Noname manuscript No.
(will be inserted by the editor)

Content-Weighted Mean-Squared Error for Quality Assessment of Compressed Images

Ke Gu, Shiqi Wang, Guangtao Zhai, Siwei Ma, Xiaokang Yang, Wenjun Zhang

Received: date / Accepted: date

Abstract Image quality assessment (IQA) has been intensively studied, especially for the full-reference (FR) scenario. However, only the mean squared error (MSE) is widely employed in compression. Why other IQA metrics work ineffectively? We first sum up three main limitations including the computational time, portability, and working manner. To address these problems, we then in this paper propose a new Content-Weighted MSE (CW-MSE) method to assess the quality of compressed images. The design principle of our model is to use adaptive Gaussian convolution to estimate the influence of image content in a block-based manner, thereby to approximate the human visual perception to image quality. Results of experiments on six popular subjective image quality databases (including LIVE, TID2008, CSIQ, IVC, Toyama and TID2013) confirm the superiority of our CW-MSE over state-of-the-art FR IQA approaches.

Ke Gu*, Guangtao Zhai, Xiaokang Yang, and Wenjun Zhang
the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China, 200240
Shanghai Key Laboratory of Digital Media Processing and Transmissions
*Tel.: 86-13764402720
*Fax: 86-21-34204511
*E-mail: guke.doctor@gmail.com

Shiqi Wang
the Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada

Siwei Ma
the Institute of Digital Media, Peking University, Beijing, China, 100871

Keywords Image quality assessment (IQA) · mean-squared error (MSE) · image compression · image content

1 Introduction

A large quantity of image quality assessment (IQA) metrics, in terms of various schemes such as subjective testing, perceptual model, saliency detection, machine learning and brain science, have been proposed during the last decade. IQA tasks are mainly divided into subjective assessment and objective assessment. The first directly measures the perception of the human visual system (HVS) to image quality and records the ratings as the overall quality scores, i.e. the mean opinion score (MOS). Subjective quality scores are quite valuable, because on one hand they can be used to justify the performance of objective IQA algorithms (e.g. LIVE [1], TID2008 [2], CSIQ [3], IVC [4], Toyama [5] and TID2013 [6]), and on the other hand they are able to instruct denoising [7], restoration [8], coding [9]-[11], and compare tone-mapped operators [12]-[14]. But subjective methods easily suffer the drawbacks of being laborious, costly and time-consuming, and thus cannot be adapted to real-time applications. As a result, recent years have witnessed a surge of objective IQA models in the literature.

Despite the emergence of hundreds of objective metrics, most of them lie in full-reference (FR) IQA approaches. The famous one is the traditional mean-squared error (MSE), which has prevailed for decades owing to its simplicity and clear physical meaning. Nonetheless, MSE was found to poorly correlate with human judgements of image quality [15]. The epoch-making structural similarity index (SSIM) [16] was developed in light

of luminance, contrast and structural similarities between the original and distorted image signals, and its variants [17]-[22] have been designed to pursue higher accuracy since then. Other advanced FR IQA methods [23]-[29] were explored via the natural scene statistics (NSS) [30], low-level vision [31], and free energy principle [32].

Researchers have reached a common agreement that IQA can assist the development and optimization of a wide variety of image processing systems, such as compression, communication, denoising, enhancement, restoration and display. Only MSE and SSIM, however, have been adopted, especially in the field of compression, whereas an enormous number of recently designed FR IQA methods are cast aside. So we first summarize several crucial reasons of the above-mentioned phenomenon, and then develop a new effective Content-Weighted MSE (CW-MSE) method.

In general, a low-pass filter can well approximate the contrast sensitivity function (CSF) of the HVS, which perceives a small patch of an image by taking into account the content correlation of itself and the neighbors. A similar basic idea has been recently used in entropy of primitives (EOP) [33] as well, but this method works too slow and not effectively. It is natural that the low-pass filter cannot be globally unified but works in a block-based way owing to the complexity of the HVS. So our CW-MSE first estimates the adaptive Gaussian kernel based on the correlation computed between each block and its surrounding blocks using the local MSE, and then adopts the adaptive Gaussian kernel to convolve each block before computing MSE of the original and distorted images, in order to derive the overall quality score.

We arrange the rest of this paper as follows. Section 2 first analyzes "why most FR IQA methods work invalidly in compression?", and then provides the proposed CW-MSE in detail. A comparison of our CW-MSE with the FR IQA algorithms stated above on six popular publicly-available LIVE [1], TID2008 [2], CSIQ [3], IVC [3], Toyama [5] and TID2013 [6] databases will be reported and discussed in Section 3. We finally conclude this paper in Section 4.

2 Proposed Quality Metric

Despite quite a lot of FR IQA methods, most image/video processing systems merely include MSE and SSIM, particularly in the important coding. Moreover, it was found that, despite several uses of SSIM [9]-[11], in most cases MSE is employed for optimization in coding. "Why the state-of-the-art FR IQA models cannot be applied to compression?" The following three points may answer this question.

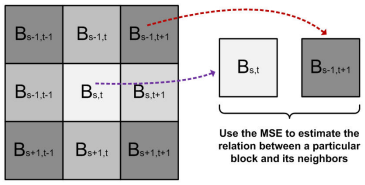


Fig. 1 Illustration of the proposed content-weighted (CW) strategy for a particular block $B_{s,t}$ of size $M \times M$. The correlation between the current block $B_{s,t}$ and its surrounding eight blocks are computed with the local MSE.

- First, real-time systems are highly in demand in coding. But most FR metrics operate with massive time, far less than the ideal technologies.
- Second, most existing systems are not well compatible with complicated models, which those above FR IQA methods are based on. That is those FR approaches have weak portability, so they are hard to be embedded into most existing applications.
- Third, coding works in the block-based manner. But above FR metrics require the entire image, letting them not available in block-matching.

Note that neither MSE or SSIM has desirable performance in assessing JPEG/ JPEG2000 (JP2K) compressed images, as given in Section 3. Also, content-based methods have been validly incorporated into compression [34]. So in this article we design the effective, efficient and portable CW-MSE metric.

The most serious disadvantage of MSE should be that it fails to account for the impact of image content. The recently developed information content weighting (IW) strategy aims to solve this problem and successfully improves MSE to derive higher-performance IW-MSE. The IW scheme, however, is motivated by the complicated MS and NSS models, and thus has substantially computational cost. Also, IW needs the whole compressed image, making it must work many times when block-matching in image coding.

In this paper, the Content-Weighted (CW) model is proposed to characterize the image content in an alternative way, and is realized in two steps. Specifically, when perceiving each small patch, the HVS jointly takes into consideration the content of each patch itself and the correlation with its surrounding patches. Although the CSF function has provided a curve of the HVS's response to varying spatial frequency, it only works for simple patterns, ignoring the effect of image content such as texture masking. Furthermore, because of the extreme complexity of the HVS, we believe the approximate low-pass filter is not unified glob-

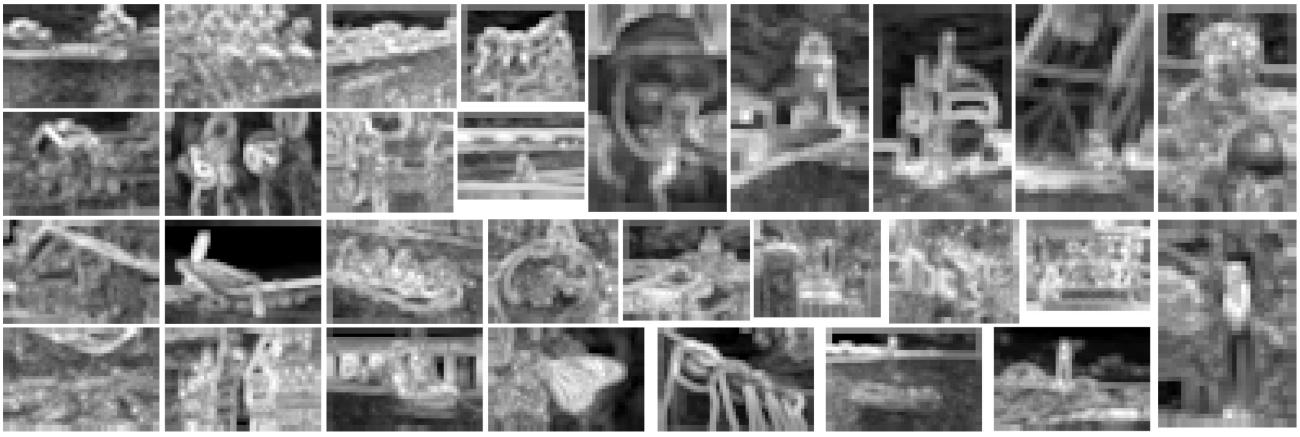


Fig. 2 The CW weighting maps of source images in LIVE [1].

ally but block-based[1]. In reality, the CSF curve suggests that the HVS will make a deeper depression on high-frequency information than that on low-frequency one. In other words, the more unlikely the current block and its neighbors are, the deeper depression we will make on. Therefore, the first step of CW-MSE is to design the adaptive Gaussian convolution (a common low-pass filter) for each block based on the content correlation of itself and its connections.

To specify, for a particular block $B_{s,t}$ of size $M \times M$ in an image I, the correlation coefficient $\phi(s,t)$ is computed by

$$\phi(s,t) = \sum_{k=1}^{K} \lambda(k) \cdot b(k) \tag{1}$$

where M is set as 8 due to the fact that the processed blocks are usually of size 8×8, K is also set as 8 since we only consider its 8-connected blocks of the current one, and $\lambda(k)$ are predefined eight constants. The similarity degree $b(k)$ of $B_{s,t}$ and its eight neighboring blocks is defined by

$$(\mathrm{I}) \begin{cases} b(k=1) = \mathrm{MSE}(\tilde{B}_{s,t}, \tilde{B}_{s-1,t}) \\ b(k=2) = \mathrm{MSE}(\tilde{B}_{s,t}, \tilde{B}_{s,t+1}) \\ b(k=3) = \mathrm{MSE}(\tilde{B}_{s,t}, \tilde{B}_{s+1,t}) \\ b(k=4) = \mathrm{MSE}(\tilde{B}_{s,t}, \tilde{B}_{s,t-1}) \end{cases} \tag{2}$$

$$(\mathrm{II}) \begin{cases} b(k=5) = \mathrm{MSE}(\tilde{B}_{s,t}, \tilde{B}_{s-1,t+1}) \\ b(k=6) = \mathrm{MSE}(\tilde{B}_{s,t}, \tilde{B}_{s+1,t+1}) \\ b(k=7) = \mathrm{MSE}(\tilde{B}_{s,t}, \tilde{B}_{s+1,t-1}) \\ b(k=8) = \mathrm{MSE}(\tilde{B}_{s,t}, \tilde{B}_{s-1,t-1}) \end{cases} \tag{3}$$

where $k = \{1,2,3,4\}$ in (I) indicate 4-connected top, right, bottom and left blocks (colored light gray in Fig. 1), whereas $k = \{5,6,7,8\}$ in (II) represent 8-connected top-right, bottom-right, bottom-left and top-left blocks (colored dark gray in Fig. 1) except the 4-connected blocks in (I). The MSE of two arbitrary blocks X and Y is defined as

$$\mathrm{MSE}(X,Y) = \frac{\sum_{p=1}^{P} \sum_{q=1}^{Q} [X(p,q) - Y(p,q)]^2}{P \cdot Q} \tag{4}$$

where X and Y are of the same size $P \times Q$. The current block $\tilde{B}_{s,t}$ and its connected eight blocks are extracted from $\tilde{I}$ that is processed by

$$\tilde{I}(i,j) = \frac{\sum_{p \in \Omega_i} \sum_{q \in \Omega_j} U(p,q) \cdot I(p,q)}{\sum_{p \in \Omega_i} \sum_{q \in \Omega_j} U(p,q)} \tag{5}$$

with

$$U(p,q) = \exp\left(\frac{-\sqrt{(i-p)^2 + (j-q)^2}}{2\sigma_w^2}\right) \tag{6}$$

where we assign σ_w as 1.5, $\Omega_i = \{i-3, i-2, i-1, i, i+1, i+2, i+3\}$ and $\Omega_j = \{j-3, j-2, j-1, j, j+1, j+2, j+3\}$.

Next, the content-adaptive Gaussian kernel, which is stable for the overall 64 pixels in each block yet distinct among various blocks, is defined for the current block $B(s,t)$ as follows:

$$V(s,t;m,n) = \exp\left(\frac{-\sqrt{m^2+n^2}}{2\sigma_v^2(s,t)}\right) \tag{7}$$

where both m and n are limited into the range of $\{-3,-2,-1,0,1,2,3\}$ in this implementation, and we adopt a three-parameters power-law function to estimate the variable $\sigma_v(s,t)$:

$$\sigma_v(s,t) = \alpha \cdot \phi(s,t)^{\gamma} + \beta \tag{8}$$

where α, β and γ are fixed parameters that control the key variable σ_v in the CW-MSE. $\phi(s,t)$ is computed by

[1] This statement will be explicitly analyzed in Section 3.

Eq. (1). To illustrate the effectiveness, we first normalize the parameter $\sigma_v(s,t)$ to generate the CW maps of 29 source images in the LIVE database [1]:

$$\mathrm{CW}(s,t) = \sigma_v(s,t)/\max(\sigma_v) \tag{9}$$

and then show the CW maps in Fig. 2 for straightforward understanding. It is obvious that the CW maps succeed in catching those blocks with important information (e.g. located near to edges) by just using the original image and MSE.

The second step is to apply the kernel $V(s,t;m,n)$ to convolve the current block $B(s,t)$:

$$B'(s,t) = \frac{\sum_{(m+s)\in\Omega_s}\sum_{(n+t)\in\Omega_t} V(s,t;m,n)\cdot B(m,n)}{\sum_{(m+s)\in\Omega_s}\sum_{(n+t)\in\Omega_t} V(s,t;m,n)}. \tag{10}$$

An important note is that for fast implementation, the kernel $V(s,t;m,n)$ is computed only once from the block $B(s,t)$ in the original image I, and the corresponding block $B_d(s,t)$ in the distorted image I_d is processed using the same kernel:

$$B'_d(s,t) = \frac{\sum_{(m+s)\in\Omega_s}\sum_{(n+t)\in\Omega_t} V(s,t;m,n)\cdot B_d(m,n)}{\sum_{(m+s)\in\Omega_s}\sum_{(n+t)\in\Omega_t} V(s,t;m,n)}. \tag{11}$$

We then evaluate the local MSE value Q_{MSE} between the block $B'(s,t)$ and $B'_d(s,t)$:

$$Q_{\mathrm{MSE}}(s,t) = \mathrm{MSE}(B'(s,t), B'_d(s,t)) \tag{12}$$

followed by pooling those MSE values of whole blocks to derive the overall quality score of the CW-MSE:

$$\text{CW-MSE} = \sum_{s=1}^{H/M}\sum_{t=1}^{W/M} Q_{\mathrm{MSE}}(s,t) \tag{13}$$

where H and W are the height and width of the image I.

3 Experimental Results

In this section, we will testify and compare the performance of the CW-MSE with classical and state-of-the-art FR IQA metrics: 1) Five classical FR MSE, SSIM [16], MS-SSIM [17], IW-MSE [18], and IW-SSIM [18]; 2) Five state-of-the-art FR FSIM [24], Q [25], GSIM [26], IGM [27], and PAMSE [29][2].

Six image databases are adopted here. 1) The LIVE database [1], which contains five image data sets with five commonly encountered distortion types, and totally 779 corrupted images from 29 references. In this research, we select 344 images belonging to JPEG and JP2K data sets. 2) The TID2008 database [2], which is a very large database including 1700 images generated from 25 pristine ones with 17 distortion categories at four distortion levels. Here, we choose 200 images corrupted by two distortion types, JPEG and JP2K. 3) The CSIQ database [3], consisting of 866 images by using six types of distortions at four to five distortion levels to corrupt 30 original versions. In our study, 300 images corresponding to JPEG and JP2K compressions are used for testing. 4) The IVC database [4], which is consisted by 185 images generated from 10 sources. Two distortion types and their associated 100 JPEG and JP2K compressed images are applied. 5) The Toyama database [5], which involves 168 distorted images. The whole database is adopted in this work. 6) The TID2013 database [6], extending the four distortion levels in TID2008 to five, making the nowadays largest database with a total number of 3000 images. Here we pick 250 JPEG and JP2K compressed images.

As suggested by the video quality experts group (VQEG) [37], we first map the objective predictions of testing IQA metrics to subjective scores using the nonlinear regression with a five-parameter logistic function:

$$Q(q) = c_1\left(\frac{1}{2} - \frac{1}{1+\exp\left(c_2(q-c_3)\right)}\right) + c_4 q + c_5 \tag{14}$$

where q and $Q(q)$ are the input and mapped scores, and $\{c_1, c_2, c_3, c_4, c_5\}$ are required to be determined during the curve fitting process. We then apply two frequently used performance measures to evaluate and compare the proposed CW-MSE with the IQA metrics tested in this study. First, Spearman rank-order correlation coefficient (SRC), which does not consider the relative distance between the data but computes their monotonicity. SRC is defined by

$$\mathrm{SRC} = 1 - \frac{6\sum_{i=1}^{N}\delta_i^2}{N(N^2-1)} \tag{15}$$

where δ_i is the difference of the i-th image's ranks in subjective and objective quality evaluations, and N stands for the number of images in the testing database. SRC is a non-parametric rank-based correlation measure, independent of any monotonic nonlinear mapping between subjective and objective quality ratings. Second, Pearson linear correlation coefficient (PCC), meaning the linear prediction performance, is computed between subjective ratings and objective scores after the nonlinear regression by Eq. (14). PCC can be calculated

[2] Interested readers can be directed to [35]-[36] for more recent IQA studies.

Table 1 Performance evaluations on JPEG compressed images. We bold two best models.

SRC	LIVE	TID2008	CSIQ	Toyama	IVC	TID2013	Average
MSE	0.8809	0.8717	0.8881	0.2868	0.7400	0.9189	0.7644
SSIM	0.9449	0.8989	0.9222	0.6263	0.8269	0.9200	0.8565
MS-SSIM	0.9794	0.9341	0.9626	0.8360	0.9261	0.9294	0.9279
IW-MSE	0.9670	**0.9529**	0.9655	0.7835	0.9354	**0.9536**	0.9263
IW-SSIM	0.9808	0.9207	0.9662	0.9204	0.9493	0.9198	0.9429
FSIM	**0.9834**	0.9370	0.9654	0.8991	**0.9582**	0.9382	0.9469
Q [25]	0.9630	0.8902	0.9319	0.8253	0.8530	0.8871	0.8918
GSIM	0.9778	0.9393	0.9632	**0.9328**	**0.9586**	0.9284	**0.9500**
IGM	**0.9808**	0.9472	**0.9686**	0.8260	0.9304	0.9487	0.9336
PAMSE	0.9567	**0.9600**	0.9526	0.6725	0.9023	**0.9575**	0.9002
CW-MSE	0.9758	0.9493	**0.9672**	**0.9384**	0.9523	0.9508	**0.9556**

PCC	LIVE	TID2008	CSIQ	Toyama	IVC	TID2013	Average
MSE	0.8891	0.8702	0.8910	0.3268	0.7464	0.9163	0.7733
SSIM	0.9510	0.9254	0.9420	0.6543	0.8368	0.9544	0.8773
MS-SSIM	0.9820	0.9618	0.9824	0.8415	0.9392	0.9613	0.9447
IW-MSE	0.9763	**0.9825**	0.9813	0.7903	0.8272	**0.9811**	0.9231
IW-SSIM	0.9823	0.9595	0.9844	0.9168	0.9568	0.9572	0.9595
FSIM	0.9838	0.9743	0.9835	0.8922	**0.9624**	0.9713	**0.9613**
Q [25]	0.9619	0.9220	0.9584	0.8252	0.8357	0.9229	0.9044
GSIM	0.9840	0.9546	0.9833	**0.9274**	**0.9605**	0.9470	0.9595
IGM	**0.9859**	**0.9795**	**0.9850**	0.8280	0.9439	0.9773	0.9483
PAMSE	0.9605	0.9768	0.9679	0.6907	0.9152	**0.9809**	0.9153
CW-MSE	**0.9852**	0.9768	**0.9857**	**0.9384**	0.9567	0.9744	**0.9695**

Table 2 Performance evaluations on JP2K compressed images. We bold two best models.

SRC	LIVE	TID2008	CSIQ	Toyama	IVC	TID2013	Average
MSE	0.8954	0.8123	0.9362	0.8605	0.8500	0.8840	0.8731
SSIM	0.9355	0.8875	0.9207	0.9148	0.8488	0.9468	0.9090
MS-SSIM	0.9654	0.9738	0.9707	0.9470	0.9319	0.9483	0.9562
IW-MSE	0.9617	0.9745	0.9777	0.9240	0.9378	0.9645	0.9567
IW-SSIM	0.9649	0.9738	0.9683	**0.9549**	**0.9495**	0.9506	0.9604
FSIM	**0.9717**	0.9773	0.9685	**0.9566**	0.9402	0.9577	**0.9620**
Q [25]	0.9353	0.9497	0.9314	0.8508	0.8823	0.9233	0.9121
GSIM	**0.9700**	0.9760	0.9648	0.9506	0.9360	0.9602	0.9596
IGM	0.9679	**0.9844**	**0.9784**	0.9306	0.9353	0.9684	0.9608
PAMSE	0.9397	0.9878	0.9702	0.9026	0.9266	**0.9759**	0.9490
CW-MSE	0.9658	**0.9794**	**0.9785**	0.9484	**0.9449**	**0.9700**	**0.9645**

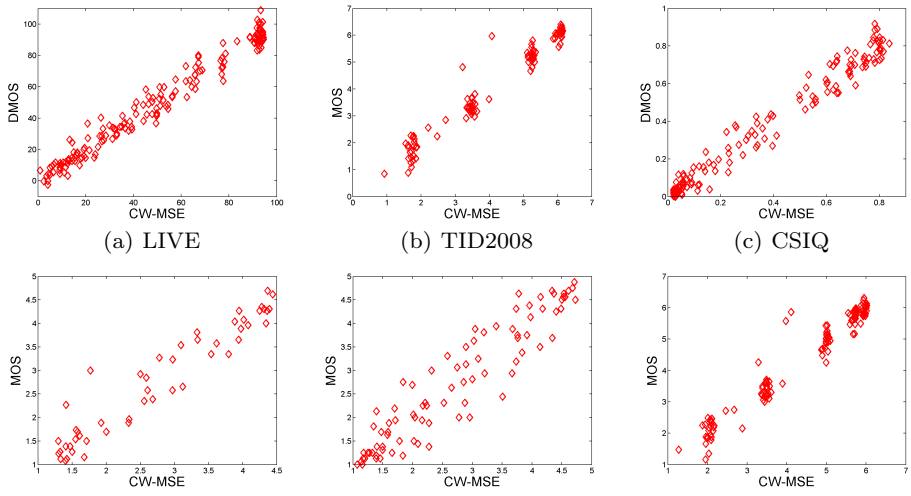
PCC	LIVE	TID2008	CSIQ	Toyama	IVC	TID2013	Average
MSE	0.8997	0.8675	0.9461	0.8577	0.8473	0.8936	0.8853
SSIM	0.9415	0.8841	0.9236	0.9211	0.8630	0.9651	0.9164
MS-SSIM	0.9711	0.9756	0.9802	0.9506	0.9341	0.9685	0.9634
IW-MSE	0.9647	0.9787	0.9835	0.9260	0.9388	0.8551	0.9411
IW-SSIM	0.9732	0.9761	0.9809	**0.9592**	**0.9540**	0.9671	**0.9684**
FSIM	**0.9775**	0.9801	0.9807	0.8871	0.9416	0.9732	0.9567
Q [25]	0.9309	0.9522	0.9492	0.8562	0.8745	0.9439	0.9178
GSIM	**0.9757**	0.9826	0.9730	**0.9569**	0.8255	0.9789	0.9488
IGM	0.9734	**0.9871**	**0.9845**	0.9360	0.9367	0.9816	0.9665
PAMSE	0.9446	0.9836	0.9775	0.9019	0.9211	**0.9865**	0.9525
CW-MSE	0.9715	**0.9862**	**0.9868**	0.9550	**0.9458**	**0.9852**	**0.9718**

by

$$\mathrm{PCC} = \frac{\sum_{i=1}^{N}(a_i - \bar{a}) \cdot (b_i - \bar{b})}{\sqrt{\sum_{i=1}^{N}(a_i - \bar{a})^2 \cdot \sum_{i=1}^{N}(b_i - \bar{b})^2}} \tag{16}$$

where a_i is the subjective ratings of the i-th image, and $\bar{a}$ is the average of the overall a_i. b_i is the converted objective quality score of the i-th image after the non-linear regression, and $\bar{b}$ is the average of the overall b_i.

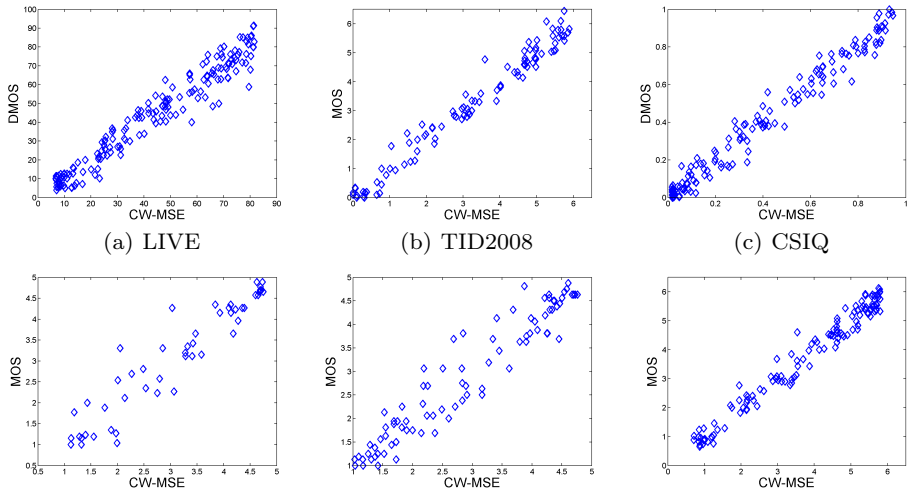
Among the above two performance measures, a good image quality metric is of a value close to 1 for SRC and PCC. We separately list and compare the performance results for JPEG and JP2K compressed images in Tables 1-2. It is apparent that the proposed



(a) LIVE (b) TID2008 (c) CSIQ

(d) IVC (e) Toyama (f) TID2013

Fig. 3 Scatter plots of DMOS/MOS versus CW-MSE on JPEG compressed images.



(a) LIVE (b) TID2008 (c) CSIQ

(d) IVC (e) Toyama (f) TID2013

Fig. 4 Scatter plots of DMOS/MOS versus CW-MSE on JP2K compressed images.

CW-MSE performs substantially well, even superior to classical and state-of-the-art approaches on average. We also draw the scatter plots of MOS/DMOS versus the CW-MSE on six image quality databases in Figs. 3-4, which further illustrates the impressive linearity and monotonicity of the proposed model. Tables 3-4 also report the quantitative comparison of our CW-MSE and the computing FR IQA techniques using a computer with 3.40GHz CPU processor and 4.00GB memory. We can clearly observe that the proposed model is of fairly high computational speed, merely less than MSE and PAMSE.

In addition to high performance, it is worth emphasizing two advantages of the proposed CW-MSE. First, the design of our technique benefits from some principles. Image structure has an extremely important influence on the visual quality of human judgement and it has been widely used in most existing IQA models [16]-[28] and some related applications [7]-[13]. In [29], the authors have pointed out that using a proper Gaussian kernel to filter the input original and distorted images can effectively extract image structure. Note that this Gaussian kernel is globally unified. According to the recently revealed free energy based brain theory [32], a so-called internal generative model in the human brain assists to infer predictions of the meaningful information of input visual signals and avoid the residual uncertainty in a constructive manner. That is the brain will largely smooth the image regions with a great amount of residual information. Actually, the free energy based distortion metric (FEDM) [38] was recently proposed using the piecewise autoregressive (AR) mod-

Table 3 Mean computational time on the whole JPEG compressed images in CSIQ.

Models	MSE	SSIM	MS-SSIM	IW-MSE	IW-SSIM
Time (milliseconds)	5.2979	45.877	81.331	351.29	410.35
Models	FSIM	GSIM	IGM	PAMSE	CW-MSE
Time (milliseconds)	462.53	28.359	12491	5.5652	16.496

Table 4 Mean computational time on the whole JP2K compressed images in CSIQ.

Models	MSE	SSIM	MS-SSIM	IW-MSE	IW-SSIM
Time (milliseconds)	4.7693	45.522	81.808	350.95	412.67
Models	FSIM	GSIM	IGM	PAMSE	CW-MSE
Time (milliseconds)	462.17	27.073	12402	5.5895	16.612

el. Though the locally adaptive AR model can be used to predict visual quality well, but it works very slowly. Consequently, the CW-MSE that uses locally adaptive Gaussian kernel to advance MSE based on the above two points is designed with such a high performance index.

Second, it is easy to use and spread the proposed CW-MSE to existing systems. On one hand, our metric works in the block-based manner and only needs MSE (the simplest quality assessment method) and the original image, and thus has low computational load and is suitable for real-time applications. Particularly, our CW-MSE only computes the weighting map once in block-matching in compression, but the IW and other models require both whole original and compressed images and thus must perform many times, leading to massive computational cost. On the other hand, the CW-MSE is very portable since MSE has been incorporated into most existing image processing systems. In summary, our CW-MSE has achieved very inspiring results, and meanwhile, can overcome those main limitations of other existing FR IQA metrics.

Finally, we stress that the proposed CW-MSE uses the same kernel V for both original and distorted images. But in some cases, e.g. high compression ratio, there exists a significant distinction between the input lossless image and the associated compressed one. Two different kernels V are highly required if the computational time is not strongly limited. At this time we only use top, left, top-left and top-right blocks, as shown in Fig. 5.

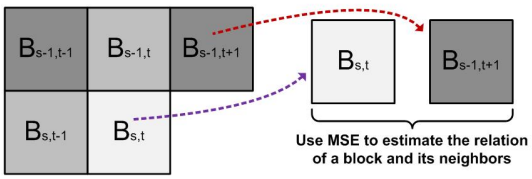


Fig. 5 Illustration of the content-weighted (CW) strategy for a particular block $B_{s,t}$ of size $M \times M$. The correlation between the current block $B_{s,t}$ and its surrounding four blocks are computed with the local MSE.

4 Conclusion

In this paper, we have first analyzed three main limitations of most existing FR IQA metrics used in coding. We then proposed a new Content-Weighted MSE (CW-MSE) to assess the quality of JPEG and JPGE2000 compressed images. The CW-MSE works by using locally adaptive Gaussian convolution to approximate the process of the HVS before computing the weighting map to improve MSE. Comparative studies of our technique are conducted on six image databases (LIVE, TID2008, CSIQ, IVC, Toyama and TID2013). Results of experiments confirm the superiority of the CW-MSE over state-of-the-art FR IQA approaches.

References

1. H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment Database Release 2," [Online]. Available: http://live.ece.utexas.edu/research/quality
2. N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008-A database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, pp. 30-45, 2009.
3. E. C. Larson and D. M. Chandler, "Categorical image quality (CSIQ) database," [Online], Available: http://vision.okstate.edu/csiq
4. A. Ninassi, P. Le Callet, and F. Autrusseau, "Subjective quality assessment-IVC database," [Online]. Available: http://www2.irccyn.ec-nantes.fr/ivcdb
5. Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. P. Sazzad, "MICT image quality evaluation database," [Online]. Available: http://mict.eng.u-toyama.ac.jp/mict/index2.html
6. N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. Jay Kuo, "Color image database TID2013: Peculiarities and preliminary results," *EUVIP2013*, pp.106-111, Jun. 2013.
7. A. Rehman and Z. Wang, "SSIM-based nonlocal means image denoising," *in Proc. IEEE Int. Conf. Image Process.*, pp. 217-220, Sept. 2011.

8. A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3378-3389, Aug. 2012.
9. S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-motivated rate distortion optimization for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 4, pp. 516-529, Apr. 2012.
10. S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "Perceptual video coding based on SSIM-inspired divisive normalization," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1418-1429, Apr. 2013.
11. T. Zhao, K. Zeng, A. Rehman, and Z. Wang, "On the use of SSIM in HEVC," *in Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, pp. 1107-1111, Nov. 2013.
12. H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657-667, Feb. 2013.
13. H. Yeganeh and Z. Wang, "High dynamic range image tone mapping by maximizing a structural fidelity measure," *in Proc. IEEE Int. Conf. Acous., Speech, Sig. Proc.*, pp. 1879-1883, May 2013.
14. K. Gu, G. Zhai, M. Liu, X. Yang, and W. Zhang, "Details preservation inspired blind quality metric of tone mapping methods," *in Proc. IEEE Int. Symp. Circuits and Syst.*, pp. 518-521, Jun. 2014.
15. Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it?-A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98-117, Jan. 2009.
16. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.
17. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," *in Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, pp. 1398-1402, Nov. 2003.
18. Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185-1198, May 2011.
19. L. He, X. Gao, W. Lu, X. Li, and D. Tao, "Image quality assessment based on S-CIELAB model," *Signal, Image and Video Processing*, vol. 7, no. 3, pp. 283-290, Sept. 2011.
20. K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new psychovisual paradigm for image quality assessment: From differentiating distortion types to discriminating quality conditions," *Signal, Image and Video Processing*, vol. 7, no. 3, pp. 423-436, May 2013.
21. E. Dumic, S. Grgic, and M. Grgic, "IQM2: new image quality measure based on steerable pyramid wavelet transform and structural similarity index," *Signal, Image and Video Processing*, Jun. 2014.
22. K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, "Structural similarity weighting for image quality assessment," *in Proc. IEEE Int. Conf. Multimedia and Expo Workshops*, pp. 1-6, Jul. 2013.
23. E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, Mar. 2010.
24. L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378-2386, Aug. 2011.
25. M. Narwaria, W. Lin, and A. Enis Cetin, "Scalable image quality assessment with 2D mel-cepstrum and machine learning approach," *Pattern Recognition*, vol. 45, no. 1, pp. 299-313, Jan. 2012.
26. A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500-1512, Apr. 2012.
27. J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 43-54, Jan. 2013.
28. K. Gu, G. Zhai, X. Yang, and W. Zhang, "An efficient color image quality metric with local-tuned-global model," *in Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, to appear.
29. W. Xue, X. Mou, L. Zhang, and X. Feng, "Perceptual fidelity aware mean squared error," *in Proc. IEEE Int. Conf. Comput. Vis.*, pp. 705-712, Dec. 2013.
30. E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annual Review of Neuroscience*, vol. 24, no. 1, pp. 1193-1216, 2001.
31. M. C. Morrone, J. Ross, D. C. Burr, and R. Owens, "Mach bands are phase dependent," *Nature*, vol. 324, pp. 250-253, Nov. 1986.
32. K. Friston, "The free-energy principle: A unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, pp. 127-138, 2010.
33. S. Wang, X. Zhang, S. Ma, and W. Gao, "Reduced reference image quality assessment using entropy of primitives," *in Picture Coding Symposium*, pp. 193-196, 2013.
34. A. Suhre, K. Kose, A. Enis Cetin, and M. N. Gurcan, "Content-adaptive color transform for image compression," *Optical Engineering*, vol. 50, no. 5, May 2011.
35. R. Soundararajan and A.C. Bovik, "Survey of information theory in visual quality assessment," *Signal, Image and Video Processing*, vol. 7, no. 3, pp. 391-401, May 2013.
36. N. Nikvand and Z. Wang, "Image distortion analysis based on normalized perceptual information distance," *Signal, Image and Video Processing*, vol. 7, no. 3, pp. 403-410, May 2013.
37. VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Mar. 2000, http://www.vqeg.org/.
38. G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41-52, Jan. 2012.