No-Reference Quality Assessment for JPEG Compressed Images

Yucheng Zhu, Guangtao Zhai, Ke Gu and Wenhan Zhu
Insti. of Image Commu. & Infor. Proce., Shanghai Jiao Tong University, Shanghai, China
{zyc420, zhaiguangtao, gukesjtuee, zhuwenhan823}@sjtu.edu.cn

***Abstract*—JPEG is a most commonly used standard of compression for digital images. Quality factor (Qfactor) for JPEG compressed image is actually a suitable indicator to the perceptual quality. However, the information of the compressor might be unknown due to various reasons. To evaluate the Qfactor, we recompress the formerly compressed image and measure the consistency between them. Then we define the fixed points (the points on the Qfactor-axis where the content of recompressed images are almost the same with that of directly compressed images) by following the Qfactor based specifications and form the image set. The quality of JPEG compressed images are measured by combining the estimated Qfactor with the features extracted from the image set. The experimental results confirm that the proposed image quality assessment technique, which is no-reference, is able to faithfully predict the visual quality of JPEG compressed images.**

***Keywords*—Image quality assessment (IQA), no-reference, JPEG compression, quality factor (Qfactor)**

I. Introduction

JPEG is a most commonly used compression method for digital images. It is an excellent way to store 24-bit photographic images. And it is widely used in multimedia signal processing, particularly for those images produced by digital photography. So the research of IQA for the JPEG compressed images is necessary. Numerous quality assessment approaches have been proposed over the last several decades. In current research of IQA, metrics are classified into full-reference (FR) [1], reduced-reference (RR) [2], and no-reference (NR) [3] methods by the degree of their accessibility to the reference image. And in the area of subjective assessment, many popular image quality databases have been proposed which can be used to validate the objective IQA metrics [4], [5], [6], [7], [8].

The JPEG compression scheme is divided into several stages, as shown in Fig. 1. Taking account of the eye's lesser sensitivity to chrominance information, JPEG uses fewer pixels for the chrominance channels. After the downsampling in chrominance channels, the image data will be divided into blocks of pixels. To discard an appropriate amount of information, after the DCT transform of pixel blocks, the compressor will quantize and round the DCT output value according to the quantization table. During compression, the injected blocking and blurring artifacts are the main factors to degrade the perceptual quality, and many methods predict the quality of JPEG compressed images by measuring the blocking and blurring.

QoMEX2017 – Erfurt, Germany; 978-1-5386-4024-1/17/$31.00 ©2017 IEEE

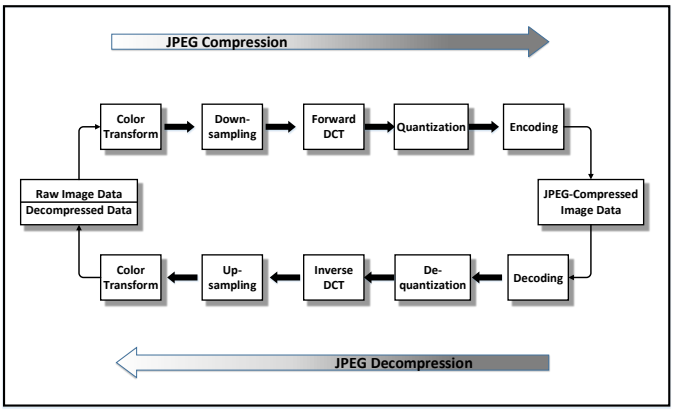

Fig. 1: JPEG compression and decompression.

More specifically, the JPEG compressor will quantize the DCT coefficients in the data unit independently. So the blocking artifacts will be introduced during the compression, and the severity of the distortion is related to the compression ratio. Lee and Park [9] found that the pixel values remained unchanged along the boundary but would change abruptly acorss the boundary. Liu and Heynderickx [10] estimated the masking effect of the texture and luminance, and they measured the local distortion by the strength of the gradient. Instead of calculating the discontinuity at boundaries of the block, Pan et al. [11] measureed the edge orientation change caused by blocking artifacts. Li et al. [12] evaluated blockiness by measuring the pseudo structures. In their method, the corner, block boundary and the color change were employed as three features to differentiate the artifact.

Each position in the DCT output block has its own quantization coefficient. The higher-order terms will be quantized more heavily than the low-order terms. The discarding of the high frequency coefficients must introduce blurring artifacts within blocks. Wang et al. [13] measured the blurring effect using two factors: one is the average absolute difference between in-block image samples, and the other is the zero-crossing rate of the differencing signal. In [14], Perra et al. extracted the edge of the image using Sobel filter, and then they measured the boundary block edges and inner block edges. In [15], Zhang and Bull predicted the quality of video by adaptively combining noticeable distortion and blurring artifacts using an enhanced nonlinear model.

The compression artifacts can also be evaluated in the transform domain. Bovik and Liu [16] constructed a new block

from two adjacent blocks in the spatial domain and measured the artifacts in the DCT domain. Chen and Bloom [17] computed the absolute difference between adjacent pixels along each column and applied one-dimensional discrete Fourier transform to measure the blockiness. Golestaneh and Chandler [18] counted the number of zero-valued DCT coefficients within the block, and they proposed the quality relevance map, which indicates whether the blocks are natural or compressed. Li et al. [19] found that high-odd-order Tchebichef kernels can capture the blocking artifacts in images.

In this paper, we devise an effective no-reference quality prediction method for JPEG compressed images via the estimation of Qfactor and content based features. We estimate the Qfactor by recompressing the formerly compressed image and measuring the consistency between them. Then we define some fixed points on the Qfactor-axis according to the recovered Qfactor and form the image set. Content based features are extracted from the image set with consideration of blocking artifacts, texture masking and pseudo-texture. Finally, the Qfactor and the content based features are combined to predict the perceptual quality.

The rest of this paper is organized as follows. Section II first presents the proposed IQA model. In Section III, the effectiveness of the algorithm is proved by the comparison with those existing relevant models. Finally, concluding remarks are given in Section IV.

II. No-Reference Quality Assessment for JPEG Compressed Images

A. The Qfactor Estimation

Ordinarily, the JPEG compression scheme applies a default set of quantization tables and scales them up or down according to the setting of Qfactor. The value of Qfactor will determine the quality of reconstructed image and the compressed file size. The perceptual quality and the size of compressed file are positively correlated to the value of Qfactor. During the compression, the ideal condition is that the Qfactor is appropriately set to achieve a balance point between file size and perceptual quality of the image.

Based on the above facts, Qfactor is an indicator to the compression ratio and can be applied to predict the quality of compressed images. However, the information of the compressor might be unknown due to various reasons. Under the condition that the quantization table is missing, and only the bitmap of JPEG compressed image is available, the Qfactor can still be accurately estimated, for instance, using the maximum likelihood approach [20], [21], [22]. In this paper, we devise a new Qfactor estimator with the fact that most JPEG files apply the standard quantization table as used in the popular implementation by the IJG (Independent JPEG Group). Given that the image has been previously compressed by JPEG, we would like to estimate the used quantization table and Qfactor. The following analysis and computation are on the basis of fact that most JPEG files apply the standard quantization table suggested by the JPEG standard [23]. For convenience, on the Qfactor-axis which ranges from 0 to 100, we define the point that equals to n as $\mathbf{q}_n$.

After DCT transform, the compressor will quantize each DCT output value according to quantization coefficients and rounds the result to an integer

$$F_Q(u,v) = IntegerRound(\frac{F(u,v)}{Q(u,v)}) \tag{1}$$

$$F^{'}(u,v) = Q(u,v)F_Q(u,v) \tag{2}$$

where $Q(u,v)$ denotes the $(u,v)th$ quantization step and $F_Q(u,v)$ denotes the $(u,v)th$ component of a quantized JPEG block. Under the ideal condition, coefficients occur only at the multiples of $Q(u,v)$, and $F^{'}(u,v)$ can be recalculated by the DCT from the decoded image block. Nevertheless, because of the rounding error, the DCT to the image block usually generates $F^{'}(u,v)^*$, which is not exactly $F^{'}(u,v)$, but an approximated version of it. The rounding error for each pixel can be assumed to be an independent and identically distributed (i.i.d.) random variable, which obeys a uniform distribution in the range of $[-0.5, 0.5)$. Then according to the 2-D DCT transform, the total rounding error is bounded by

$$\begin{aligned} |F^{'}(u,v) - F^{'}(u,v)^*| \leq \alpha(u)\alpha(v) \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \\ |0.5cos[\frac{\pi(2x+1)u}{2N}]cos[\frac{\pi(2y+1)v}{2N}]| \end{aligned} \tag{3}$$

$$where \qquad \alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & for \quad u = 0 \\ \sqrt{\frac{2}{N}} & for \quad u \neq 0 \end{cases} \tag{4}$$

When $u, v = 0$ *or* 4, the righthand of the Eq. 3 reaches its maximum which equals to 4. In other words, the total rounding error of each DCT coefficient is less than 4.

The summation of independent and identically distributed random variables can be well approximated as having a Gaussian distribution by the central limit theorem. Therefore, $F^{'}(u,v)$ should be approximately distributed as Gaussian. Because of the unitary nature of DCT, the mean of the Gaussian distribution is zero, while the variance of the Gaussian distribution is proportional to the variance of pixels in the block [24]. Most of the AC coefficients of 24-bit images distribute on the range from 0 to 100. If we recompress these images have been once compressed, the distribution will change. But when the Qfactor of the recompressor equals to that of the former compression process, only small amount of DCT coefficients with large rounding errors would be misclassified into the adjacent entry. By above analysis, we can recompress the original image and measure the consistency between them to estimate the Qfactor.

$$F_Q^*(u,v) = IntegerRound(\frac{F^{'}(u,v)^*}{Q(u,v)/2}) \tag{5}$$

In the second compression, we define the inconsistency between the image and its second compressed version as recompression distortion. When the second compression ratio equals to the first, the recompression distortion would reach its minimum. When we recompress the formerly compressed images using higher Qfactor, there would be a point where the

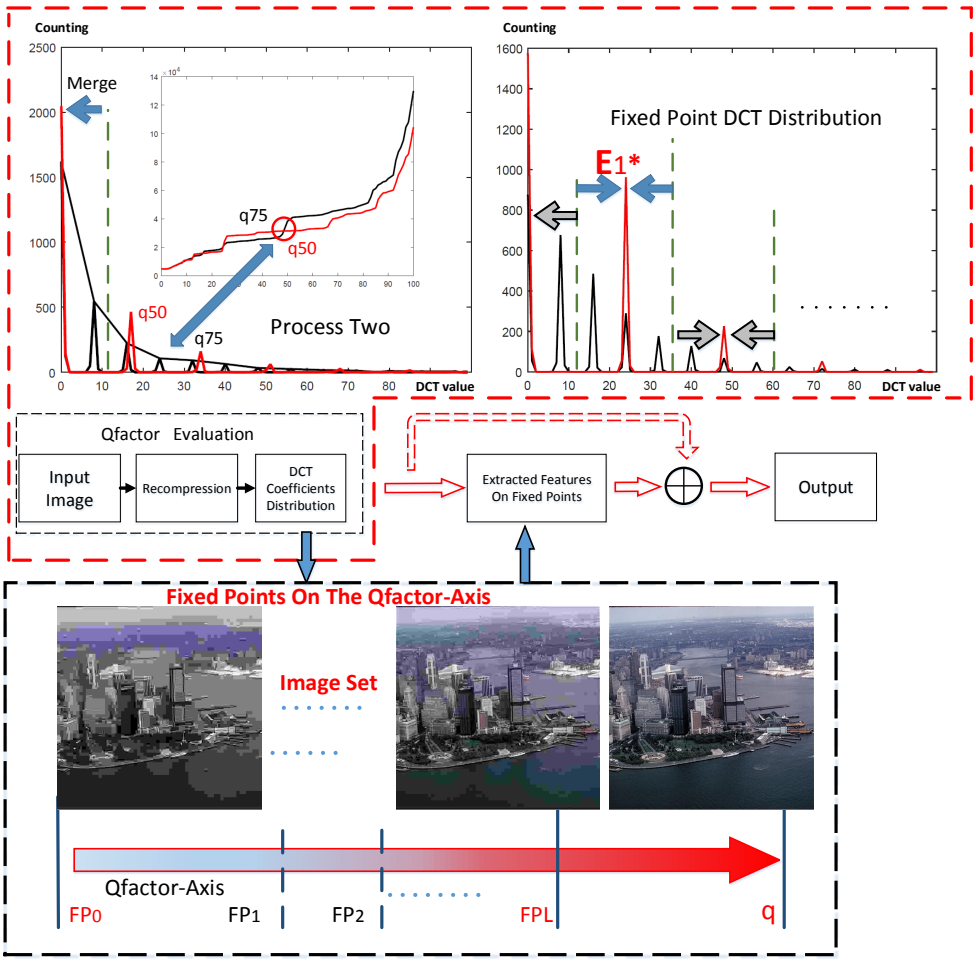

Fig. 2: *The diagram of the proposed IQA algorithm for JPEG compressed images: The upper part is about the estimation of the Qfactor, and the lower part is about the extraction of content based features from the image set. The final measurement of the quality is devised by combining the Qfactor with the extracted content-based features from the image set, which is formed by fixed points on the Qfactor-axis.*
The left top graph shows the histogram of DCT coefficients and the changing process during the recompression (Process Two: $E_0 \rightarrow E_0^* \leftarrow E_1, E_2 \rightarrow E_1^* \leftarrow E_3, \ldots$). And the red circled point is the LP1 ($\mathbf{q}_{50}$) for the black line ($\mathbf{q}_{75}$). This can be used to differentiate the $\mathbf{q}_{50}$ (red line) from its RP (black line $\mathbf{q}_{75}$). The right top graph illustrates the change process of DCT coefficients distribution on the fixed points.

quantization step in the table is half of that of the first compression. As shown in Eq. 1 and Eq. 5, the $F_Q^*(u,v) \approx 2F_Q(u,v)$. And in this case, $F_Q^*(u,v)Q(u,v)/2 \approx 2F_Q(u,v)Q(u,v)$. So the coefficients after the dequantization are almost the same with that of the first compression. For example, when an image with $\mathbf{q}_{50}$ is recompressed by $\mathbf{q}_{75}$, the recompression distortion between the original image and the recompressed image is small. When an image with $\mathbf{q_1}$ is recompressed using $\mathbf{q_2}$, if every quantization step in the table of the $\mathbf{q_2}$ is approximately half of the $\mathbf{q_1}$'s, the dct coefficients of the recompressed image will be almost unchanged. For $\mathbf{q_2} > \mathbf{q_1}$, we define $\mathbf{q_2}$ as $\mathbf{q_1}$'s right point (RP).

We recompress the image and the size of file decreases when we use a lower Qfactor. Most AC coefficients of 24-bit images range from 0 to 100, and the width of the DCT coefficient distribution shrinks as we go to higher frequencies. In the range, when Qfactor goes down to a certain value, the quantized coefficients will first undergo process one to switch their positions ($E_0^* = E_0, E_1 \rightarrow E_1^* \leftarrow E_2, \ldots$). And then after the small decrease of Qfactor, the quantized coefficients will undergo process two ($E_0 \rightarrow E_0^* \leftarrow E_1, E_2 \rightarrow E_1^* \leftarrow E_3, \ldots$) where the E_n and E_n^* are the nth DCT entries of the first and second compressions, as shown in Fig. 2. The two processes occur shortly near around the switching point and will enable the compressor use shorter bitstrings, which corresponds to the codes of shorter length in the Huffman table. This means that the two processes will cause a large and abrupt decrease of file size.

When we recompress the image along the Qfactor-axis from value $\mathbf{q}_a$ to zero, there are some points where the filesize will abruptly decrease. We define such kind of points as the Left Points ($LP_{s=1,2\ldots}$) of $\mathbf{q}_a$. LP_1 is adjacent to $\mathbf{q}_a$. In most cases, $\mathbf{q}_a$ is the RP of its LP_1. More specifically, when we estimate the Qfactor of one compressed image, we may get the RP of the image as the result. To differentiate the RP from the real Qfactor, we can compute the image's LP_1. For example, as shown in the Fig. 2, when we estimate the Qfactor of $\mathbf{q}_{50}$ by traversing the axis, the recompression distortion of two points ($\mathbf{q}_{50}$ and $\mathbf{q}_{75}$) are small. But the LP_1 of $\mathbf{q}_{50}$ is $\mathbf{q}_{25}$ and the LP_1 of $\mathbf{q}_{75}$ is $\mathbf{q}_{50}$. After calculation, there is an abrupt change of file size at $\mathbf{q}_{25}$ rather than $\mathbf{q}_{50}$. So the LP_1 of the image is $\mathbf{q}_{25}$ and the real Qfactor is $\mathbf{q}_{50}$. By detecting the LP_1, we can get the accurate Qfactor.

B. Content-Based Degradation Measures

When we recompress the image, the recompression distortion will be introduced. Compared with the image $I_{\mathbf{q}_j}$ which is directly compressed to $\mathbf{q}_j$, the recompressed image ($I_{\mathbf{q}_i \rightarrow \mathbf{q}_j}$) undergoes two compressions which are from $\mathbf{q}_i$ to $\mathbf{q}_j$. So there exists difference between contents of two images even though they are of the same Qfactor after the compression. However,

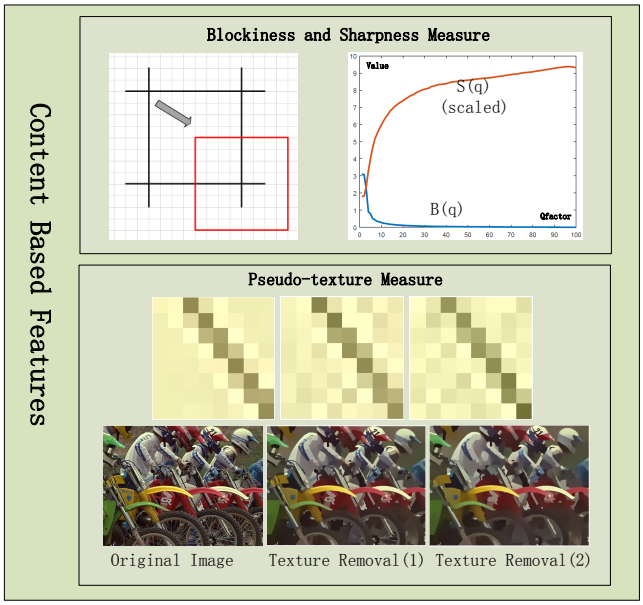

Fig. 3: Content based features.

when we recompress the image, there will be several points $\mathbf{q}_s$ for $\mathbf{q}_i$. On these points, the $I_{\mathbf{q}_i \to \mathbf{q}_s}$ are almost the same with $I_{\mathbf{q}_s}$. It can also be explained through the change of DCT coefficients distribution, as illustrated in Fig. 2. These fixed points are spaced by adjacent LPs, and $\mathbf{q}_0$ is one fixed point for all Qfactors. We define the left fixed point which is closest to the LP_1 as FPL and define $\{FPL, FP0 = \mathbf{q}_0\}_{\mathbf{q}}$ as the fixed point set for $\mathbf{q}$. The following analysis is for the directly compressed images and the images in the set share the same properties.

Entropy is a statistical measure of randomness. We calculate the entropy of compressed images using Eq. 6

$$Entropy\,of\,I_{\mathbf{q}} = -\sum_{i=1}^{N} p[C_{\mathbf{q}}(i)] \log p[C_{\mathbf{q}}(i)] \tag{6}$$

Where $p[C_{\mathbf{q}}(i)]$ is the the probability for ith luminance value in image of Qfactor $\mathbf{q}$. The entropy of image indicates the gross information content which varies with Qfactors. The curves are steep when the Qfactor is low and are flat when the Qfactor is high, which means the amount of gross information content is distinguishable in the low-value range and is difficult to predict in the high-value range. In order to differentiate image quality of high Qfactors, we extract the content based features from the image set.

The JPEG compressor will quantize the DCT coefficients in the data unit independently. So the blocking artifacts will be introduced after the compression. Meanwhile, the discard of the high frequency coefficients will introduce blurring artifacts within blocks. To measure the blurring and blocking artifacts, we calculate the standard deviation within the block. Then as shown in Fig. 3, we place the left-up corner of 8×8 block at the center of the original block and calculate the standard deviation of the new block. We use ($\overline{Std_o}$, $\overline{Std_m}$) to denote the averaged standard deviations of the two kinds of blocks. The blockiness $B(I)$ and sharpness $S(I)$ of image I are measured as Eq. 7. As shown in Fig. 3, when we compress

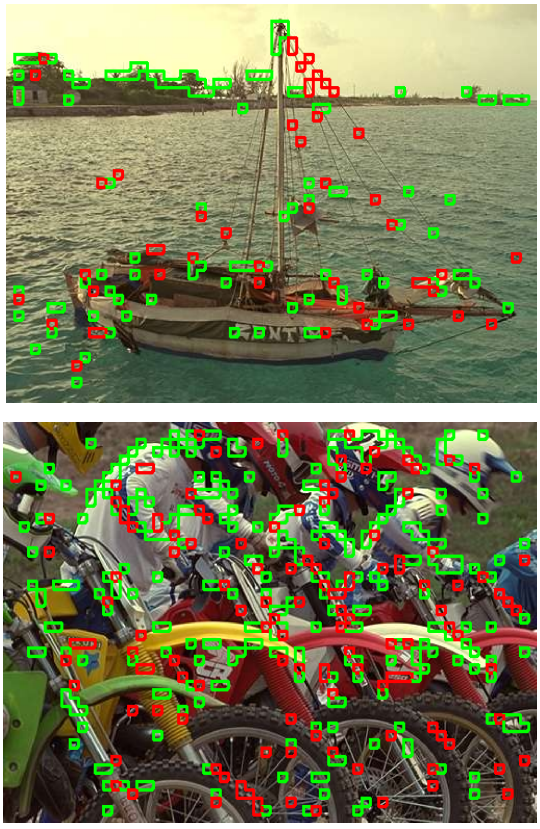

Fig. 4: Areas that are defined in Table II: Red areas belong to $\mathbf{P}$. Green areas belong to $\mathbf{N}$.

I with different Qfactors, the two curves change monotonously.

$$\begin{aligned} B(I_{\mathbf{q}}) &= (\overline{Std_m(I_{\mathbf{q}})} - \overline{Std_o(I_{\mathbf{q}})})/\overline{Std_o(I_{\mathbf{q}})} \\ S(I_{\mathbf{q}}) &= (\overline{Std_o(I_{\mathbf{q}})} - \overline{Std_o(I_{FP0_{\mathbf{q}}})})/\overline{Std_o(I_{FP0_{\mathbf{q}}})} \end{aligned} \tag{7}$$

More details are added when the compressor uses the higher Qfactor, so does the pseudo-texture. This kind of distortion will occur when an image with distinct edges is compressed. Unfortunately, photographic images are such kind of cases because of the image signal processor (ISP) in digital cameras. To enhance the sharpness of images, the ISP will use halos to emphasize the transitions between the dark and light areas [25]. But if the 8×8 block contains the sharpened edges, there will easily appear the pseudo-texture when images are compressed, as shown in Fig. 3. The light and black areas around the edge will mix up. And in the flat area, these pseudo-texture will be noticeable and degrade the quality. The synthetic images will be worse because there is much pure color area in images where the pseudo-texture will be more noticeable.

The edge of main structure will produce much pseudo-texture after compression. Considering the masking effect of the complex texture, we extract the meaningful structures from textured surfaces using the method proposed in [26]. As shown in Fig. 3, the edge and the main structures are well preserved. And by adjusting the parameters, we can remove different amount of texture. What's more, some other areas are also liable to cause noticeable pseudo-texture. Table II gives the detailed calculation to get the average of gradient magnitude

TABLE I: PERFORMANCE COMPARISONS ON IMAGE DATABASES. TOP-THREE PLCC AND SROCC RESULTS ARE MARKED IN BOLD.

Database		Bovik [16]	Wang [13]	Perra [14]	Pan [11]	Liu [10]	Chen [17]	Lee [9]	NJQA [18]	Li [12]	PSS [28]	Ours Qfactor	Ours All
LIVE	PLCC	0.931	0.937	0.873	0.889	0.957	0.916	0.942	**0.963**	**0.967**	0.936	0.958	**0.971**
	SROCC	0.924	0.931	0.869	0.872	0.938	0.905	0.930	**0.956**	**0.957**	0.924	0.946	**0.961**
	RMSE	8.834	7.860	15.543	11.113	7.040	9.714	8.120	8.623	8.259	8.531	5.823	5.812
CSIQ	PLCC	0.969	**0.979**	0.896	0.889	0.811	0.942	0.977	0.954	0.949	**0.981**	0.978	**0.984**
	SROCC	0.945	**0.953**	0.852	0.863	0.924	0.923	0.948	0.925	0.931	0.952	**0.958**	**0.960**
	RMSE	0.076	0.068	0.138	0.140	0.179	0.103	0.066	0.092	0.097	0.059	0.058	0.053
TID2013	PLCC	0.931	0.951	0.817	0.864	0.936	0.912	0.921	0.948	0.916	**0.959**	**0.957**	**0.962**
	SROCC	0.882	**0.927**	0.762	0.814	0.862	0.856	0.876	0.886	0.865	0.901	**0.939**	**0.955**
	RMSE	0.549	0.473	0.868	0.758	0.532	0.618	0.586	0.481	0.606	0.429	0.295	0.288
SIQAD	PLCC	0.466	0.741	0.311	0.420	0.505	0.130	0.742	0.621	0.623	**0.763**	**0.798**	**0.802**
	SROCC	0.465	0.739	0.241	0.399	0.413	0.163	0.746	0.615	0.626	**0.756**	**0.766**	**0.795**
	RMSE	8.312	6.318	8.931	8.527	8.109	9.317	6.296	7.357	7.343	6.079	6.042	5.625

in areas that are liable to cause pseudo-texture. Fig. 4 shows some examples of areas that are defined in Table II.

TABLE II: MEASUREMENT OF THE PSEUDO-TEXTURE.

Measurement of the pseudo-texture (input image $I_{\mathbf{q}}, I_{FPL_{\mathbf{q}}}$, output average of gradient magnitude $E(I_q)$.)
1. $\mathbf{O} = \{$ $I_{FPL_{\mathbf{q}}}$ is grouped into 8×8 blocks $\}$.
2. Calculate the sum of gradient magnitude ($\mathbf{GM}$) in each block.
$\mathbf{GM} = \sqrt{[\mathbf{I} \bigotimes \mathbf{h_x}]^2 + [\mathbf{I} \bigotimes \mathbf{h_y}]^2}$,
where $\bigotimes$ indicates the convolution operation.
$\mathbf{h_x} = \frac{1}{16}\begin{bmatrix} +3 & 0 & -3 \\ +10 & 0 & -10 \\ +3 & 0 & -3 \end{bmatrix}$ $\mathbf{h_y} = \frac{1}{16}\begin{bmatrix} +3 & +10 & +3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix}$
3. $\mathbf{C} = \{Block(i,j) \mid Block(i,j) \in \mathbf{O}$, and $\sum_{Block(i,j)} GM(x,y) < T_1\Big\}$.
4. $\mathbf{N(i,j)} = \{Block(i+\delta_i, j+\delta_j) \mid \delta_i, \delta_j \in \{-1,0,1\}$, and $(\delta_i, \delta_j) \neq (0,0)\}$.
5. $\mathbf{D1} = \{Block(i,j) \mid Block(i,j) \in \mathbf{C}$, and $\exists B \in \mathbf{N(i,j)}$ s.t. $\sum_B GM(x,y) > T_2$ $\Big\}$.
6. Apply $1-5$ to $I_{\mathbf{q}}$ and get $\mathbf{D2}$. Let $\mathbf{P} = \mathbf{D2} - \mathbf{D1}$, $\mathbf{N} = \mathbf{D1} - \mathbf{D2}$.
7. Apply the structure extraction [26] to $I_{\mathbf{q}}$ and get $\mathbf{S}$.
8. $\mathbf{E(I_q)} = \frac{1}{M}(\sum GM(\mathbf{S}) + \sum GM(\mathbf{P}) - \sum GM(\mathbf{N}))$

When Qfactor is small, the blur and blockiness are the major artifacts. As we increase the Qfactor, the blockiness and blur become more and more inconspicuous. But the pseudo-texture will appear and be the main factor to degrade the quality. We use Eq. 8 as a quality prediction of image I with Qfactor $\mathbf{q}$.

$$Q(I_{\mathbf{q}}) = \begin{cases} \mathbf{q} + \min\left[\alpha(S(I_{\mathbf{q}}) - B(I_{\mathbf{q}})), 1\right] for \mathbf{q} < \mathbf{q}_0 \\ \mathbf{q} - \min\left[\beta E(I_{\mathbf{q}}), 5\right] for \mathbf{q} \geq \mathbf{q}_0 \end{cases} \quad (8)$$

We use the image set $\{FPL, FP0 = \mathbf{q}_0\}_{\mathbf{q}}$ to modify the prediction.

$$Qf(I_{\mathbf{q}}) = Q(I_{\mathbf{q}}) + \gamma Q(I_{FPL_{\mathbf{q}}}) \quad (9)$$

where the $Qf(I_{\mathbf{q}})$ is the final quality measure of the JPEG compressed image $I_{\mathbf{q}}$.

III. EXPERIMENTAL RESULTS AND ANALYSIS

The JPEG subset of four databases, including LIVE [4], CSIQ [5], TID2013 [6], and one screen content image database SIQAD [27], are used as testing beds. The LIVE database has 233 JPEG compressed images which are saved as the BMP format. The distorted images were generated by compressing the reference images (full color) using JPEG at bit rates ranging from 0.15 bpp to 3.34 bpp. The CSIQ database contains 150 JPEG compressed images of five compression ratio levels which are saved as the PNG format. The TID2013 database includes 125 JPEG compressed images of five distortion levels and uses BMP format to store the images. The SIQAD database contains 140 JPEG compressed images of seven degradation levels which are of the PNG format.

The quality factor parameter is not available in LIVE, CSIQ and TID2013. But the SIQAD provides the Qfactor of the seven distortion levels. We compare the obtained Qfactor from our algorithm with the given values from the SIQAD. The experimental results show that the estimation of Qfactor on the SIQAD is absolutely correct. After that, we compress different natural images using Qfactor from 1 to 100 and use our method to estimate the Qfactor. The experimental results show that the Qfactor under 95 can be accurately predicted.

In our experiments, two commonly used metrics are employed to validate the performance. Pearson linear correlation coefficient (PLCC) measures the linear correlation between two variables, while Spearman's rank-order correlation coefficient (SROCC) is a measure of the rank correlation. We present the absolute value of PLCC and SROCC. So the higher values indicate better performance.

In this paper, we compare the proposed model with several state-of-the-art NR JPEG IQA methods: Bovik [16], Wang [13], Perra [14], Pan [11], Liu [10], Chen [17], Lee [9], NJQA [18], Li [12] and PSS [28]. We first use the estimated Qfactor to predict the quality of JPEG compressed images and then add the content based features to validate their effectiveness. As shown in Table I, the Qfactor and the overall technique achieve high performance on the databases. For natural scene images, our algorithm is comparable to the best performed metrics. And for screen content images, our algorithm shows the best performance.

Despite the good performance of Qfactor and the Qfactor-based metric, it is also important to note that the Qfactor, as we have analyzed in this paper, helps to measure the quality of JPEG compressed images only under condition that IJG

scaling scheme is employed. In real world, fortunately, the IJG implementation is dominant in compression scheme. For other compression methods, similar experiments should be conducted to validate the effectiveness of quality factor. And the comparisons will clarify whether the good performance of Qfactor comes from the appropriate scaling method designed by IJG.

Also when the image is compressed for multiple times, the Qfactor may not work well. Since Qfactor of the last compression is available, the estimation of previous Qfactor levels should be conducted. Similar case scenario includes images being contaminated by other kinds of distortions, such as blurring and noise injection, along with JPEG compression. Under this condition, apart from Qfactor, other types of metrics should be incorporated into the frame. In fact, multiple-distortion is another emerging research direction for perceptual quality assessment [29]. Nevertheless, in this paper, rather than proposing a method considering all these factors, our metric highlights the fact that, it is a good strategy to incorporate quality factor into the quality assessment scheme for JPEG compressed images.

IV. Conclusion

In this paper, we propose a quality prediction metric for the JPEG compressed images via the estimation of quality factor (Qfactor) and content based features. We evaluate the Qfactor by measuring the consistency between the directly compressed images and the recompressed images. Then the image set is formed by following the Qfactor based specifications, and the content based features of the image set are extracted. The quality is measured by combining the predicted Qfactor and the features. The experimental results demonstrate the effectiveness of our method, which is no-reference and consistent across different types of image content, and highlight the fact that, it is a good strategy to incorporate quality factor into the quality assessment scheme for JPEG compressed images.

Acknowledgment

This work was supported in part by the National Science Foundation of China under Grants 61422112, 61371146, 61521062, 61527804, the Foundation for the Author of National Excellent Doctoral Dissertation of China under Grant 201339, Science and Technology Commission of Shanghai Municipality under Grant 15DZ0500200 and National High-tech R&D Program of China under Grant 2015AA015905.

References

[1] Zhou Wang, A C Bovik, H R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[2] Guangtao Zhai, Xiaolin Wu, Xiaokang Yang, Weisi Lin, and Wenjun Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 41–52, 2012.

[3] Ke Gu, Guangtao Zhai, Xiaokang Yang, and Wenjun Zhang, "Using free energy principle for blind image quality assessment," *IEEE Transactions on Multimedia*, vol. 17, no. 1, pp. 50–63, 2015.

[4] H R Sheikh, Wang Z, Cormack L, and A C Bovik, "Live image quality assessment database release 2," [Online]. Available:http://live.ece.utexas.edu/research/quality.

[5] Eric C Larson and Damon M Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 11006, 2010.

[6] Nikolay N Ponomarenko, Lina Jin, Oleg Ieremeiev, Vladimir V Lukin, Karen Egiazarian, Jaakko Astola, Benoit Vozel, Kacem Chehdi, Marco Carli, Federica Battisti, et al., "Image database tid2013: Peculiarities, results and perspectives," *Signal Processing-image Communication*, vol. 30, pp. 57–77, 2015.

[7] Xiongkuo Min, Guangtao Zhai, Zhongpai Gao, and Ke Gu, "Visual attention data for image quality assessment databases," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 2014, pp. 894–897.

[8] Y. Zhu, G. Zhai, K. Gu, and M. Liu, "Blindly evaluating stereoscopic image quality with free-energy principle," in *IEEE International Symposium on Circuits and Systems*, May 2016, pp. 2222–2225.

[9] Sangwoo Lee and Sang Ju Park, "A new image quality assessment method to detect and measure strength of blocking artifacts," *Signal Processing-image Communication*, vol. 27, no. 1, pp. 31–38, 2012.

[10] Hantao Liu and Ingrid Heynderickx, "A perceptually relevant no-reference blockiness metric based on local image characteristics," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, pp. 263540, 2009.

[11] Feng Pan, X Lin, Susanto Rahardja, Ee Ping Ong, and W S Lin, "Using edge direction information for measuring blocking artifacts of images," *Multidimensional Systems and Signal Processing*, vol. 18, no. 4, pp. 297–308, 2007.

[12] Leida Li, Weisi Lin, and Hancheng Zhu, "Learning structural regularity for evaluating blocking artifacts in jpeg images," *IEEE Signal Processing Letters*, vol. 21, no. 8, pp. 918–922, 2014.

[13] Zhou Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of jpeg compressed images," in *IEEE International Conference on Image Processing*, 2002, vol. 1, pp. I–477–I–480.

[14] C. Perra, F. Massidda, and D. D. Giusto, "Image blockiness evaluation based on sobel operator," in *IEEE International Conference on Image Processing*, Sept 2005, vol. 1, pp. I–389–92.

[15] F Zhang and David Bull, "A perception-based hybrid model for video quality assessment," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 6, pp. 1017–1028, 2016.

[16] A. C. Bovik and Shizhong Liu, "Dct-domain blind measurement of blocking artifacts in dct-coded images," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001, vol. 3, pp. 1725–1728.

[17] Chunhua Chen and Jeffrey A Bloom, "A blind reference-free blockiness measure," *in Pacific Rim Conference on Multimedia*, 2010, pp. 112–123.

[18] S. A. Golestaneh and D. M. Chandler, "No-reference quality assessment of jpeg images via a quality relevance map," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 155–158, Feb 2014.

[19] Leida Li, Hancheng Zhu, Gaobo Yang, and Jiansheng Qian, "Referenceless measure of blocking artifacts by tchebichef kernel analysis," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 122–125, 2014.

[20] Zhigang Fan and R. L. de Queiroz, "Identification of bitmap compression history: Jpeg detection and quantizer estimation," *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 230–235, Feb 2003.

[21] R. Neelamani, R. de Queiroz, Zhigang Fan, S. Dash, and R. G. Baraniuk, "Jpeg compression history estimation for color images," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1365–1378, June 2006.

[22] W. Luo, J. Huang, and G. Qiu, "Jpeg error analysis and its applications to digital image forensics," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 480–491, Sept 2010.

[23] G. K. Wallace, "The jpeg still picture compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, Feb 1992.

[24] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the dct coefficient distributions for images," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1661–1666, Oct 2000.

[25] Y. Zhu, G. Zhai, K. Gu, and Z. Che, "No-reference image quality assessment for photographic images of consumer device," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2016, pp. 1085–1089.

[26] Li Xu, Qiong Yan, Yang Xia, and Jiaya Jia, "Structure extraction from texture via relative total variation," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 139, 2012.

[27] Huan Yang, Yuming Fang, and Weisi Lin, "Perceptual quality assessment of screen content images," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4408–4421, 2015.

[28] X. Min, G. Zhai, K. Gu, Y. Fang, X. Yang, X. Wu, J. Zhou, and X. Liu, "Blind quality assessment of compressed images via pseudo structural similarity," in *IEEE International Conference on Multimedia and Expo*, July 2016, pp. 1–6.

[29] Guangtao Zhai, "Recent Advances in Image Quality Assessment," in *Visual Signal Quality Assessment*, Springer International Publishing, 2015, pp. 73–97.