No-Reference Image Quality Assessment Based on Localized Gradient Statistics: Application to JPEG and JPEG2000

Hantao Liu*[a], Judith Redi[b], Hani Alers[a], Rodolfo Zunino[b], Ingrid Heynderickx[a,c]
[a]Department of Mediamatics, Delft University of Technology, Delft, The Netherlands
[b]Department of Biophysical and Electronic Engineering, University of Genoa, Genoa, Italy
[c]Group Visual Experiences, Philips Research Laboratories, Eindhoven, The Netherlands

ABSTRACT

This paper presents a novel system that employs an adaptive neural network for the no-reference assessment of perceived quality of JPEG/JPEG2000 coded images. The adaptive neural network simulates the human visual system as a black box, avoiding its explicit modeling. It uses image features and the corresponding subjective quality score to learn the unknown relationship between an image and its perceived quality. Related approaches in literature extract a considerable number of features to form the input to the neural network. This potentially increases the system's complexity, and consequently, may affect its prediction accuracy. Our proposed method optimizes the feature-extraction stage by selecting the most relevant features. It shows that one can largely reduce the number of features needed for the neural network when using gradient-based information. Additionally, the proposed method demonstrates that a common adaptive framework can be used to support the quality estimation for both compression methods. The performance of the method is evaluated with a publicly available database of images and their quality score. The results show that our proposed no-reference method for the quality prediction of JPEG and JPEG2000 coded images has a comparable performance to the leading metrics available in literature, but at a considerably lower complexity.

Keywords: Image quality assessment, objective metric, JPEG, JPEG2000, neural network

1. INTRODUCTION

The development of electronic imaging and multimedia techniques has pushed the demand for reliable quality assessment. Understanding and evaluating image quality has become increasingly important for a broad range of applications, such as the optimization of digital imaging systems, the benchmarking of image and video coding algorithms, and the quality monitoring and control in displays [1]. Traditionally, image quality has been evaluated by human subjects, and a mean opinion score (MOS) has been used to represent the image quality perceived by an averaged viewer. When conducted properly, subjective experiments are considered as the most reliable means of assessing image quality. However, performing subjective experiments is very time-consuming, and as a consequence, very expensive and too slow to be useful in real-world applications. Therefore, during the last decades, a lot of research effort has been devoted to the development of objective metrics that automatically and quantitatively can predict image quality as perceived by an averaged viewer.

Objective metrics reported in literature range from dedicated metrics that measure a specific image distortion to general metrics that assess the overall perceived quality. Both the dedicated and general metrics can be classified into full-reference (FR) or no-reference (NR) metrics, depending on whether the distorted image is compared to the original image or video. FR metrics measure the similarity or fidelity between the distorted image and its original version, where the latter is considered as a distortion-free reference. The most widely used FR metrics are the mean squared error (MSE) and the peak signal-to-noise ratio (PSNR), both aiming for an overall quality assessment. Improved alternatives of these two basic general metrics include e.g. the structural similarity (SSIM) index [2] and the visual information fidelity (VIF) index [3]. Since FR metrics require the access to the original, which is not available in (most) real-time systems, their applicability is limited to in-lab (off-line) testing of image and video processing algorithms. Instead for real-time applications, NR metrics, in which the quality prediction is based on the distorted image only, i.e. without any reference, are more practical. Designing NR metrics, however, is still challenging mainly due to the limited understanding of how the human visual system (HVS) affects image quality assessment.

Human Vision and Electronic Imaging XV, edited by Bernice E. Rogowitz, Thrasyvoulos N. Pappas,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7527, 75271F · © 2010 SPIE-IS&T · CCC code:
0277-786X/10/$18 · doi: 10.1117/12.838982

During the last decades, there is considerable progress in the development of NR metrics, as can be seen from some successful methods reported in the literature [4]-[11]. A large number of NR metrics, such as e.g. in [4]-[7], are dedicated metrics measuring a specific type of artifact created by a specific image distortion process, such as a metric measuring sensor noise, ringing or blockiness as a consequence of signal compression, or blur caused during acquisition. In such a scenario, the design of the NR metric can make use of the specific characteristics of the artifact, and therefore, generally obtains a higher reliability with respect to the perceived quality degradation. The design of specific NR metrics is particularly beneficial for e.g. video chain optimization [12]. For the prediction of the overall perceived quality, approaches such as pooling the local distortions into an overall quality score, and combining different artifacts inherent in an image have been reported [6] and [10], but are studied only to a limited extent. In [11], natural scene statistics were used to blindly measure the overall quality of images compressed by JPEG2000. The approach relied on the assumption that typical natural images exhibit strong statistical regularities, and therefore, reside in a tiny area of the space containing all possible images. Based on this assumption, the approach quantified image quality by detecting variations in the statistics of image features in the wavelet domain.

Instead of precisely modeling specific types of artifacts or natural scene statistics, some approaches such as e.g. in [8] and [9] attempt to formulate NR image quality assessment as a machine learning problem. They treat the HVS as a black box, whose input-output relationship between image characteristics and a quality rating is to be learned by computational intelligent tools, such as neural networks. The problem is generally formulated as a regression or function approximation approach, and the training data are obtained from extensive subjective experiments. The goal is to train the model so that the error between the desired output (i.e. the subjective quality rating) and the model prediction is minimized. The approach was proved to be effective for the overall quality prediction of JPEG compressed images (see [8] and [9]), but at the expense of the extraction of a considerable number of image features, such as general pixel-based features in [8] or HVS-based features in [9], as input to the neural network. This potentially increases the model's complexity, and consequently, may affect the prediction accuracy of the metric.

Because of the widespread use of compression, a lot of research is devoted to measuring the image quality after JPEG and JPEG2000 coding and decoding. In this paper, we further rely on the approach of using neural networks for the NR assessment of perceived overall quality of JPEG/JPEG2000 coded images. We optimize the feature-extraction stage by efficiently selecting and calculating the most relevant features, thus providing a simple yet efficient alternative for real-time implementation. It should be noted that the whole process is built up on the luminance component of images only in order to further reduce the computational load.

2. FEATURE EXTRACTION

Figure 1. Illustration of blocking and blur artifacts in a JPEG and a JPEG2000 compressed image, respectively.

Literature has shown that the image quality of JPEG compressed images is highly correlated with the occurrence of blocking artifacts [7]; while the quality of JPEG2000 compressed images is highly correlated with the occurrence of blur [10]. A blocking artifact manifests itself as an artificial discontinuity in the image content, which is a direct consequence of the fact that the quantization in JPEG is block-based and that the blocks are quantized independently [13]. A blur artifact occurs in JPEG2000 mainly due to the loss of high frequency transform coefficients in the wavelet-based coding, as a result of which the image signal is smoothened [10]. Figure 1 illustrates the occurrence of blocking artifacts in a JPEG compressed image, and of blur artifacts in a JPEG2000 compressed image.

Quality degradation as a consequence of compression would hence easily be predictable from the extraction of blockiness and blur related image features, provided that an adaptive tool such as a neural network is used to empirically learn the highly non-linear relationship between these features and the quality rating. To efficiently characterize the local behavior of artifacts and thus to feed the neural network with relevant features for image quality prediction, a gradient-based feature-extraction scheme is proposed. It contains three basic components: (1) the localization of artifacts, (2) the local feature extraction using local gradients in relation to their neighborhood, and (3) the assembling of global statistical descriptors as inputs to the neural network. How to implement each of these steps for the quality prediction of JPEG and JPEG2000 compressed images is detailed below.

2.1 Local feature extraction: JPEG

Due to the underlying algorithm for JPEG compression, the spatial location of blocking artifacts is very regular. In principle, they occur on a grid of blocks of 8×8 pixels, starting at the top-left corner of an image. In real-life applications, however, the grid size may be different and its starting position may be shifted due to deviations in the incoming signal or as a consequence of spatial scaling. As a consequence, a NR metric runs the risk of calculating blockiness at wrong pixel positions, which consequently degrades its accuracy. To ensure that the metric is calculated at the exact position of the block boundaries, a grid detector is adopted. The blocking grid detection method proposed in [7] is implemented in this paper, but it should be noted that the feature extraction approach proposed here is independent of the particular choice of grid detector, and so, any other alternative can be conveniently used instead. The blocking grid detection method of [7] first maps an image onto a 1-D signal profile, in which the periodic property of blocking artifacts is maintained. Then the exact block size as well as the grid offset are easily extracted from the discrete Fourier transform (DFT) of this 1-D signal profile.

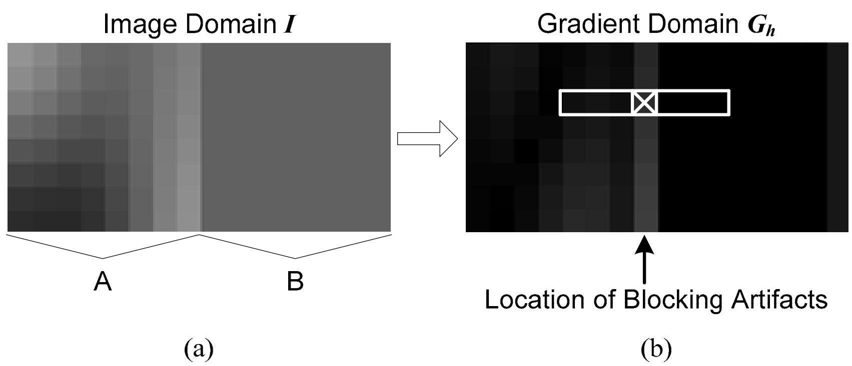

Figure 2. Illustration of the template for calculating the local blockiness: (a) two adjacent 8×8 blocks (i.e. A and B) extracted from a real JPEG image, and (b) the gradient profile of the image patch of (a).

When the blocking artifacts are (exactly) located, their local behavior can be explicitly modeled. In this paper, the local feature extraction for the JPEG compressed images is based on calculating the visual strength of a blocking artifact within a local area of the image content [7]. Since a blocking artifact is a local edge that stands out from its spatial vicinity, it can be simply defined relating the energy present in the gradient at the artifact to the energy present in the gradient in its neighboring pixels. When we denote the luminance channel of an image of M×N (height × width) pixels as I(i, j) for i∈[1, M], j∈[1, N], the local blockiness $L_{blockiness\text{-}h}$ along the horizontal direction at location (i, j) is quantified as

$$L_{blockiness\text{-}h}(i,j)=\frac{G_h(i,j)}{\frac{1}{2n}\sum_{x=-n,\ldots,n,x\neq 0} G_h(i,j+x)} \qquad (i,j)\in\{\text{location of blocking}\} \tag{1}$$

where $G_h(i, j)$ indicates the gradient map along the horizontal direction, and it is computed as

$$G_h(i,j)=\left|I(i,j+1)-I(i,j)\right| \; j\in[1,N-1] \tag{2}$$

An example of the template for calculating the $L_{blockiness\text{-}h}$ is shown in Figure 2, where two adjacent blocks of 8×8 pixels (i.e. A and B) are extracted from a real JPEG image. The local blockiness along the vertical direction $L_{blockiness\text{-}v}$ can be similarly calculated. The higher the values of $L_{blockiness\text{-}h}$ and $L_{blockiness\text{-}v}$ the larger the distortion of the blocking artifact is. It should, however, be noted that this does not necessarily mean that the blocking artifact is also more visible. The local visibility of a blocking artifact may be affected by texture and luminance masking, which typically occur in the HVS. It has been shown in literature that taking into account these masking effects can be greatly beneficial for the prediction performance of a NR blocking metric [7]. However, modeling the HVS introduces more computational power. So, in this paper we avoid the calculation of masking, and rely on the neural network to learn the unknown functional relationship between the extracted gradient-based features and the rating of overall image quality.

2.2 Local feature extraction: JPEG2000

In JPEG2000 compression, blur artifacts are perceptually prominent along edges or in textured areas. Measuring the smoothing or smearing effect on strong edges has been proved to be an effective approach to approximate the overall perceived quality of JPEG2000 compressed images [10]. In this paper, the local feature extraction for the JPEG2000 compressed images is built upon calculating the degree of blur at an edge within a local area of image content. To detect strong edges, and consequently to identify the spatial location of blur artifacts, a variety of techniques has been proposed in literature (e.g. PEM in [5]). However, to maintain a low complexity of the system, just a Sobel edge detector is adopted here. The location of the strong edges is then extracted by applying a threshold to the resulting gradient image (i.e. by removing noise and insignificant edges).

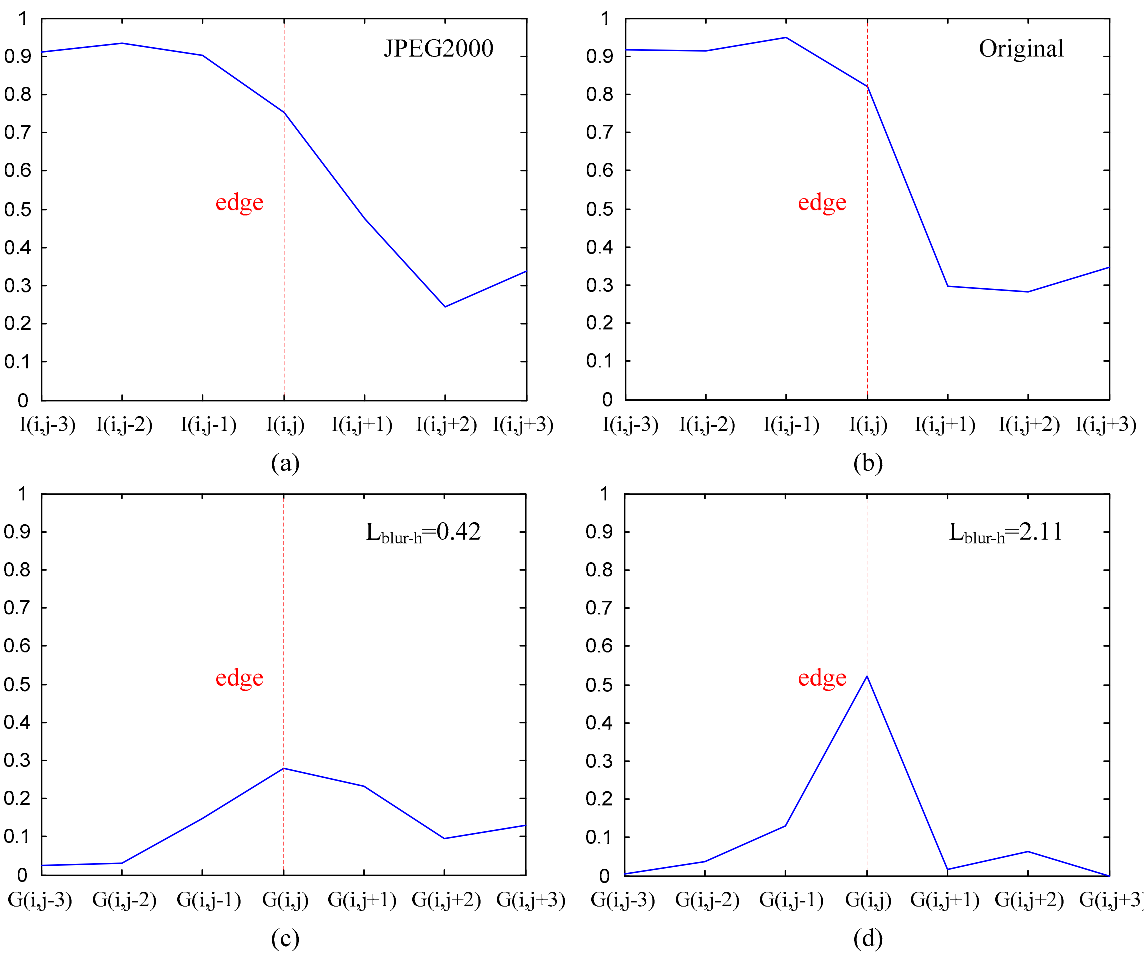

Figure 3. Illustration of the calculation of the local blur: (a) the intensity profile over a detected edge (i.e. at location (113, 259) in Figure 1 (b)), (b) the intensity profile over the corresponding pixels of (a) in the original image of Figure 1 (b), (c) the gradient profile of (a), and (d) the gradient profile of (b).

For all pixels located along the detected edges, local blur is defined as the sharpness of the edge in the gradient domain instead of calculating the distance between the start and end position of an edge (i.e. the edge spread along the horizontal/vertical direction) as proposed in [10]. When describing blur simply as the relative gradient energy of an edge compared to its direct vicinity, it can be quantified in the same manner as used in (1), i.e.:

$$L_{blur\text{-}h}(i,j) = \frac{G_h(i,j)}{\frac{1}{2n}\sum_{x=-n,\ldots,n,x\neq 0} G_h(i,j+x)} \qquad (i,j) \in \{\text{location of blur}\} \tag{3}$$

where $L_{blur\text{-}h}$ indicates the local blur along the horizontal direction. $L_{blur\text{-}v}$, i.e. the local blur in the vertical direction, can be calculated similarly. The lower the value of $L_{blur\text{-}h}$ and $L_{blur\text{-}v}$, the larger the distortion of the blur artifact is. Figure 3 explains the reasoning behind the proposed approach of using gradient energy to detect blur. Figure 3(a) shows the intensity profile over a detected edge (i.e. at location (113, 259) of the image in Figure 1(b)), while Figure 3(b) shows the corresponding pixel intensity values in the original image of Figure 1(b). The difference in sharpness between the two edges is clearly revealed in the gradient domain (see Figure 3(c) and (d)). In correspondence, the values of $L_{blur\text{-}h}$ indicate that the edge of Figure 3(a) is more blurred than the edge of Figure 3(b).

2.3 Global descriptor of the image features

To gather the local distortion information in a more compact format, in order to build an informative global descriptor to feed the neural network, a refined pooling strategy of the one proposed in [8] is adopted. It results in a vector that represents each image. This vector includes a statistical representation of the distortion distribution in the image. Having computed the feature values m_i (i=1, ..., N_M) per image (i.e. $L_{blockiness}$ calculated on the blocking grid or L_{blur} calculated on the detected edges), these values are sorted in ascending order of magnitude. The envelope of the obtained distribution is then expressed in the global distortion descriptor **m** by taking 11 of its percentiles φ:

$$\mathbf{m} = \{\varphi_\alpha ; \alpha \in \{0,10,20,30,40,50,60,70,80,90,100\}\}; \quad \varphi_\alpha = \left\lfloor \frac{N_M}{100}\alpha + \frac{1}{2} \right\rfloor \tag{4}$$

Compared to simply taking the average of the feature values, this spatial pooling strategy allows feeding the non-linear regression with a more complete overview of the amount and behavior of the considered distortion in the image.

3. NR IMAGE QUALITY ESTIMATOR

The image quality assessment community is accustomed to make use of non-linear mapping strategies to improve the correlation between objective metric predictions and human quality assessments [1]. Several approaches can be used to this end, ranging from logistic fitting [1] to machine learning (ML) [8], [9] and [14]. In the case of perception-related problems, machine learning methods are powerful tools. Even though they don't require any a-priori information, ML methods still allow inferring highly non-linear relationships between numerical descriptions of images and their subjective assessment, where the mapping is learnt from examples (i.e. real-world observations). Moreover, when the computational constraints are tight, such as in real-time applications, the accuracy in the mapping granted by ML methods allows designing simple metrics with low computational requirements.

In the system proposed here, a feed-forward neural network maps the blockiness and blur features into the associated estimates of perceived quality. Neural networks learn the non-linear dependency of target values $\hat{t}$ from inputs $\mathbf{x}$ as a series expansion of n_h basis functions a_h, often sigmoids [15].

$$\hat{t} = y(\mathbf{x}) = \sum_{h=1}^{n} w_h a_h(\mathbf{x}) \tag{5}$$

The MultiLayer Perceptron (MLP) paradigm [15] belongs to this class of networks, and has been proved to perform effectively in problems where the target-mapping function can be retrieved by a few computing units endowed with global scope. MLPs aim at implementing a stimulus-response behavior by arranging such computing units ("neurons") into a layered network: each unit involves a non-linear transformation of weighted inputs; each layer outputs are further non-linearly transformed by the neurons in the next layer.

The “Circular Back Propagation” (CBP) network [16] is an extension of the MLP paradigm, based on the addition of one more input value, being the sum of the squared values of all the network inputs. It has been proven that this addition allows the network to switch between the classic, sigmoidal behavior, and a smoother, bell-shaped radial function [16]. The modeling process is entirely data-driven, and fully defined in the training phase. Thanks to this adaptive behavior, CBP networks are quite appropriate for perception modeling, where often the underlying cognitive mechanisms are not clear.

The CBP architecture can be formally described as follows. Given an input vector $\mathbf{x}=\{x_1, \ldots, x_{ni}\}$ for a given stimulus and its corresponding target value t of dimensionality n_o, the input layer is structured in n_i neurons, each connected to every neuron in the following “hidden” layer. The j-th hidden neuron performs a non-linear transformation of a weighted combination of the input values, with coefficients $w_{j,i}$ (j=1, …, n_h; i=1, …, n_i):

$$a_j = sigm\left(w_{j,0} + \sum_{i=1}^{n_i} w_{j,i} x_i + w_{j,n_i+1} \sum_{i=1}^{n_i} x_i^{\,2} \right) \tag{6}$$

where $sigm(z)=(1+e^{-z})^{-1}$, and a_j is the neuron activation (i.e. the output of the basis function). The output layer provides the actual network predictions, y_k, (k = 1, …, n_o), in a similar way:

$$y_k = sigm\left(w_{k,0} + \sum_{j=1}^{n_h} w_{k,j} a_j \right) \tag{7}$$

In the case of quality assessment, n_o=1, and the target value t is just the quality score represented by a scalar value.

The structural CBP modification does not interfere with the possibility of using the well-known back propagation algorithm [15] for model training. The cost to be minimized is expressed through a quadratic function comparing the target and the predicted value:

$$E = \frac{1}{n_o n_p} \sum_{l=1}^{n_p} \sum_{k=1}^{n_o} \left(t_k^{(l)} - y_k^{(l)} \right)^2 \tag{8}$$

where n_p is the number of training patterns. Together with the weighting coefficients $w_{j,i}$ and $w_{k,j}$, the other degree of freedom of the neural network that has to be set is the number of hidden units, n_h. For this task, the present research followed an empirical approach [17] mainly because of its simplicity and proved effectiveness. The whole parameter setting process is completed in the training phase, and hence does not entail computational overhead at runtime.

4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed approach, the LIVE image quality assessment database [18] was used. It consists of a set of twenty nine high-resolution and high-quality color source images that reflect adequate diversity in image content. These images were compressed using JPEG at a bit rate ranging from 0.15 bits per pixel (bpp) to 3.34 bpp, resulting in a database of 233 JPEG compressed stimuli (including the originals). The same source images were also compressed using JPEG2000 at a bit rate ranging from 0.028 bpp to 3.15 bpp, yielding a database of 227 JPEG2000 compressed stimuli (including the originals). An extensive psychovisual experiment was conducted to assign a difference mean opinion score (DMOS) to each stimulus. The DMOS was measured on a continuous linear scale that was divided into five intervals marked with the adjectives “Bad”, “Poor”, “Fair”, “Good” and “Excellent”.

In our performance evaluation, the source images were divided into two groups, i.e. 20 out of 29 source images were used for training and the remaining 9 images were used for testing. This resulted in 161 stimuli for training and 72 stimuli for testing in the JPEG database for evaluating our JPEG metric, and 156 stimuli for training and 71 stimuli for testing in the JPEG2000 database for evaluating our JPEG2000 metric. For each stimulus, a vector containing eleven percentiles of the distribution of the local blockiness/blur features was taken as the input to the neural network, which was equipped with 3 hidden neurons. Figure 4 shows the scatter plots of the DMOS versus the quality prediction based on our neural network approach, for the JPEG and JPEG2000 database of test images.

The performance of an objective metric can be evaluated with respect to its ability to predict subjective quality ratings (the DMOS). Two values can be employed to characterize this ability: i.e. the Pearson linear correlation coefficient, and the root mean square error (RMSE) [19]. Based on these two values the JPEG/JPEG2000 metrics proposed in this paper are compared to state-of-the-art NR metrics, including four metrics for JPEG [6]-[9] and two metrics for JPEG2000 [10] and [11]. Table 1 and 2 list the Pearson correlation coefficient and the RMSE for our metrics as well as for the state-of-the-art metrics. It should be noted that we had no access to the test environment of the metrics in [6]-[11], and the comparison among metrics might be biased due to e.g. a different selection of the disjoint sets for training and testing. However, the metrics mentioned above were all evaluated with the LIVE database; hence the performance comparison is considered fairly valuable. For the NR metrics for JPEG compression, our proposed metric outperforms the metrics of [6] and [7] in predicting the overall perceived quality. Its performance is comparable to that of the metrics of [8] and [9], which are also based on a neural network approach. The advantage of our metric lies in its simplicity in the extracted features, compared to the metrics of [8] and [9]. For the NR metrics for JPEG2000 compression, our proposed metric clearly outperforms the metric in [10] mainly due to the powerful neural network, which can efficiently approximate the functional relationship between the extracted local blur features and the rating of overall quality. The performance of our proposed JPEG2000 metric is comparable to the metric in [11], but it is indeed a simple yet efficient alternative useful for real-time implementation.

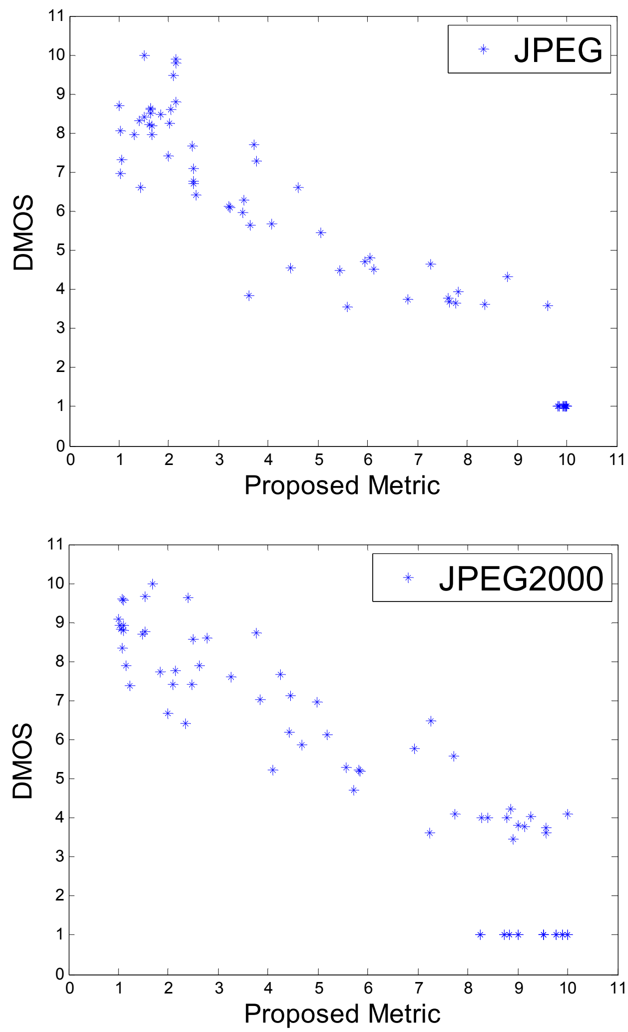

Figure 4: Scatter plots of DMOS versus the proposed metric (based on the neural network approach) for JPEG and JPEG2000.

Table 1. Performance comparison of our approach with state-of-the-art metrics for JPEG.

NR Metric (JPEG)	Pearson Correlation Coefficient	Root Mean Square Error (RMSE)
Proposed	0.952	0.234 (scale [-1, 1])
Wang et al [6]	0.931	N/A
Liu et al [7]	0.918	N/A
Gastaldo et al [8]	0.943	0.153 (scale [-1,1])
Babu et al [9]	N/A	0.570 (scale [1, 10]

Table 2. Performance comparison of our approach with state-of-the-art metrics for JPEG2000.

NR Metric (JPEG2000)	Pearson Correlation Coefficient	Root Mean Square Error (RMSE)
Proposed	0.92	0.289 (scale [-1, 1])
Marziliano et al [10]	0.85	N/A
Sheikh et al [11]	0.93	N/A

5. CONCLUSIONS

In this paper, we present a novel approach to assess the overall perceived quality of JPEG/JPEG2000 compressed images, without the access to their original version. The approach extracts the most relevant features that are representative for the quality degradation as a consequence of compression, i.e. the local blockiness in the JPEG compressed images, and the local blur in the JPEG2000 compressed images. The highly non-linear relationship between the extracted features and the quality rating is empirically learned by a CBP neural network. This approach intrinsically takes advantage of the prior knowledge on the specific characteristics of JPEG/JPEG2000 compression artifacts, thus optimizing the metric by efficiently calculating the image features. The performance of the proposed NR JPEG/JPEG2000 metrics is evaluated with the use of the LIVE database, and compared to several leading alternatives in literature. Experimental results show that our metrics result in a strong correlation with subjective data at a reduced computational load. As such, the proposed approach is promising in terms of both computational efficiency and practical reliability for real-time applications.

REFERENCES

[1] Wang, Z. and Bovik, A. C., [Modern Image Quality Assessment], Synthesis Lectures on Image, Video & Multimedia Processing, Morgan & Claypool Publishers (2006).

[2] Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P., "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing, vol.13, no.4, pp. 600- 612 (2004).

[3] Sheikh, H. R. and Bovik, A. C., "Image information and visual quality," IEEE Transactions on Image Processing, vol.15, no.2, pp. 430- 444 (2006).

[4] Ferzli, R. and Karam, L. J., “A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB),” IEEE Transactions on Image Processing, vol. 18, pp. 717–728 (2009).

[5] Liu, H., Klomp, N. and Heynderickx, I., "A No-Reference Metric for Perceived Ringing," in Proc. Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (2009).

[6] Wang, Z., Sheikh, H. R. and Bovik, A. C., "No-reference perceptual quality assessment of JPEG compressed images", Proc. IEEE International Conference on Image Processing (2002).

[7] Liu, H. and Heynderickx, I., “A Perceptually Relevant No-Reference Blockiness Metric Based on Local Image Characteristics,” EURASIP Journal on Advances in Signal Processing, vol. 2009 (2009).
[8] Gastaldo, P. and Zunino, R., “Neural networks for the no-reference assessment of perceived quality,” Journal of. Electronic Imaging, 14 (3), 033004 (2005).
[9] Babu, R. V., Suresh, S. and Perkis, A., “No-reference JPEG-image quality assessment using GAP-RBF,” Signal Processing, vol. 87, no.6, pp.1493-1503 (2007).
[10] Marziliano, P., Dufax, F., Winkler, S. and Ebrahimi, T., “Perceptual blur and ringing metrics: Application to JPEG2000,” Signal Processing: Image Communication, vol. 19, pp. 163-172 (2004).
[11] Sheikh, H. R., Bovik, A. C. and Cormack, L. K., "No-Reference Quality Assessment Using Natural Scene Statistics: JPEG2000," IEEE Transactions on Image Processing, vol. 14, no. 12 (2005).
[12] Zon, K. and Ali, W., “Automated video chain optimization,” IEEE Transactions on Consumer Electronics, vol. 47, pp. 593-603 (2001).
[13] Yuen, M. and Wu, H. R., “A survey of hybrid MC/DPCM/DCT video coding distortions,” Signal Processing, vol. 70, no. 3, pp. 247-278 (1998).
[14] Redi, J., Gastaldo, P., Zunino, R. and Heynderickx, I., “Reduced reference assessment of perceived quality by exploiting color information”, Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (2009).
[15] Rumelhart, D. E. and McClelland, J. L., [Parallel distributed processing], MIT Press, Cambridge, MA (1986).
[16] Ridella, S., Rovetta, S. and Zunino, R., “Circular back-propagation networks for classification,” IEEE Trans. on Neural Networks, vol. 8, pp. 84-97 (1997).
[17] Widrow, B. and Lehr, M. A., “30 Years of Adaptive Neural Networks: Perceptron, Madaline and Back Propagation,” Proc. IEEE. 78(9), 1415-42 (1990).
[18] Sheikh, H. R., Wang, Z., Cormack, L. and Bovik, A. C., "LIVE Image Quality Assessment Database Release 2," http://live.ece.utexas.edu/research/quality
[19] VQEG (2003, Aug.): Final report from the video quality experts group on the validation of objective models of video quality assessment. Available: http://www.vqeg.org