No-reference Perceptual Quality Assessment of JPEG Images Using General Regression Neural Network

Yanwei Yu, Zhengding Lu, Hefei Ling, and Fuhao Zou

College of Computer Science and Technology,
Huazhong University of Science & Technology, Wuhan 430074, China
yuyanwei198112@163.com, zdlu4409@public.wh.hb.cn, lhefei@163.com,
fuhao_zou@yahoo.com.cn

Abstract. No-reference perceptual quality assessment for JPEG images in real time is a critical requirement for some applications, such as in-service visual quality monitoring, where original information can not be available. This paper proposes a no-reference perceptual quality-assessment method based on a general regression neural network (GRNN). The three visual features of artifacts introduced in JPEG images are formulated block by block individually so that our method is computational and memory-efficient. The GRNN is used to realize the mapping of these visual features into a quality score in real time because of its excellent approximation and very short training time (one-pass learning). Experimental results on an on-line database show that our estimated scores have an excellent correlation with subjective MOS scores.

1 Introduction

Perceptual quality assessment of digital image and video is essential because they are subject to various distortions during acquisition, processing, or reproduction. Often, the subjective evaluation, for example, the mean opinion score (MOS), is the most reliable way of assessing the quality of an digital image or video, but the subjective evaluation method is not suitable for its inconvenience, slowness, high cost for most applications. Thus, in the past three to four decades, researchers turn to developing objective quality metrics. According to the availability of the original data, objective quality metrics are generally divided into three classes: full-reference (FR) where the reference signal is fully available, reduced-reference (RR) which requires certain features of the original signal, and no-reference (NR) in which a reference image or video is not available [1]. Most of current objective quality metrics belong to full-reference quality metrics, most of which try to simulate the characteristics of human visual error perception [1-5]. But in some applications, such as in-service visual quality monitoring, the reference data is not available. So it is urgent to develop a no-reference objective quality metric.

The present existing NR quality metrics mainly assessed the perceived quality by using the statistical features of distorted images [6-8]. Wang et al. [6] proposed a no-referenced perceptual quality assessment of JPEG compressed images. The metrics in [6] are computed locally, thus they are computational and memory efficient. But their fatal shortcoming is that the model parameters are derived from the training data, so the generalization is poor. An objective quality-assessment method based on a Circular Back-Propagation (CBP) neural structure was proposed in [7]. Their method is unique that they aim at reproducing perceived image quality, rather than at defining a comprehensive model of the human visual system, but the method requires more storage space because the numerical features as CBPNN inputs derived from co-occurrence matrixes can not be computed locally.

In this paper, we propose a novel method (namely GRNN_NRQA for short) using general regression neural network (GRNN) for no-reference perceptual quality assessment of JPEG images. We quantify blocking artifacts and blurring artifacts as the most significant artifacts in JPEG images according to the visual characteristic of the distortion, such as loss of image details and edge smoothness. Then we emulate the human "understanding" mechanism to predict perceptual quality scores from numerical features by using GRNN because of its excellent approximation and very short training time (one-pass learning). The effectiveness of the proposed method has been verified by the experiments on the LIVE image quality assessment database [9]. Experimental results show that the proposed method outperforms the methods in [6-7] in terms of prediction accuracy, monotonicity and consistency in assessing the perceived visual quality of image and video sequence.

2 The Proposed Perceptual Quality Metric

The proposed perceptual quality evaluation scheme consists of two parts mainly. The first part describes blocking artifacts and blurring artifacts in JPEG distorted image using the mathematical formula. The second part estimates the JPEG distorted image quality from the numeric features extracted during the former phase by using GRNN. Fig. 1 gives the block diagram of the GRNN_NRQA scheme.

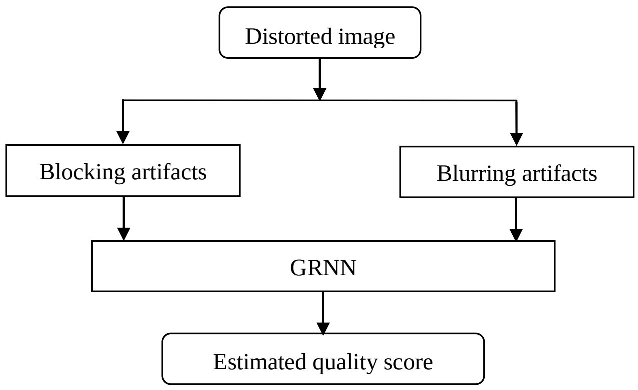
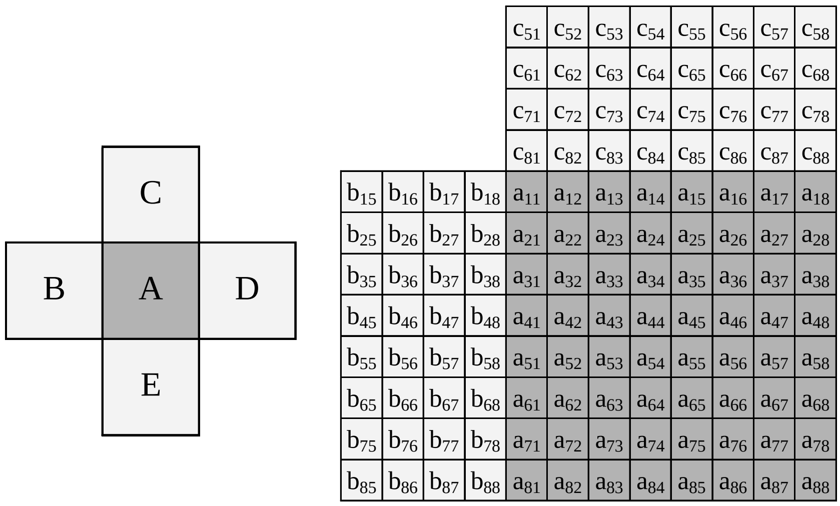


Fig. 1. Block diagram of the GRNN_NRQA scheme

2.1 The Formulation of Artifacts in JPEG Images

JPEG is a lossy image coding technique based on the block-based discrete cosine transform (BDCT) [10]. During the quantization of BDCT coefficients, blurring artifacts and blocking artifacts are introduced. The blocking artifact manifests an artificial discontinuity between adjacent blocks resulting from the independent quantization of the BDCT coefficients. Blurring artifacts are due to the loss of DCT high-frequent coefficients.

Because each macro block in the image may have different quantization steps and the same extent of blockiness may result in different degree of disturbance in human eyes due to the masking effect, we detect blocking artifacts locally block by block individually in this paper. Fig. 2 shows an example of 8×8 block and its neighboring blocks. It is enough for measuring blocking artifacts of Block A to consider its relationship with the neighboring blocks, i.e., Block B, C, D and E. However, at the time of processing Block A, only Block B and C are considered. The other two borders of Block A are to be considered when measuring horizontal blocking artifacts for Block D and vertical blocking artifacts for Block E. Note that we only process the 8×8 region near the block boundary to reduce computational complexity when measuring the artifacts at the block boundary because the artifacts at the block boundary is mainly affected by its neighboring region.



(a) (b)

Fig. 2. An example of 8×8 block and its neighboring blocks

2.1.1 blocking artifacts measure

Because the blocking artifact manifests an artificial discontinuity between adjacent blocks, it is reasonable to quantify the blocking artifacts of Block A by just examining the inter-pixel difference between Block A and its neighbors Block B and C.

By take into account the different masking effect of each block, the inter-pixel difference between Block A and its neighbors Block is divided by the sum of inter-pixel difference in the 8×8 region near the block boundary to denote the blocking artifact measure.

We define the horizontal inter-block difference $F1\_h$ between Block A and B as

$$F1\_h = \sum_{i=1}^{8} F1\_h(i) \qquad (1)$$

$$F1\_h(i) = \begin{cases} \dfrac{|a_{i1} - b_{i8}|}{\sum_{j=5}^{7} |b_{i(j+1)} - b_{ij}| + \sum_{j=1}^{3} |a_{i(j+1)} - a_{ij}| + |a_{i1} - b_{i8}|} & a_{i1} \neq b_{i8} \\ 0 & a_{i1} = b_{i8} \end{cases} \qquad (2)$$

Where $F1\_h(i)$ denotes the horizontal inter-block difference at the i^{th} row between Block A and B.

The vertical inter-block difference $F1\_v$ between Block A and Block C are defined in the similar way. Assuming that the sensitivity of HVS to horizontal and vertical blocking artifacts are similar, then the blocking artifacts of an 8×8 block $F1_{blk}$ can be summarized as,

$$F1_{blk} = \frac{F1\_h + F1\_v}{2} \qquad (3)$$

2.1.2 blurring artifacts measure

The blurring artifacts manifest the loss of details in the block and "large splash" in the less coarse texture at the low bit-rate because of the unobviousness of the border between the neighboring blocks. So blurring artifacts can be measured by two aspects:

1) intra-block blurring artifacts measurement

The burring artifacts measure in the block can be measured by the descendant contrast between neighboring pixels in the block.

We define the horizontal contrast between neighboring pixels $F2\_h$ in the block A as,

$$F2\_h = \frac{1}{56} \sum_{i=1}^{8} \left(\sum_{j=1}^{7} |a_{i(j+1)} - a_{ij}| \right) \qquad (4)$$

The vertical contrast between neighboring pixels $F2\_v$ in the block A is defined in the similar way. Then the blurring artifacts in an 8×8 block $F2_{blk}$ can be represented as,

$$F2_{blk} = \frac{F2\_h + F2\_v}{2} \tag{5}$$

2) inter-block blurring artifacts measurement

The burring artifacts at the border between the neighboring blocks can be measured in terms of the proportion of zero crossings from pixel to pixel locally. That means only the 8x8 region near the block boundary is considered, ranging from 4 pixels to the left of the boundary to 4 pixels to the right of the boundary. The total number of zero-crossings is obtained divided by the total number of crossings to give the flatness measure in that particular region.

Defining a function for zero crossing as

$$ZC(x, y) = \begin{cases} 1, & |x - y| = 0 \\ 0, & else \end{cases} \tag{6}$$

Then

$$F3\_h = \frac{1}{56}\sum_{i=1}^{8}\left(\sum_{j=1}^{7}\left|b_{i(j+1)} - b_{i,j}\right| + \sum_{j=1}^{3}\left|a_{i(j+1)} - a_{i,j}\right|\right) + \frac{1}{56}\left|a_{i1} - b_{i8}\right| \tag{7}$$

Where $F3\_h$ is the horizontal inter-block blurring artifacts measure across block A and block B.

The vertical inter-block blurring artifacts measure across block A and block C can be attained similarly. Then the inter-block blurring artifacts $F3_{blk}$ near the regions of block A can be quantified as

$$F3_{blk} = \frac{F3\_h + F3\_v}{2} \tag{8}$$

Since all the blocks are subject to similar distortions and all the above artifacts for each single block have considered the masking effects , the overall measure of artifacts for an image block is attained by averaging the corresponding artifact of all blocks.

$$F1 = average(F1_{blk}), \qquad F2 = average(F2_{blk}), \qquad F3 = average(F3_{blk}) \tag{9}$$

2.2 Perceptual Quality Assessment Using GRNN

The GRNN is then used to emulate the understanding system of human beings and predict the perceptual quality scores from the aforementioned visual features. The GRNN is a memory based neural network based on the estimation of a probability density function. The GRNN have the advantages of very short training time (one-pass learning) and guaranteed performance even with sparse data and does not need any priori information about the form of the regression functions in comparison to conventional regression techniques [11-12]. In mathematical terms, if we have a vector random variable x, a scalar random variable y, let X be a particular measured value of x, then the conditional mean of y given X can be represented as:

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y_i \exp(-\frac{D_i^{2}}{2\sigma^{2}})}{\sum_{i=1}^{n} \exp(-\frac{D_i^{2}}{2\sigma^{2}})} \tag{10}$$

Where $D_i^{2} = (X - X_i)^T (X - X_i)$. X_i and Y_i are the sample values of the random x and y. n denotes the number of samples. σ is the width of the kernel which is the only unknown parameter in the above equation (10). After repeating experiments, we found the estimating performance is the best when σ is 0.018.

3 Experimental Results

The system for image quality assessment has been tested by using a database of JPEG images available at the LIVE Quality Assessment Database [9]. The database of JPEG images included 233 test images, 175 of which were generated by compressing (with JPEG) 29 high-resolution 24-bits/pixel RGB color images (typically 768 x 512) with different compression ratios. The Mean Opinion Score (MOS) of each image in the LIVE database is provided on a linear scale ranging from 1 (“bad”) to 100 (“excellent”). In the present experiment, the subjective scores have been converted into the interval [-1,1] linearly.

The effectiveness of the GRNN_QA system is measured by using a conventional cross-validation approach. The available samples are randomly divided into a training set including 100 images and a test set including 133 images. During the training process of quality-assessment system only the training set is used, whereas the test set is applied only to predict the system generalization ability.

The effectiveness of objective quality metrics is generally reflected by the correlation with subjective quality scores. The higher the correlation is, the more effective the objective quality metric is. In this paper, the correlation is quantified by the Pearson Correlation and the Root Mean Squared error (RMS).

The scatter plot in Fig. 4 shows the results obtained for the test set, with the estimated objective quality scores Q_GRNN as the x axis and the MOS scores as the y axis. It is obvious that the estimated scores correlate well with MOS scores and our proposed method is effective.

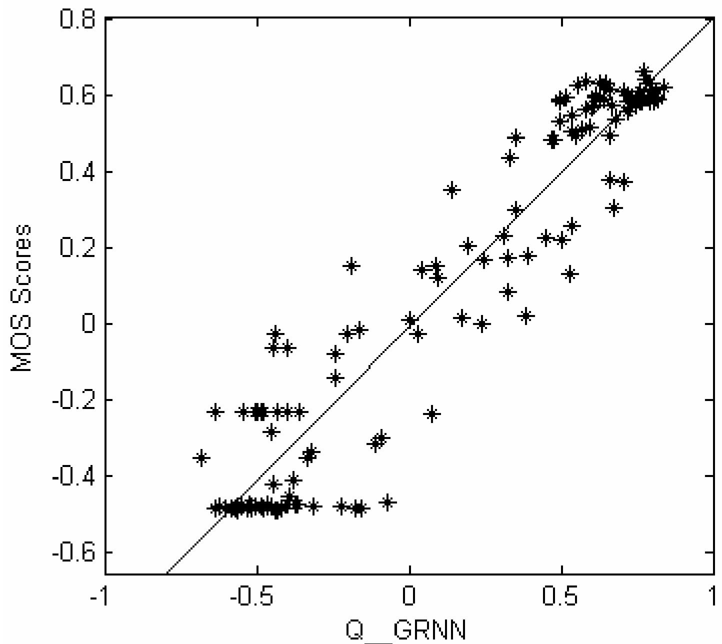


Fig. 3. Estimated perceptual quality scores Q_GRNN versus MOS Scores

Table 1 shows the Pearson Correlation and the Root Mean Squared error (RMS) between various quality metrics and the subjective ratings of the JEPG database provided by LIVE [9]. It can be seen that our objective scores Q_GRNN have an excellent correlation with the subjective MOS scores and our method outperform the methods in [6-7].

Table 1. Pearson Correlation and RMS for LIVE database

algorithm	Pearson Correlation	RMS
Q_GRNN	0.954983	0.14053
S [6]	0.918092	0.7256
Q_CBP [7]	0.950152	0.14

4 Conclusions

In this paper, we proposed a no-reference perceptual quality assessment system GRNN_NRQA for JPEG images. Visual features effectively capturing the artifacts introduced by JPEG were formulated block by block individually and processed by GRNN which output the corresponding quality estimate. Experimental results confirmed that the system provides a satisfactory approximation of the subjective MOS scores.

The method is computationally efficient since no complicated transform is computed and the algorithm can be implemented without storing the entire image in memory, which makes embedded implementations easier. The basic methodology of the proposed method can also been used to develop NR quality assessment methods for MPEG compressed video. The next work is to apply our proposed method to our project to estimate the quality of watermarked video in real time.

Acknowledgement

This work is supported by the National Natural Science Foundation of China under Grant No. 60502024, the Natural Science Foundation of Hubei Province under Grant No. 2005ABA267, the Electronic Development Fund of Ministry of Information Industry of China and the Innovation Fund for Technology Based Firms of Ministry of Science and Technology of China under Grant No. 04C26214201284.

References

1. Z. Wang, H. R. Sheikh and A C. Bovik: Objective Video Quality Assessment. The Handbook of Video Databases: Design and Applications (B. Furht and O. Marqure, eds.), CRC Press, (2003) 1041-1078
2. M. P. Eckert and A. P. Bradley: Perceptual Quality Metrics Applied to Still Image Compression. Signal Processing, 70 (1998) 177-200
3. VQEG: Final Report from The Video Quality Experts Group on The Validation of Objective Models of Video Quality Assessment. http://www.vqeg.org/, 2000
4. T. N. Pappas and R. J. Safranek: Perceptual Criteria for Image Quality Evaluation. Handbook of Image and Video Processing (A. Bovik, ed.), Academic Press, 2000
5. Z. Wang and A. C. Bovik: Why Is Image Quality Assessment So Difficult? in Proc. IEEE Int. Conf.Acoust., Speech, and Signal Processing, 2002
6. Z. Wang, H. R. Sheikh and A. C. Bovik: No-Reference Perceptual Quality Assessment of JPEG Compressed Images. IEEE International Conference on Image Processing, (2002) I-477-480
7. Gastaldo, P., Zunino, R.: No-reference Quality Assessment of JPEG Images by Using CBP Neural Networks. Proc. of the 2004 International Symposium on Circuits and Systems (ISCAS'04), 5 (2004) V-772 - V-775
8. Pan, F., Lin, X., Rahardja, S., Lin, W., Ong, E., Yao, S., Lu, Z, Yang, X.: A Locally-adaptive Algorithm for Measuring Blocking Artifacts in Images and Videos. Proc. of the 2004 International Symposium on Circuits and Systems (ISCAS'04), 3 (2004) 23-26
9. H. R. Sheikh, Z. Wang, A. C. Bovik, and L. K. Cormack: LIVE Image Quality Assessment Database. http://live.ece.utexas.edu/research/quality/
10. Y. Q. Shi, H. Sun: Image and Video Compression for Multimedia Engineering – Fundamentals, Algorithms and Standards. CRC Press, 1999
11. D. F. Spetch: A General Regression Neural Network. IEEE Transactions on Neural Netwolks, 2(1991) 568-576
12. Parzen E: On Estimation of a Probability Density Function and Mode. Ann. Math Statistics, 33(1962) 1065-1076