

# NO-REFERENCE QUALITY ASSESSMENT FOR JPEG2000 COMPRESSED IMAGES\*

Hanghang Tong<sup>1</sup>, Mingjing Li<sup>2</sup>, Hong-Jiang Zhang<sup>2</sup>, Changshui Zhang<sup>3</sup>

<sup>1,3</sup>Department of Automation, Tsinghua University, Beijing 100084, China

<sup>2</sup>Microsoft Research Asia, 49 Zhichun Road, Beijing 100080, China

<sup>1</sup>walkstar98@mails.tsinghua.edu.cn, <sup>2</sup>{mjli, hjzhang}@microsoft.com, <sup>3</sup>zcs@tsinghua.edu.cn

## ABSTRACT

No-Reference quality assessment is a relatively new topic and has been attracting more and more attention in recent years. Due to the limited understanding of the human vision system, most of the existing methods focus on measuring to what extent the image has been distorted. In this paper, by viewing all edge points in JPEG2000 compressed images as ‘distorted’ or ‘un-distorted’, we propose using principal component analysis (PCA) to extract the local feature of a given edge point, which indicates both blurring and ringing. We also propose using the probabilities of the given edge point being ‘distorted’ and ‘un-distorted’ to model the local distortion metric, which is straightforward and can be easily applied to any type of local feature. Experimental results demonstrate the effectiveness of our scheme.

## 1. INTRODUCTION

Image quality assessment plays an important role in monitoring image quality, benchmarking image processing systems and optimizing algorithms [2, 9, 10, 11]. Most existing methods are essentially fidelity assessment since they need the original un-distorted image as a reference [9]. On the other hand, it is very easy for human observers to assess image quality without using any reference image [3, 9, 11]. What is more, in many situations, the original un-distorted image might not exist [3, 9, 10]. In recent years, No-Reference (NR) image quality assessment has attracted the attention of more and more researchers [5, 6, 7, 9, 10].

Due to the limited understanding of the human vision system (HVS), most, if not all, of the existing NR assessment algorithms focus on measuring to what extent the image has probably been distorted. Whether implicitly or explicitly, the general flow of these algorithms can be summarized as follows [5, 6, 7, 9, 10]: 1) find some local feature  $X(i, j)$  which exhibits different

characteristics for the original images and the distorted versions, where  $(i, j)$  denotes the position of the pixel; 2) use  $X(i, j)$  to assess local distortion  $Ld(i, j)$ ; 3) average  $Ld(i, j)$  over the whole image to get a overall distortion metric  $Dm$ ; 4) use  $Dm$  to predict image quality score  $Ps$  which is consistent with human perception. Finding suitable local feature  $X(i, j)$  and modeling the local distortion metric  $Ld(i, j)$  are two key steps within the whole algorithm.

For JPEG2000 compressed images, the main distortion that might be introduced is blurring and ringing. There exist two major types of NR assessment algorithms for JPEG2000 compressed images: in [6, 7], the authors proposed using edge spread as the local blurring feature  $X(i, j)$  which is used directly as the local distortion metric  $Ld(i, j)$ ; in [9], the authors proposed using wavelet coefficients as the local feature  $X(i, j)$ , and the local distortion metric  $Ld(i, j)$  is simply denoted as ‘significant’ or ‘insignificant’ by a threshold. The drawbacks of these algorithms are: 1) edge spread only indicates the blurring distortion and calculating wavelet coefficients can be very time consuming; 2) both types of algorithms seem to over simply the relationship between the local feature  $X(i, j)$  and the local distortion metric  $Ld(i, j)$ .

To address the drawbacks of the existing local feature, we propose using PCA to extract a new type of local feature by viewing all edge points  $(i, j)$  of images as two classes, i.e. ‘distorted’ and ‘un-distorted’. The new feature is learned from the images and thus might be able to indicate both blurring and ringing. Also, extracting such feature is fast since it does not need to calculate wavelet coefficients. We also propose a new model to describe the relationship between the local feature  $X(i, j)$  and the local distortion metric  $Ld(i, j)$ . In our model, the local distortion metric  $Ld(i, j)$  is calculated using the probabilities of  $(i, j)$  being ‘distorted’ and ‘un-distorted’. Compared with the existing methods, it can better

---

\* This work was performed at Microsoft Research Asia

describe the relationship between the local feature  $X(i, j)$  and the local distortion metric  $Ld(i, j)$ . The proposed new model is straightforward and can be easily applied to any type of local feature.

The rest of this paper is organized as follows: in Section 2, we present our scheme in detail; experimental results are given in Section 3; finally, we conclude the paper in Section 4.

## 2. PROPOSED NR QUALITY ASSEMENT SCHEME

As mentioned in the introduction section, the main distortion that might be introduced in JPEG2000 compressed images is blurring and ringing. Blurring is perceptually apparent around edges while ringing usually appears near sharp edges [1, 5].

Inspired by the above observation, we initially use the pixels in the neighborhood of a given edge point  $(i, j)$  to indicate both its blurring and ringing. Furthermore, by viewing all edge points as two classes ('distorted' and 'un-distorted'), we can use PCA to make the effect of distortion explicit. Also, viewing all edge points as two classes enables us to model the relationship between the local feature  $X(i, j)$  and the local distortion metric  $Ld(i, j)$  through the probabilities of the edge point  $(i, j)$  being 'distorted' and 'un-distorted'.

In this section, we first propose the method of extracting local feature  $X(i, j)$  in subsection 2.1; then the new model for the local distortion metric  $Ld(i, j)$  will be given in subsection 2.2; finally, subsection 2.3 will give the details of the proposed NR quality assessment algorithm for JPEG2000 compressed images.

### 2.1. The local feature $X(i, j)$

Since the information about blurring and ringing of a given edge point is contained in its neighborhood, we perform the following operation as a preprocessing step:

**Operation 1:** for an edge point  $(i, j)$ , assign it to the center of a window and arrange all the pixels within this window in a vector  $X_0(i, j)$ . Let  $r$  denote the radius of the window. Vector  $X_0(i, j)$  contains the information about to what extent  $(i, j)$  is distorted.

However, we can not determine the distortion effect directly from the vector  $X_0(i, j)$ . To address this issue, firstly, all edge points are viewed as 'distorted' or 'un-distorted'; then we resort to PCA to find some principle axes so that the 'distorted' and 'un-distorted' edge points will be separated as far as possible in the subspace spanned by these axes. The detailed algorithm to find such discriminative axes is given below:

### Algorithm 1: projection axes extraction

Step 1: prepare some original images and their JPEG2000 compressed versions. The compression ratio should be high enough so that every edge point in the compressed versions can be viewed as distorted. Denote all the images as 'dataset for preprocessing';

Step 2: for every image in 'dataset for preprocessing', use canny operator to detect all edge points.

Step 3: for every edge point  $(i, j)$ , use Operation 1 to get vector  $X_0(i, j)$ . Let  $S^+$  and  $S^-$  denote the subset of  $X_0(i, j)$  acquired from the original images and distorted ones, respectively;

Step 4: calculate covariance matrix  $\Sigma$  [4]:

$$\Sigma = (N^- \cdot \Sigma^- + N^+ \cdot \Sigma^+) / (N^- + N^+) \quad (1)$$

where  $N^+$  and  $N^-$  denote the number of vectors in  $S^+$  and  $S^-$ ; and  $\Sigma^+$  and  $\Sigma^-$  are the covariance matrices for  $S^+$  and  $S^-$ , respectively;

Step 5: perform PCA on  $\Sigma$ . Let  $\lambda_i$  and  $u_i$  denote the  $i^{\text{th}}$  eigenvalue and principle axis, respectively ( $i=1, 2, \dots, (r+1)^2$ );

Step 6: re-sort  $u_i$  by Eq.2. (largest  $J$  ranked first) [4]:

$$J = (u_i^T \cdot S_b \cdot u_i) / \lambda_i \quad (2)$$

where  $S_b$  denotes the inter-class dissimilarity.

Note that by taking the covariance matrix as Eq.1., we can make use of the information of the label in PCA. Moreover, by re-sorting  $u_i$  by Eq.2., those principle axes containing most discriminative information are ranked first [4]. However, we find out by experiments that the first principle axis is not discriminative enough. Therefore, in this paper, we use the second and third principle axes as projection axes in a somewhat arbitrary manner and leave this problem to the future work. The final local feature for a given edge point  $(i, j)$  can be calculated by the following operation:

**Operation 2:** for an edge point  $(i, j)$ , use Operation 1 to get a vector  $X_0(i, j)$ . Let  $x_2$  and  $x_3$  denote the projection of  $X_0(i, j)$  on  $u_2$  and  $u_3$ , respectively. Its local feature is denoted as  $X(i, j) = [x_2, x_3]^T$ .

### 2.2. Modeling the local distortion metric $Ld(i, j)$

By viewing all edge points as 'distorted' or 'un-distorted', we propose using Eq.3. to calculate the local distortion metric  $Ld(i, j)$  for a given edge point  $(i, j)$ :

$$Ld(i, j) = \frac{P(S^- | X(i, j))}{P(S^+ | X(i, j)) + P(S^- | X(i, j))} \quad (3)$$

where  $P(S^+|X(i,j))$  and  $P(S^-|X(i,j))$  are the posterior probabilities, respectively.

Using Bayes rule, Eq.3. can be written as:

$$Ld(i,j) = \frac{P(X(i,j)|S^-) \cdot ratio}{P(X(i,j)|S^+) + P(X(i,j)|S^-) \cdot ratio} \quad (4)$$

where  $P(X(i,j)|S^+)$  and  $P(X(i,j)|S^-)$  are the conditional probabilities respectively, and  $ratio = P(S^-)/P(S^+)$ .

Note that our model of the local distortion metric  $Ld(i,j)$  is not dependent on the type of local feature  $X(i,j)$ . It can be easily generalized to other types of local feature.

In this paper, we assume that the distribution of conditional distributions are Gaussian with unknown mean  $(\mu^+, \mu^-)$  and variance  $(var^+, var^-)$ . The unknown parameters can be estimated from  $S^+$  and  $S^-$  by maximum likelihood (ML) [4]:

**Algorithm 2: parameter estimation for  $Ld(i,j)$**

Step 1: prepare  $S^+$  and  $S^-$  as Algorithm 1;

Step 2: use  $S^-$  to estimate  $\mu^-$  and  $var^-$  by ML;

Step 3: use  $S^+$  to estimate  $\mu^+$  and  $var^+$  by ML.

### 2.3. Proposed NR quality assessment scheme

Based on the previous preparation, the details of the proposed NR assessment algorithm can be given as follows:

**Algorithm 3: NR quality assessment for JPEG2000 compressed image**

Step 1: use Algorithm 1 to get projection axes  $u_2$  and  $u_3$ ;

Step 2: use Algorithm 2 to estimate  $(\mu^+, \mu^-)$  and  $(var^+, var^-)$ ;

Step 3: use Canny operator to detect all edge points of the testing image;

Step 4: use Operation 2 to calculate the local feature  $X(i,j)$  for every edge point;

Step 5: calculate the distortion metric  $Dm = \frac{1}{N_{edge}} \sum_{(i,j) \in edge\ points} Ld(i,j)$ , where  $N_{edge}$  is the number

of the total edge points in the testing image;

Step 6: predict the quality score of the testing image:  $Ps = \alpha + \beta \cdot Dm^\gamma$  [7], where  $\alpha$ ,  $\beta$  and  $\gamma$  are unknown parameters.

## 3. EXPERIMENTAL RESULTS

In this section, we will show some experimental results. The image database we use in the experiment is from [8], which consists of 29 original high-resolution 24-bits/pixel RGB color images and their JPEG2000 compressed

versions with different compressed ratios. The total number of images in the database is 227. According to [8], about 25 human observers rated each image as ‘Bad’, ‘Poor’, ‘Fair’, ‘Good’ and ‘Excellent’. Mean human scores are acquired after normalizing the original raw scores and removing outliers (for more details, refer [8, 9]). In our experiment, the database is divided randomly into two sets: 14 training images together with their compressed versions and 15 testing images together with their compressed versions. ‘Dataset for preprocessing’ used in Algorithm 1 and 2 consists of 14 original images from the training set and their compressed versions with the highest compressed ratio.

Five parameters need to be set:  $r, ratio, \alpha, \beta$  and  $\gamma$ . A larger  $r$  can provide more information about the local distortion effect, however, it also need more processing time. In our experiment, we set it to be 4. The remaining four parameters are obtained by minimizing the MSE (mean-square-error) between prediction score and mean human score on the training set:

$$MSE = \frac{1}{N} \sum_k (Ps_{(k)} - Mhs_{(k)})^2 \quad (5)$$

where  $Ps_{(k)}$  and  $Mhs_{(k)}$  are the prediction score and the mean human score of the  $k^{th}$  image, respectively;  $N$  is the total number of the images in the training set.

The prediction result using our algorithm versus mean human score on the testing set is shown in Fig.1. In [7], the authors proposed using edge spread as the local blurring feature  $X(i,j)$  and in their algorithm, this feature is used directly as the local distortion metric  $Ld(i,j)$ . The result obtained by using their algorithm on the same testing set is shown in Fig.2. To demonstrate the effectiveness of the proposed model of the local distortion metric, we also apply it to edge spread. The linear correlation values between the prediction result and mean human score on the testing set for the above three algorithms are illustrated in Table 1. An example of applying our algorithm to assess the quality of JPEG2000 compressed images is shown in Fig.3.

Assessment algorithm	Linear correlation value
algorithm 1	74.0%
algorithm 2	78.7%
algorithm 3	85.7%

Table 1. Linear correlation values between the prediction result and mean human score. Algorithm 1: the one proposed in [7]; algorithm 2: applying our local distortion metric model to edge spread; algorithm 3: our proposed algorithm.

## 4. CONCLUSION

In this paper, we have proposed a new NR quality assessment algorithm for JPEG2000 compressed images.

By viewing all edge points as two classes ('distorted' and 'un-distorted'), we take the advantage of PCA to extract local feature of a given edge point. The proposed local feature can indicate both blurring and ringing in JPEG2000 compressed images. Such feature is fast to extract compared with that through wavelet decomposition. We have also proposed using the probabilities of the given edge point being 'distorted' and 'un-distorted' to model the local distortion metric. Compared with the existing methods, it can better describe the relationship between the local feature and the local distortion metric. The proposed new model is straightforward and can be easily generalized to other types of local feature. Further work includes extending the proposed algorithm to assess other kinds of distortion.

### 5. ACKNOWLEDGEMENTS

This work was supported by National High Technology Research and Development Program of China (863 Program) under contract No.2001AA114190.

### 6. REFERENCES

[1] M.D. Adams, "The JPEG-2000 still image compression standard", ISO/IEC JTC 1/SC 29/WG 1, Document N2412, Sep. 2001.  
 [2] C.J. van den Branden, "Special issue on image and video quality metrics", *Signal Processing*, vol 70, Nov. 1998.  
 [3] J. Caviedes and S. Gurbuz, "No-reference sharpness metric based on local edge kurtosis", *International Conference on Image Processing*, vol. 3, pp. 53-56, Sep. 2002.  
 [4] R.O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification (2nd Edition)*, Wiley-Interscience, Oct. 2000.

[5] X. Li, "Blind image quality assessment", *International Conference on Image Processing*, vol. 1, pp. 24-28, Sep. 2002.  
 [6] P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi, "A no-reference perceptual blur metric", *International Conference on Image Processing*, vol. 3, pp. 57-60, Sep. 2002.  
 [7] E.P. Ong, W.S. Lin, Z.K. Lu, S.S. Yao, X.K. Yang and L.F. Jiang, "No-reference JPEG-2000 image quality metric", *International Conference on Multimedia and Expo*, vol. 1, pp. 6-9, July. 2003.  
 [8] H.R. Sheikh, Z. Wang, L. Cormack and A. C. Bovik, "LIVE image quality assessment database", <http://live.ece.utexas.edu/research/quality>.  
 [9] H.R. Sheikh, Z. Wang, L. Cormack and A. C. Bovik, "Blind quality assessment for JPEG2000 compressed images", *International Conference on Image Processing*, vol. 2, pp. 3-6, Sep. 2002.  
 [10] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment", <http://www.vqeg.org/>, Mar. 2000.  
 [11] S. Winkler, "Issues in vision modeling for perceptual video quality assessment", *Signal Processing*, vol. 78, pp. 231-251, Oct. 1999.

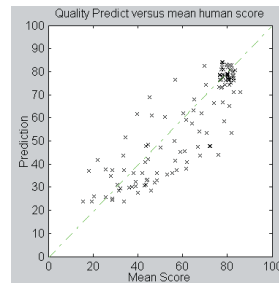


Fig.1. Quality prediction using our algorithm versus mean human score

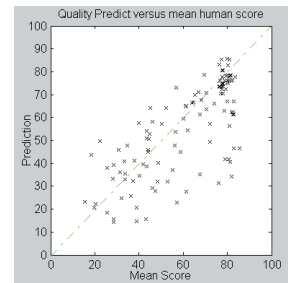


Fig.2. Quality prediction using the algorithm proposed in [7] versus mean human score



(a)  $P_s = 80.94$   $M_{hs} = 82.49$



(b)  $P_s = 30.76$   $M_{hs} = 32.11$



(c)  $P_s = 27.20$   $M_{hs} = 28.47$

Fig.3. An example of applying our algorithm to assess the quality of JPEG2000 compressed images. (a) the original uncompressed image; (b) some distortion in the compressed image; (c) lots of distortion in the compressed image.  $P_s$  is the prediction result;  $M_{hs}$  is the mean human score.