140	0.7055	0.7110	0.6124	0.7378	0.7411	**0.7517**	0.7407	0.7480	0.7494	**0.7507**	**0.7519**
	PLCC	0.8171	0.8619	0.8675	0.8624	0.8933	0.8905	0.8953	0.8318	0.9040	**0.9154**	**0.9254**	0.9110	0.9098	0.9106	**0.9127**	0.9126

Table 3. Comparison of score prediction on different IQA methods on four databases.

– In each row, the double-underlined, underlined and simple bold numbers represent the first, second and third-ranked performances respectively

– For RMSE a lower score is better whereas higher values are better for the SROCC, KROCC and PLCC metrics.

Dataset	Metric	CEQI[28]	SCQI	CSCQI	SSCQI	CSSCQI
TID 2013	SROCC	0.8168	0.9052	0.9058	0.9045	0.9057
	KROCC	0.6492	0.7327	0.7337	0.7323	0.7346
	PLCC	0.8742	0.9071	0.9075	0.9082	0.9087
	RMSE	0.6019	0.5219	0.5207	0.5187	0.5175
TID 2008	SROCC	0.9069	0.9051	0.9072	0.9076	0.9088
	KROCC	0.7307	0.7294	0.732	0.7329	0.7357
	PLCC	0.9014	0.8899	0.8917	0.8933	0.8941
	RMSE	0.581	0.612	0.6073	0.603	0.601
CSIQ	SROCC	0.9563	0.9434	0.9442	0.9497	0.9473
	KROCC	0.8138	0.7863	0.7883	0.7974	0.7937
	PLCC	0.9565	0.9268	0.9278	0.9373	0.9338
	RMSE	0.0766	0.0986	0.0979	0.0915	0.0939
LIVE	SROCC	0.9577	0.9438	0.944	0.946	0.9455
	KROCC	0.8307	0.7929	0.7932	0.7968	0.7961
	PLCC	0.9534	0.9373	0.9373	0.9386	0.9379
	RMSE	6.973	8.0593	8.0567	7.9758	8.0204
OVERALL	SROCC	0.8798	0.916	0.917	0.9175	0.918
	KROCC	0.7193	0.748	0.7494	0.7507	0.7519
	PLCC	0.904	0.9098	0.9106	0.9127	0.9126
	RMSE	1.4825	1.6197	1.6174	1.6024	1.6084

Table 4. Performance comparison among CEQI, SCQI, CSCQI, SSCQI and CSSCQI

– In each row, the double-underlined, underlined and simple bold numbers represent the first, second and third-ranked performances respectively

– For RMSE a lower score is better whereas higher values are better for the SROCC, KROCC and PLCC metrics.

The performance comparison was done using four commonly adopted metrics— SROCC, KROCC, PLCC, and RMSE, elaborated as Spearman's rank-order correlation coefficient, Kendall's rank-order correlation coefficient, Pearson's linear correlation coefficient (PLCC), and the Root mean square error.

Table 3 shows the results on four benchmark databases among different IQA models for all of the four metrics mentioned above. The top three values are indicated using the double-underline, single-underline and no-underline respectively with a boldfaced font. However, in the case of RMSE, the lowest value is double-underlined, since a lower RMSE implies a better method. All of our proposed methods perform better than the SCQI, as a result, it is evident that both center-emphasis and visual saliency can improve SCQI independently. Again, we see that, for the biggest database, TID2013, CSSCQI outperforms all of the compared methods in KROCC, PLCC and RMSE metrics whereas CSCQI tops in SROCC. For the other three databases, proposed methods achieve competitive performance. We calculated the weighted averages of the SROCC, KROCC, PLCC, and RMSE using the number of distorted images to find the overall performance, as suggested in [6]. It can be noticed that, compared to VSI, HPSIe, and MDSI, our proposed CSSCQI shown better prediction accuracy with (0.76%, 1.08%, 0.11%)-point,(1.41%, 1.08%, 0.02%)-point higher overall SROCC and KROCC values, respectively.

To investigate the improvements over SCQI, Table 4 shows the comparison only among the methods proposed by us (CEQI [28], CSCQI, SSCQI and CSSCQI) along with the SCQI. CEQI was proposed in our previous workwhereas CSCQI, SSCQI and CSSCQI are discussed in this paper. We find that CSSCQI have the highest performance on the biggest two databases TID2008 and TID2013 SSCQI is the second highest. Although CEQI shows top performance for the smaller databases CSIQ and LIVE, the overall performance is the lowest. On the other hand, SSCQI achieves second highest proving the successful merge of SCQI with spectral residual saliency using the mixed-mode pooling strategy. All three proposed methods in this paper perform better than SCQI in almost all possible cases. The overall ranking based on performance is shown in Table 5. CSSCQI holds the highest ranks in SROCC and KROCC but ranked 4 in PLCC, however, the improvement of the proposed approaches over the SCQI is clearly noticed.

Table 6 shows the SROCC performance comparison for all distortion types for TID2013, CSIQ and LIVE databases; Table 2 provides the necessary description of an abbreviation. All images are not affected equally by a specific type of distortion rather it varies with the color, salient regions and several other factors related to an image. As a result, we see that methods are giving discrete performances for different distortions and performance even varies between databases. Nevertheless, Table 6 gives us a good understanding of whether an IQA method is biased to any specific noise type or not. It can be seen that the proposed CSCQI, SSCQI and CSSCQI perform consistently well for all types of distortion; they are not too biased to any specific type of distortion while retaining satisfactory average performance.

IQA	SROCC	KROCC	PLCC
SSIM	15	15	16
MS-SSIM	13	14	14
IW-SSIM	14	13	12
MAD	12	12	13
FSIMc	9	9	10
GMSD	11	11	11
MCSD	10	10	9
VIF	16	16	15
VSI	7	8	8
HPSIe	8	6	2
MDSI	4	2	1
LLM	6	7	5
SCQI	5	5	7
CSCQI	3	4	6
SSCQI	2	3	3
CSSCQI	1	1	4

Table 5. Table: Ranking of IQA methods based on overall performance

–Number1isthebestand16istheworst.

Dataset	Noise	SSIM	MS-SSIM	IW-SSIM	MAD	FSIMc	GMSD	MCSD	VIF	VSI	HPSIe	MDSI	SCQI	CSCQI	SSCQI	CSSCQI
TID 2013	AGN	0.8762	0.8773	0.8581	0.8820	0.9162	**0.9445**	0.9425	0.9302	**0.9464**	0.9405	**0.9491**	0.9438	0.9437	0.9433	0.9437
	ANC	0.7825	0.7836	0.7685	0.8053	0.8659	**0.8741**	0.8716	0.8466	0.8711	**0.8738**	**0.8842**	0.8636	0.8629	0.8634	0.8628
	SCN	0.8823	0.8844	0.8499	0.9055	0.9105	0.9401	0.9486	0.9261	0.9493	0.9429	**0.9575**	**0.9529**	0.9517	**0.9539**	0.9507
	MN	0.8373	0.8509	0.8440	0.8041	0.8760	0.8284	0.8322	**0.8908**	0.8715	0.8471	**0.8843**	0.8794	**0.8816**	0.8698	0.8725
	HFN	0.8248	0.8358	0.8530	0.8723	0.8854	0.8684	0.8754	**0.8948**	0.8783	**0.8976**	**0.8908**	0.8827	0.8840	0.8862	0.8847
	IN	0.7608	0.7397	0.7312	0.4903	0.8355	0.8369	0.8356	0.8578	**0.8846**	0.8662	**0.8851**	0.8743	**0.8774**	0.8730	0.8765
	QN	0.8748	0.8662	0.8245	0.8067	0.8701	0.8782	0.8798	**0.8961**	**0.8990**	0.8790	**0.9093**	0.8887	0.8911	0.8805	0.8883
	GB	0.9396	0.9458	0.9407	0.8878	0.9285	0.9240	0.9198	0.8327	0.9267	0.8969	0.9365	**0.9624**	**0.9624**	**0.9639**	0.9618
	IDN	0.8920	0.8906	0.8947	0.9212	0.9247	**0.9318**	0.9239	0.8684	0.9207	**0.9440**	0.9126	0.9247	0.9266	0.9237	**0.9303**
	JPEG	0.8917	0.9000	0.8898	0.9126	0.8918	0.9042	0.9033	0.8700	**0.9188**	**0.9227**	0.9125	0.9173	0.9171	**0.9180**	0.9177
	JP2K	0.9461	0.9541	0.9499	0.9526	0.9630	**0.9720**	0.9674	0.9443	**0.9707**	0.9636	0.9651	0.9648	0.9664	0.9670	**0.9680**
	JGTE	0.9137	0.9216	0.9123	0.9272	0.9239	0.8950	0.8744	0.8785	**0.9555**	0.9378	0.9405	0.9411	**0.9459**	0.9418	**0.9446**
	J2TE	0.8789	0.8838	0.8542	0.8291	0.8944	0.8751	0.8560	0.8840	0.9252	0.9095	0.9187	**0.9299**	**0.9346**	0.9291	**0.9327**
	NEPN	0.8667	0.8503	0.7587	0.6555	0.8587	0.8085	0.8012	0.8333	**0.8807**	0.8502	0.8388	**0.8759**	**0.8760**	0.8711	0.8663
	LBD	0.4099	0.7713	0.6889	0.8618	0.7576	0.8601	**0.8911**	0.6727	0.8530	0.8333	**0.8884**	0.8730	0.8665	**0.8744**	0.8711
	MS	0.3118	0.8260	0.8259	0.7745	0.8800	**0.8967**	**0.8978**	0.7556	0.8308	0.8658	**0.9049**	0.8709	0.8622	0.8594	0.8649
	CTC	**0.6363**	0.5742	0.5702	0.4839	0.5561	0.5738	0.5067	**0.7290**	0.5856	0.5636	**0.5990**	0.5887	0.5868	0.5897	0.5969
	CCS	0.0725	0.6860	**0.6899**	0.4094	0.3905	**0.7344**	**0.7159**	0.5446	0.5546	0.5390	0.5504	0.5837	0.5864	0.5722	0.5800
	MGN	0.4613	0.4596	0.4453	0.3744	0.5707	0.4145	0.4117	0.4501	0.7947	0.5896	0.7330	**0.8204**	0.8189	**0.8193**	**0.8192**
	CMN	0.7915	0.7796	0.7841	0.8470	0.8736	0.9055	0.9047	0.7869	**0.9235**	0.8955	0.8998	0.9179	**0.9190**	0.9156	**0.9188**
	LCN	0.8901	0.8895	0.9043	0.9335	0.9338	**0.9487**	**0.9512**	0.9145	0.9400	**0.9435**	0.9378	0.9309	0.9320	0.9343	0.9332
	ICQ	0.9023	0.9041	0.9007	0.9195	0.9466	**0.9621**	**0.9592**	0.9032	0.9386	**0.9591**	0.9528	0.9470	0.9498	0.9461	0.9473
	CA	0.8620	0.8678	0.8513	0.8861	0.8756	0.9059	0.9170	0.8439	0.8951	0.8873	0.8953	**0.9258**	**0.9257**	0.9250	**0.9261**
	SS	0.8805	0.8820	0.8717	0.8288	0.8870	0.8483	0.8350	0.8852	0.8845	0.8475	0.8719	**0.8940**	**0.8943**	0.8933	**0.8950**
	AVG	18.3856	19.8242	19.4618	18.9711	20.2161	20.5312	20.4220	19.8393	20.9989	20.5960	21.0183	**21.1538**	**21.1630**	21.1140	**21.1531**
CSIQ	AGN	0.7212	0.7310	0.7187	0.7611	0.8360	0.7847	0.7906	0.6559	**0.8980**	0.8611	0.8788	0.8841	**0.8881**	0.8836	**0.8890**
	JPEG	0.8341	0.8759	0.8990	0.9103	0.9158	0.9075	0.8973	0.7628	0.9132	**0.9321**	0.9198	0.9167	0.9182	**0.9199**	**0.9209**
	JP2K	0.7778	0.8556	0.8794	0.8483	0.8984	0.8706	0.8588	0.7427	**0.9236**	**0.9325**	**0.9434**	0.9212	0.9216	0.9213	0.9231
	AGPN	0.6871	0.7300	0.7095	0.7462	0.8056	0.7484	0.7582	0.7081	0.8701	0.8564	0.8391	**0.8864**	**0.8888**	0.8818	**0.8848**
	GB	0.8071	0.8133	0.7950	0.7776	0.8770	0.8103	0.7999	0.7205	**0.9051**	0.8839	0.8759	**0.8999**	**0.8998**	0.8937	0.8976
	CTC	0.7907	0.8268	0.8072	0.8017	0.8661	0.8249	0.8051	0.7186	**0.9196**	0.9027	0.9107	0.9132	0.9121	**0.9157**	**0.9155**
	AVG	0.7697	0.8054	0.8015	0.8075	0.8665	0.8244	0.8183	0.7181	**0.9049**	0.8948	0.8946	0.9036	**0.9048**	0.9027	**0.9052**
LIVE	JP2K	0.7267	0.7569	0.7426	0.7806	0.8444	0.8073	0.8056	0.6867	**0.8973**	0.8725	0.8820	0.8886	**0.8928**	0.8900	**0.8942**
	JPEG	0.8185	0.8727	0.9027	0.8823	0.9166	0.9024	0.8925	0.7855	0.9257	**0.9428**	**0.9453**	0.9245	0.9256	0.9258	**0.9285**
	AWGN	0.7413	0.7772	0.7603	0.7758	0.8373	0.7762	0.7858	0.7333	0.8832	0.8726	0.8589	**0.8897**	**0.8909**	0.8859	**0.8874**
	GB	0.7827	0.8116	0.7962	0.8093	0.8559	0.8078	0.7872	0.6753	**0.9004**	0.8843	0.8795	**0.9007**	**0.9002**	0.8981	0.9000
	FF	0.7259	0.7572	0.7540	0.8213	0.8623	0.7525	0.7698	0.7076	0.9045	0.8649	0.8836	**0.9178**	**0.9160**	**0.9172**	0.9157
	AVG	0.7590	0.7951	0.7912	0.8139	0.8633	0.8092	0.8082	0.7177	0.9022	0.8874	0.8899	**0.9043**	**0.9051**	0.9034	**0.9052**

Table 6. Distortion-wise comparison of SROCC performed on three databases

– In each row, the double-underlined, underlined and simple bold numbers represent the first, second and third-ranked performances respectively

–The distortion acronyms are defined in Table 2, AVG refers to the aggregated average over all noises in a database

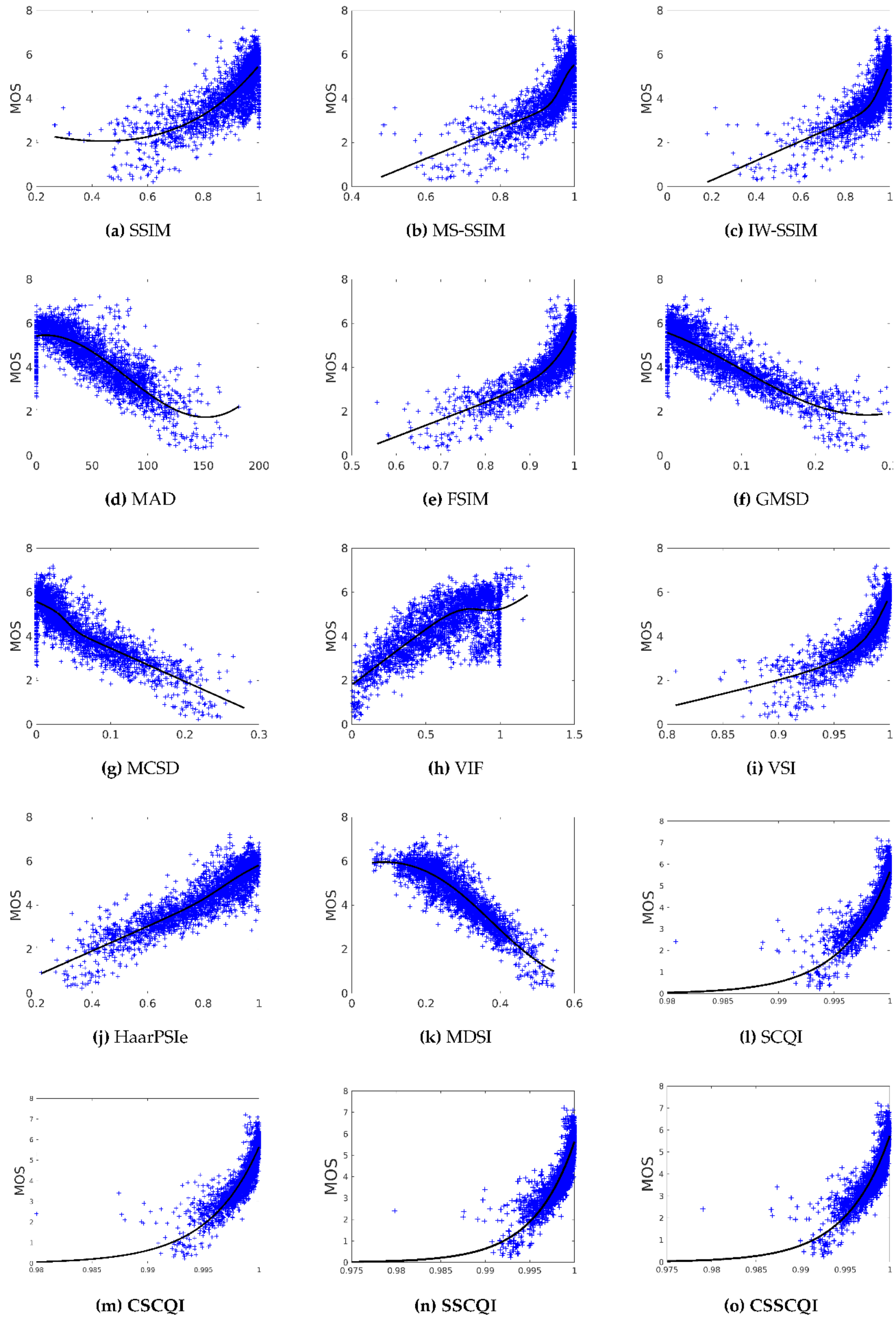


Figure 3. Predicted scores with the MOS on TID2013 database. The black curves are obtained by a nonlinear fitting.

Scatter plots in Figure 3 demonstrates the predicted scores for different IQA approaches with the subjective scores for the TID2013 database. We see that our proposed approaches are quite consistent in predicting scores as compared to other methods and providing better correlation with MOS/DMOS.

IQA	Average Running Time (ms)	Average Images Per Second
SSIM	0.04	22.64
MS-SSIM	0.15	6.51
IW-SSIM	0.94	1.07
MAD	3.94	0.25
FSIMc	0.65	1.55
GMSD	0.02	45.09
MCSD	0.04	26.56
VIF	3.23	0.31
VSI	0.40	2.51
HPSIe	0.22	4.60
MDSI	0.06	17.15
SCQI	0.26	3.83
CSCQI	0.26	3.82
SSCQI	0.32	3.16
CSSCQI	0.35	2.88

Table 7. Running time comparison on IQA models

The principal purpose of designing an IQA model is the performance of its prediction. However, in some cases processing time is also a major concern, especially in a real-time system. The run-time comparison was performed on various IQA models with MATLAB R2018b on a computer having Intel(R) Core(TM) i5-4670 CPU with a

3.40GHz processor and 16GB of RAM. The MATLAB codes written by the authors of each method was collected from their websites. We run the codes and elapsed time was recorded; Table 7 shows the results. We see that the improvements from SCQI take some time cost. SCQI and CSCQI have almost the same running time of 0.26 milliseconds with 3.83 and 3.82 images per second respectively. However, SSCQI and CSSCQI are able to process 3.16 and 2.88 images per second respectively. Notwithstanding, CSSCQI can process more images than IW-SSIM, MAD, FSIMc, VIF and VSI among the compared methods.

4. Conclusion

In this paper, we considered the center bias of HVS and proposed several approaches for full-reference image quality assessment method (CSCQI, SSCQI and CSSCQI). The merging of SCQI with spectral residual visual saliency was successful, thanks to the novel mixed-mode pooling strategy. Also, giving extra emphasis on the center part of the image was also improved the performance of quality assessment. The proposed approaches were compared with other state-of-the-art IQA models and they outperform most of the competing methods. Comparing individual distortion types, proposed methods give consistent scores. Incorporating additional features took some extra time, still, the proposed approaches stay above compared state-of-the-art approaches. In our study we have already found that this center emphasized approach enhances the performance of few other existing IQA models and we believe that same will happen with most of the other no-reference and reduced-reference models. In our future work, we will further investigate these possibilities.

Acknowledgements

This work was supported by the Information Technology Research and Resource Center (ITRRC, web: http://itrrc.com), and the JnU research grant (জবি/গবেষণা/গপ্র/২০২০-২০২১/বিজ্ঞান/৩৩) Jagannath University, Dhaka, Bangladesh.

References

Bae, S.H.; Kim, M. A novel generalized DCT-based JND profile based on an elaborate CM-JND model for variable block-sized transforms in monochrome images. IEEE Transactions on Image processing 2014, 23, 3227–3240.

Bae, S.H.; Kim, M. A novel image quality assessment with globally and locally consilient visual quality perception. IEEE Transactions on Image Processing 2016, 25, 2392–2406.

Bindemann, M. Scene and screen center bias early eye movements in scene viewing. Vision research 2010, 50, 2577–2587.

Chandler, D.M.; Hemami, S.S. VSNR: A wavelet-based visual signal-to-noise ratio for natural images. IEEE transactions on image processing 2007, 16, 2284–2298.

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

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

Ding, Y. General Framework of Image Quality Assessment. In Visual Quality Assessment for Natural and Medical Image; Springer, 2018; pp. 45–62.

Duan, L.; Wu, C.; Miao, J.; Qing, L.; Fu, Y. Visual saliency detection by spatially weighted dissimilarity. Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011, pp. 473–480.

Hou, X.; Zhang, L. Saliency detection: A spectral residual approach. Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. IEEE, 2007, pp. 1–8.

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

Langford, R.C. How People Look at Pictures, A Study of the Psychology of Perception in Art. 1936.

Larson, E.C.; Chandler, D. Categorical image quality (CSIQ) database, 2010.

Layek, M.; Uddin, A.; Le, T.P.; Chung, T.; Huh, E.N.; others. Center-Emphasized Visual Saliency and a Contrast-Based Full Reference Image Quality Index. Symmetry 2019, 11, 296.

Li, A.; She, X.; Sun, Q. Color image quality assessment combining saliency and fsim. Fifth International Conference on Digital Image Processing (ICDIP 2013). International Society for Optics and Photonics, 2013, Vol. 8878, p. 88780I.

Ma, Q.; Zhang, L. Saliency-based image quality assessment criterion. International Conference on Intelligent Computing. Springer, 2008, pp. 1124–1133.

Mannan, S.; Ruddock, K.; Wooding, D. Fixation sequences made during visual examination of briefly presented 2D images. Spatial vision 1997, 11, 157–178.

Nafchi, H.Z.; Shahkolaei, A.; Hedjam, R.; Cheriet, M. Mean deviation similarity index: Efficient and reliable full-reference image quality evaluator. IEEE Access 2016, 4, 5579–5590.

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

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

Ponomarenko, N.; Ieremeiev, O.; Lukin, V.; Jin, L.; Egiazarian, K.; Astola, J.; Vozel, B.; Chehdi, K.; Carli, M.; Battisti, F.; others. A new color image database TID2013: Innovations and results. International Conference on Advanced Concepts for Intelligent Vision Systems. Springer, 2013, pp. 402–413.

Ponomarenko, N.; Lukin, V.; Zelensky, A.; Egiazarian, K.; Carli, M.; Battisti, F.TID2008-a database for evaluation of full-reference visual quality assessment metrics. Advances of Modern Radio electronics 2009, 10, 30–45.

Reisenhofer, R.; Bosse, S.; Kutyniok, G.; Wiegand, T. A Haar wavelet-based perceptual similarity index for image quality assessment. Signal Processing: Image Communication 2018, 61, 33–43.

Sheikh, H. LIVE image quality assessment database release 2. http://live. ece. utexas. edu/research/quality 2005.

Sheikh, H.R.; Bovik, A.C. A visual information fidelity approach to video quality assessment. The First International Workshop on Video Processing and Quality Metrics for Consumer Electronics, 2005, pp. 23–25.

Sheikh, H.R.; Bovik, A.C. Image information and visual quality. IEEE Transactions on Image Processing 2006, 15, 430–444. doi:10.1109/TIP.2005.859378.

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

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.

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

Wang, H.; Fu, J.; Lin, W.; Hu, S.; Kuo, C.C.J.; Zuo, L. Image quality assessment based on local linear information and distortion-specific compensation. IEEE Transactions on image processing 2017, 26, 915–926.

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 Processing: Image Communication 2016, 45, 1–9.

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

Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing 2004, 13, 600–612.

Wang, Z.; Li, Q. Information content weighting for perceptual image quality assessment. IEEE Transactions on Image Processing 2011, 20, 1185–1198.

Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale structural similarity for image quality assessment. The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003. Ieee, 2003, Vol. 2, pp. 1398–1402.

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

Zhang, L.; Li, H. SR-SIM: A fast and high performance IQA index based on spectral residual. Image Processing (ICIP), 2012 19th IEEE International Conference on. IEEE, 2012, pp. 1473–1476.

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

Zhang, L.; Zhang, L.; Mou, X.; Zhang, D.; others. FSIM: a feature similarity index for image quality assessment. IEEE transactions on Image Processing 2011, 20, 2378–2386.