Learning No-Reference Quality Metric by Examples[*]

Hanghang Tong[1], Mingjing Li[2], Hong-Jiang Zhang[2], Changshui Zhang[3], Jingrui He[1], Wei-Ying Ma[2]
[1, 3]Department of Automation, Tsinghua University, Beijing 100084, China
[2]Microsoft Research Asia, 49 Zhichun Road, Beijing 100080, China
[1]{walkstar98, hejingrui98}@mails.tsinghua.edu.cn, [2]{mjli, hjzhang, wyma}@microsoft.com,
[3]zcs@tsinghua.edu.cn

Abstract

In this paper, a novel learning based method is proposed for No-Reference image quality assessment. Instead of examining the exact prior knowledge for the given type of distortion and finding a suitable way to represent it, our method aims to directly get the quality metric by means of learning. At first, some training examples are prepared for both high-quality and low-quality classes; then a binary classifier is built on the training set; finally the quality metric of an un-labeled example is denoted by the extent to which it belongs to these two classes. Different schemes to acquire examples from a given image, to build the binary classifier and to model the quality metric are proposed and investigated. While most existing methods are tailored for some specific distortion type, the proposed method might provide a general solution for No-Reference image quality assessment. Experimental results on JPEG and JPEG2000 compressed images validate the effectiveness of the proposed method.

1. Introduction

Image quality assessment aims to automatically provide an objective measurement for the quality of a given image which is consistent with the result given by human observers [2, 12, 16, 18, 20]. With the prevalence of digital images, automatic image quality assessment is highly desirable in the following ways [14, 16, 18]: 1) to monitor and control image quality for quality control systems; 2) to benchmark image processing systems; 3) to optimize algorithms and parameters; 4) to help home users better manage their digital photos and evaluate their expertise in photographing.

According to the prior knowledge used in the assessment, we can categorize existing image quality metrics into three classes [16, 18]: full-reference (FR), reduce-reference (RR) and no-reference (NR). Both FR and RR are essentially fidelity assessment since they need the original un-distorted image as a reference either fully or partially [14, 16, 18]. However, in many situations, the original un-distorted image might not exist or be very hard to obtain [9, 13, 14]. On the other hand, it is very easy for human observers to assess image quality without using any reference image. In recent years, NR image quality assessment has attracted the attention of more and more researchers [3, 8, 10, 13, 17, 19].

Due to the limited understanding of the human vision system (HVS), most, if not all, of the existing NR assessment algorithms focus on distortion measurement, in which the quality metric is described by the extent to which the image has probably been distorted [13, 14]. No matter whether explicitly or implicitly, the general flow of these algorithms can be summarized as follows [14]: 1) find some discriminative local feature; 2) use local feature to model local distortion metric; 3) average local distortion metric over the whole image to get a overall distortion metric Qm ; 4) use Qm to predict image quality score Ps which is consistent with human perception. Finding suitable local feature and modeling the local distortion metric are two key steps within the whole algorithm.

Most of existing methods focus on blurring, blockiness and ringing. For example, the authors in [17, 19] proposed using blockiness difference and activity of the image signal as local distortion feature for blockiness and blurring, and using a nonlinear combination of them to model the local blurring metric; the authors in [9, 10] proposed using edge spread as the local blurring feature which is used directly as the local distortion metric; the authors in [13] proposed using wavelet coefficients as the local feature for blurring

[*] This work was performed at Microsoft Research Asia.

and ringing in JPEG2000 compressed images, and the local distortion metric is simply denoted as "significant" or "insignificant" by a threshold. The main drawbacks of the existing methods are [14]: 1) extracting local distortion feature is quite distortion-type dependent (they need the exact prior knowledge for the given type of distortion and a suitable way to represent it); 2) the way they model the local distortion metric seems to over simplify the relationship between the local feature and the local distortion metric.

To address the drawbacks of existing NR methods for JPEG2000 compressed images, by viewing all edge points as either "distorted" or "un-distorted", we proposed in [14] using principal component analysis (PCA) to extract the local feature of a given edge point, which indicates both blurring and ringing. We also proposed using the probabilities of the given edge point being "distorted" and "un-distorted" to model the local distortion metric by Bayes rule. However, there are still some limitations: 1) both the way we select projection axis and the Gaussian assumption for the conditional probabilities are somewhat arbitrary; and 2) the local distortion metric takes the ratio between the priors of "distorted" and "un-distorted" as a parameter, which is hard to obtain in practice.

In this paper, we extend our work in [14] and propose a learning based method to model the quality metric. In our method, instead of examining the exact prior knowledge for the given type of distortion and finding a suitable way to represent it, we aim to directly get the quality metric by means of learning. The basic idea of our method is that images of similar quality should share some common law in their low-level features and this common law might be learned from a set of given examples. To achieve this goal, we first prepare some examples to compose two classes: "high quality" and "low quality"; then a binary classifier is built on these two classes so that the two classes will be separated as far as possible; finally, the quality metric of an un-labeled example is denoted by the extent to which it belongs to these two classes. In contrast to the traditional methods which are tailored for some specific distortion type, ours might provide a general solution for NR image quality assessment. Systematic experiments on JPEG and JPEG2000 compressed images validate the effectiveness of the proposed method.

The examples for the training set can be one point or block within the given image, where the distortion might occur. To build a binary classifier, we propose two different schemes: one resorts to boosting to perform feature selection and classifier training simultaneously; the other incorporates the label information into PCA for feature re-extraction and feature de-correlation; followed by Maximum Marginal Diversity (MMD) [15] for feature selection and Bayesian classifier for classification. While the second scheme is based on our previous work [14] on the whole, its limitations mentioned above are addressed in this paper. Furthermore, according to the different forms of the trained classifiers, we also propose two schemes to model the quality metric.

The organization of this paper is as follows. Section 2 presents the flowchart of the proposed method. We address the issues of training set preparation, classifier building and quality metric modeling in Section 3, Section 4 and Section 5 respectively. Systematic experimental results are presented in Section 6. Finally, we conclude the paper in Section 7.

2. The flowchart of the proposed method

The basic idea of our method is that images of similar quality should share some common law in their low-level features and this common law might be learned from a set of given examples. The flowchart of the proposed learning-based method for NR image quality assessment is summarized in Fig. 1. Its details are given as follows.

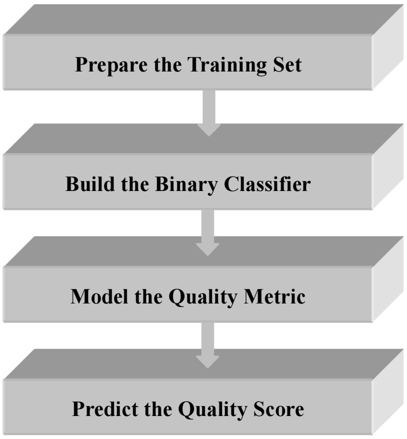


Fig. 1. The flowchart of the proposed method

- First, we need to prepare some examples $E(i)$ both of "high quality" and of "low quality" to compose the training set, where $i = 1, 2, \cdots, N$ and N is the total number of examples. Those "high quality" examples $E^+(i)$ compose S^+ (the subset for positive examples), where $i = 1, 2, \cdots, N^+$ and N^+ is the total number of "high quality" examples. Those "low quality" examples $E^-(i)$ compose S^-

(the subset for negative examples), where $i=1,2,\cdots,N^-$, N^- is the total number of "low quality" examples and $N^+ + N^- = N$. Here, each example $E(i)$ can be one point or block of a given image and it is denoted as an initial feature vector: $E(i) \rightarrow F(i)\,(i=1,2,\cdots,N)$.

- Next, a binary classifier is built on $\{F(i),Y(i)\}$ $(i=1,2,\cdots,N)$, which separates the positive and negative examples as far as possible, where $Y(i)=+1$ if $E(i)\in S^+$ and $Y(i)=-1$ otherwise.
- Then, the quality metric $Qm(j)$ for a new example j can be modeled through its probabilities of being "positive" and being "negative":

$$Qm(j) = f(P(S^+ \mid F(i)), P(S^- \mid F(i))) \qquad (1)$$

where $P(S^+ \mid F(i))$ and $P(S^- \mid F(i))$ are the posteriors of the example j being "positive" and "negative" respectively which can be acquired from the trained classifier; and $f: R^2 \rightarrow R$ is some kind of function which maps the posteriors to a quality metric. Average $Qm(j)$ over the whole image I, we get an overall distortion metric $Qm(I)$ for the given image.

- Finally, predict the quality score $Ps(I)$ of the given image so that it will be consistent with the result given by human observers [9, 10, 14]:

$$Ps(I) = \alpha + \beta \cdot Qm(I)^{\gamma} \qquad (2)$$

where α , β and γ are unknown parameters and can be determined by minimizing the MSE (mean-square-error) between prediction score and mean human score:

$$MSE = \frac{1}{N_{aho}} \sum_{I=1}^{N_{aho}} (Ps(I) - Mhs(I))^2 \qquad (3)$$

where $Mhs(I)$ is the mean human score of the I^{th} image; N_{aho} is the number of the images used to determine the parameters.

3. Prepare the training set

For most existing types of distortion, we can identify them locally. For example, blurring is perceptually apparent around edges; ringing usually appears near sharp edges [1, 9, 10]; while blockiness often occurs in JPEG compressed images and is visible at the boundary of two adjacent blocks (usually 8×8) used in the compression stage [17, 19]. Based on the above observation, we propose the following two operations to acquire examples and form their corresponding initial features from a given image:

Opt. 1*: detect all edge points of a given image. Every edge point is viewed as an example* $E(i)$. *For each edge point* $E(i)$, *assign it to the center of a block and arrange all the pixels within this block in a vector which is used as the corresponding initial feature* $F(i)$. *Let* r *denote the size of the block and* $F(i)$ *is* r^2 *dimensional.*

Opt. 2*: divide a given image into small blocks. Every block is viewed as an example* $E(i)$. *For each block* $E(i)$, *arrange all the pixels within it in a vector which is used as the corresponding initial feature* $F(i)$. *Let* r *denote the size of the block and* $F(i)$ *is* r^2 *dimensional.*

Note that Opt. 1 is designed for blurring and ringing, while Opt. 2 is mainly designed for blockiness and it also provides a rough description for blurring and ringing. Opt. 2 is more efficient than Opt. 1 in terms of processing time since it does not require edge detection as a preprocessing step.

Based on the above preparation, the training set can be set up as follows:

Algorithm 1: Prepare the training set

1. Prepare some original un-distorted images and their distorted versions. There must be enough distortion in the distorted image so that every example in it can be viewed as "low quality";
2. For each image, use Opt. 1 or Opt. 2 to obtain all examples $E(i)$ and their corresponding initial features $F(i)$;
3. Add $E(i)$ to the training set. To be specific, if $E(i)$ comes from an original image, add it to S^+ ; else add it to S^- .

4. Build the binary classifier

It is always a challenge to select a good feature set for classification. We propose two different schemes for our task in this paper.

4.1. Boosting based scheme

Recent developments in the field of machine learning have demonstrated that boosting based methods may have a satisfactory performance by combining weak learners [5, 6]. Furthermore, the boosting procedure can also be viewed as a feature selection process if the weak learner uses a single feature in each stage. Benefiting from such cherished properties, our first scheme is very simple. That is, we simply use some boosting based method to train on the initial feature set and in this context, boosting performs both feature selection and classifier training simultaneously.

4.2. Feature re-extraction based scheme

Theoretically, Bayesian classifier can produce the minimum classification error. However, we can not directly apply it to our task since the dimensionality of the initial feature vector is very high, which makes it very difficult to estimate probability distribution that is necessary for Bayesian classifier. Therefore, we have to select a small subset from the initial feature vector, whose elements are most discriminative.

On the other hand, we find out by experiments that the discriminative power of each dimension in the initial feature vector is very weak, which means a small subset of it might not be adequate for a satisfactory classification performance.

Based on the above observations, some more powerful features should be re-extracted and selected from the initial feature vector. In [14], we have proposed using PCA to perform feature re-extraction. However, the way we select the feature (the associated projection axis) in [14] was somewhat arbitrary. In this paper, MMD [15] is adopted to perform feature selection. The detailed scheme is given as follows:

Algorithm 2: Feature re-extraction based scheme

```
1. Normalize $F(i)$ $(i=1,2,\cdots,(N^{+}+N^{-}))$ on each dimension to $[0,1]$;
2. Calculate covariance matrix $\Sigma$ [4, 7]:
       $\Sigma=\left(N^{-}\cdot\Sigma^{-}+N^{+}\cdot\Sigma^{+}\right)/\left(N^{-}+N^{+}\right)$    (4)
   where $\Sigma^{-}$ and $\Sigma^{+}$ are covariance matrix of $S^{-}$ and $S^{+}$, respectively
3. Perform PCA on $\Sigma$. Let $u_j(j=1,2,\cdots,r^2)$ denote the $j^{th}$ principle axis;
4. The new feature set is denoted as $F'(i)=\left[x_1,x_2,\cdots,x_{r^2}\right]^T$, where $x_j$ denotes the projection of $F(i)$ on $u_j$;
5. Use MMD to select the $M$ most discriminative features $F_s^{'}(i)\left(s=1,\cdots,M\right)$;
6. Feed $F_s^{'}(i)$ to Bayesian classifier.
```

Note that by taking the covariance matrix as Eq. 4., we can make use of the label information in PCA to re-extract some more discriminative features from the initial feature vector. Moreover, de-correlation on different dimensions by PCA also makes the subsequent feature selection step more reliable.

5. Model the quality metric

After the binary classifier is built, the quality metric $Qm(j)$ for a new example j can be modeled through its probabilities of being "positive" and being "negative" as Eq. 1. To be specific, we propose two schemes according to the different forms of the trained classifiers. In both cases, the overall quality metric $Qm(I)$ for a given image I is obtained by averaging $Qm(j)$ over the whole image.

5.1. Quality metric for Boosting

Among the many choices, we favor Real-AdaBoost [5, 6] here for its relative simplicity and clear physical meaning: since in Real-AdaBoost, the output of every weak learner indicates the probability of a given example j being of "high quality" or being of "low quality", by combining the outputs of all the weak learners obtained in the training stage, we get a confident coefficient for its quality metric:

$$Qm\left(j\right)=\sum_{t=1}^{T}h_t\left(F\left(j\right)\right) \qquad (5)$$

where $h_t(t=1,2,\cdots,T)$ denote the t^{th} weak learner of Real-AdaBoost; T is the total number of weak learners.

5.2. Quality metric for Bayesian classifier

In this case, we get a Bayesian classifier and the quality metric $Qm(j)$ can be modeled through its posterior probabilities:

$$Q(j)=\frac{P(S^{+}\mid F_s^{'}(j))}{P(S^{-}\mid F_s^{'}(j))+P(S^{+}\mid F_s^{'}(j))} \qquad (6)$$

where $P(S^{+}\mid F_s^{'}(j))$ and $P(S^{-}\mid F_s^{'}(j))$ are the posterior probabilities, respectively.

Using Bayes rule, Eq. 6. can be converted to:

$$Q(j)=\frac{P(F_s^{'}(j)\mid S^{+})\cdot ratio}{P(F_s^{'}(j)\mid S^{-})\cdot ratio+P(F_s^{'}(j)\mid S^{+})} \qquad (7)$$

where $P(F_s^{'}(j)|S^{-})$ and $P(F_s^{'}(j)|S^{+})$ are the conditional probabilities respectively, and $ratio=P(S^{+})/P(S^{-})$.

In [14], $ratio$ is viewed as an additional parameter and the final quality metric is a function of $ratio$. Although we can obtain it by optimizing Eq. 2. together with α, β and γ theoretically, it is very difficult in practice. So in this paper, we simply set $ratio = N^{+}/N^{-}$.

Another issue with Eq. 7. is the estimation of the conditional probabilities. In [14], we simply assume them as Gaussian distribution, the parameters of which can be estimated by maximum likelihood (ML) [4, 7]. In this paper, in addition to this simple (and somewhat arbitrary) strategy, we also propose using multi-dimensional histogram (MDH) to estimate the conditional probabilities, since the dimensionality of $F_s^{'}(j)$ is quite low (2 in our experiments).

6. Experimental results

What we try to propose in this paper is a general solution for NR quality assessment. In this Section, we will examine the performance of the proposed method for JPEG and JPEG2000 compressed images.

6.1. Operation and parameter settings

In our experiments, there are two training sets and one testing set: "training set 1" for training the classifier, "training set 2" for determining the parameters (α, β and γ) which are necessary to predict quality scores in Eq. 2., and "testing set" for examining the performance of the proposed algorithms.

We use the linear correlation value (LCV) and MSE between the prediction results and mean human score to evaluate the performance of various methods. Different schemes to acquire examples from a given image, to build the binary classifier and to model the quality metric will be evaluated. Moreover, to estimate the conditional probabilities for Bayesian classifier, both ML and MDH will be investigated.

A set of parameters need to be determined:

- A larger block size r can provide more information about the local distortion effect, however, it also need more processing time. In our experiment, it is set to 9 for Opt. 1 and 12 for Opt. 2;
- Two parameters in Real-AdaBoost (the bin number n_{Bin} and the weak learner number T) are determined by cross-validation (5 folds) [4, 7];
- The number of selected features in Algorithm 2 M is set to 2;
- In the case that the conditional probabilities are estimated by MDH, the bin number on each dimension is 20.

6.2. Assessment results for JPEG images

The image database that we use in this part is from [11], which consists of 29 original high-resolution 24-bits/pixel RGB color images and their JPEG compressed versions with different compressed ratios. The total number of the images in the database is 234. According to [11], about 25 human observers rated each image as "Bad", "Poor", "Fair", "Good" or "Excellent". Mean human scores are acquired after normalizing the original raw scores and removing outliers. (For more details of the subjective experiment, refer to [11, 19].)

In our experiment, the database is randomly divided into two sets: 15 images together with their compressed versions compose "training set 2" and 14 images together with their compressed versions compose "testing set". "Training set 1" is set up by 15 original images from "training set 2" and their compressed versions with the highest compressed ratio.

The main distortion in JPEG compressed images is blockiness and blurring [17, 19]. In order to form the training set, we adopt Opt. 2 to acquire the examples from a given image for reasons given in Section 3.

LCV and MSE by various methods are listed in Table 1. In [19], the authors proposed using blockiness difference and activity of the image signal as local distortion feature for blockiness and blurring, and using a nonlinear combination of them to model the local blurring metric. The result obtained by their algorithm on the same testing set is also shown in Table 1 for comparison. It can be seen that all the results by our methods are comparable with those by [19]. For Real-AdaBoost, it even outperforms those by [19]. While the algorithm in [19] is based on the exact prior knowledge about what blockiness and blurring are and how to describe them, such knowledge is not required in our methods.

Table 1. Assessment results for JPEG images

Result / **Method**		LCV	MSE
Real-AdaBoost		**92.3%**	**9.1**
Bayesian	ML	86.4%	12.4
	MDH	88.0%	11.5
Algorithm in [19]		90.1	9.8

The prediction result using Real-AdaBoost versus mean human score on "testing set" is shown in Fig. 2. An example of applying Real-AdaBoost to assess the quality of JPEG compressed images is shown in Fig. 3.

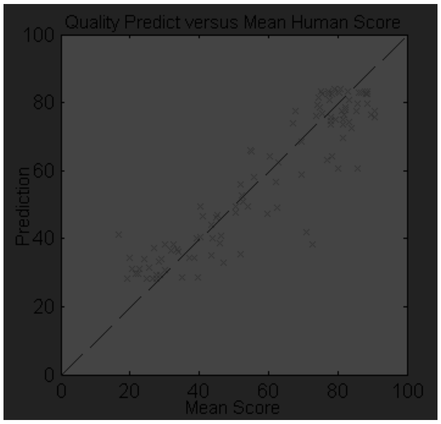


Fig. 2. Quality prediction versus mean human score for JPEG compressed images

6.3. Assessment results for JPEG2000 images

The image database that we use in this part is also from [11], 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 the images in the database is 227. The subjective experiment is similar with that of JPEG compressed images.

In our experiment, the database is randomly divided into two sets: 14 images together with their compressed versions compose "training set 2" and 15 images together with their compressed versions compose "testing set". "Training set 1" is set up by 14 original images from "training set 2" and their compressed versions with the highest compressed ratio.

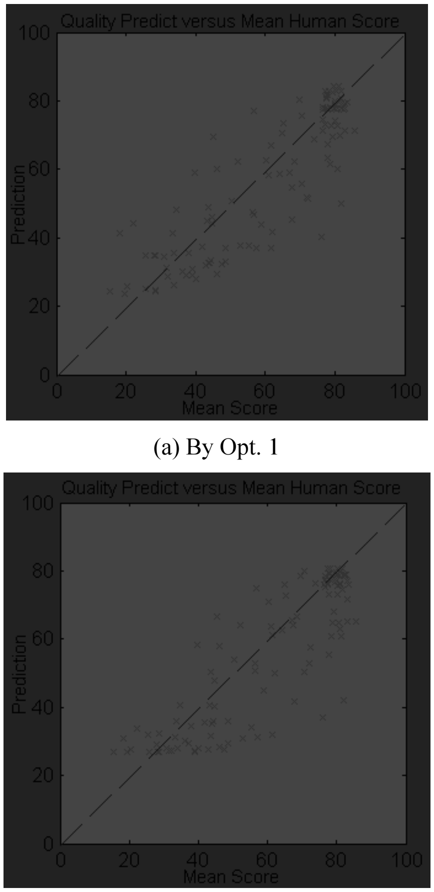
The main distortion in JPEG2000 compressed images is blurring and ringing [1, 9, 10, 13]. In order to form the training set, we adopt both Opt. 1 and Opt. 2 to acquire the examples from a given image.

LCV and MSE by various methods are listed in Table 2. In [10], the authors proposed using edge spread as the local blurring feature which is used directly as the local distortion metric. The result obtained by their algorithm on the same testing set is also shown in Table 2 for comparison. In all cases, our methods outperform the one in [10] by a large margin. Comparing the different operations to acquire the examples, it can be seen that Opt. 1 outperforms Opt. 2. However, it is worth noticing that in Opt. 2, there is no edge detection step which is time-consuming so that it can serve as the fast version of Opt. 1 in this context.

Table 2. Assessment results for JPEG2000 images

Result **Method**			LCV	MSE
Opt.1	Real-AdaBoost		**86.3%**	**11.1**
	Bayesian	ML	85.0%	11.7
		MDH	85.3%	11.4
Opt.2	Real-AdaBoost		85.6%	11.8
	Bayesian	ML	81.6%	13.0
		MDH	80.6%	13.2
Algorithm in [10]			74.0%	15.9

The prediction result using Real-AdaBoost versus mean human score on "testing set" is shown in Fig. 4. An example of applying Real-AdaBoost to assess the quality of JPEG2000 compressed images is shown in Fig. 5.



(a) By Opt. 1

(b) By Opt. 2

Fig. 4. Quality prediction versus mean human score for JPEG2000 compressed images

7. Conclusion

In this paper, we have extended our previous work in [14] and proposed a novel learning method for NR image quality assessment. In our method, we first prepare some examples to compose two classes: "high quality" and "low quality"; then a binary classifier is trained on these two classes; finally, the quality metric of an un-labeled example is denoted by the extent to which it belongs to these two classes. Different schemes to acquire examples from a given image, to building the binary classifier and to model the quality metric are proposed and investigated. In contrast to the traditional methods which are tailored for some specific distortion type, our method might provide a general solution for NR image quality assessment. Systematic subjective experiments on JPEG and JPEG2000 compressed images demonstrate the effectiveness of the proposed method. Future work includes testing on other types of distortion and integrating prior knowledge with the proposed method.

8. Acknowledgements

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

9. 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] Y. Freund, and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Journal of Computer and System Sciences*, 55(1): 119-139, 1997.

[6] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting", *The Annual of Statistics*, 28(2): 337-374, 2000.

[7] T. Hastie, R. Tibshirani, and J.H. Friedman, *The Elements of Statistical Learning*. Springer Verlag, Aug. 2001.

[8] X. Li, "Blind image quality assessment", *International Conference on Image Processing*, vol. 1, pp. 24-28, Sep. 2002.

[9] 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.

[10] 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.

[11] H.R. Sheikh, Z. Wang, L. Cormack and A. C. Bovik, "LIVE image quality assessment database", http://live.ece.utexas.edu/research/quality.

[12] R. Shaw, "A century of image quality", *IS&T's PICS Conference*, pp. 221-224, Savannah, GA, 1999.

[13] 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.

[14] H.H. Tong, M.J. Li, H.J. Zhang, and C.S. Zhang, "No-reference quality assessment for JPEG2000 compressed images", *To appear in International Conference on Image Processing,* 2004.

[15] N. Vasconcelos, "Feature selection by maximum marginal diversity: optimality and implications for visual recognition", *Proceeding on Computer Vision and Pattern Recognition,* vol. 1, pp. 762-769, 2003.

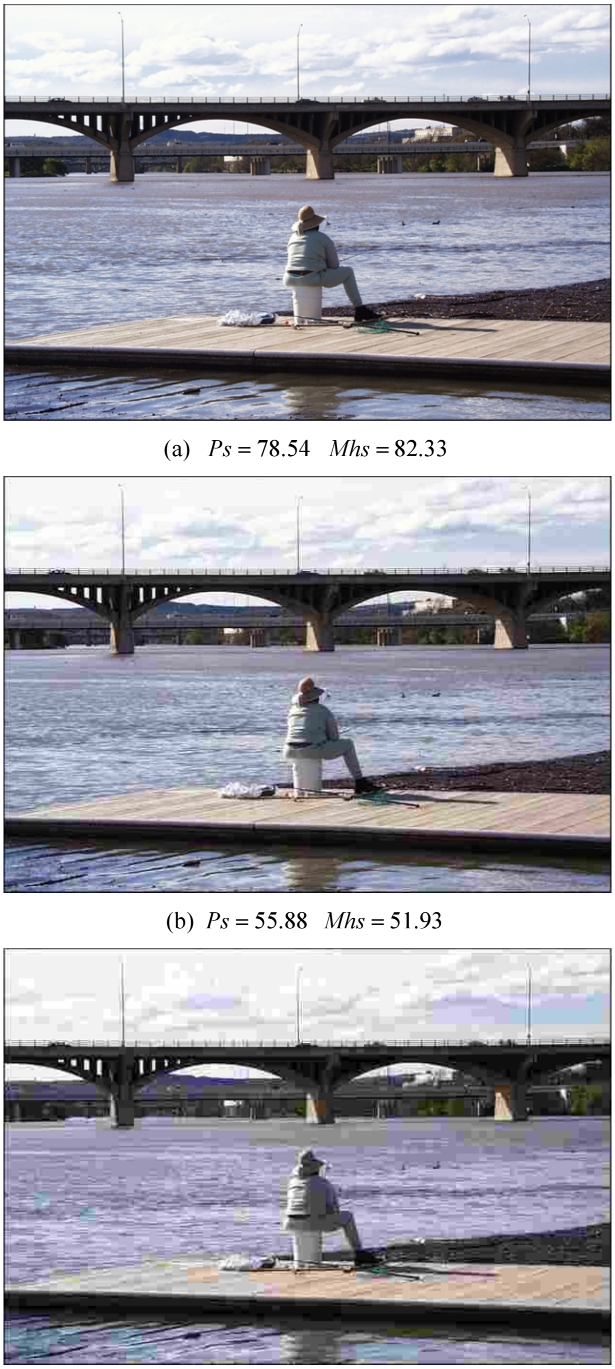
[16] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment", http://www.vqeg.org/, Mar. 2000.

[17] Z. Wang, A.C. Bovik, and B.L. Evans, "Blind measurement of blocking artifacts in images", *International Conference on Image Processing*, pp. 981-984, Sept. 2000.

[18] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", *IEEE Transaction on Image Proc.,* vol. 13, no. 4, Apr. 2004.

[19] Z. Wang, H.R. Sheikh, and A.C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images". *International Conference on Image Processing*, vol. 2, pp. 3-6, Sep. 2002.

[20] S. Winkler, "Issues in vision modeling for perceptual video quality assessment", *Signal Processing*, vol. 78, pp. 231-251, Oct. 1999.

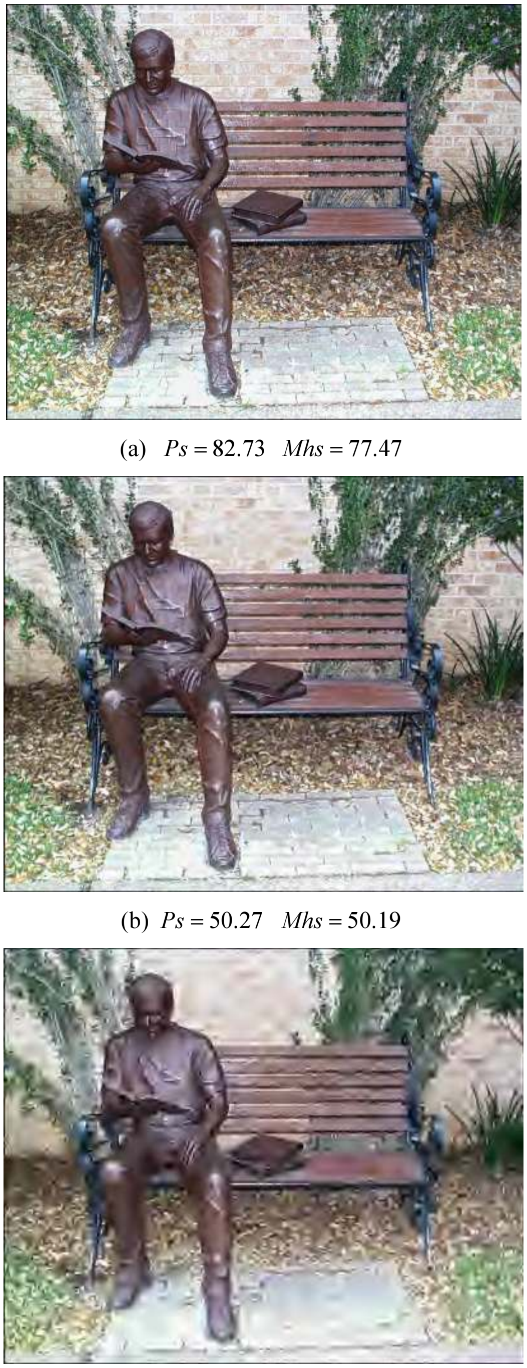


(a) $Ps = 78.54$ $Mhs = 82.33$

(b) $Ps = 55.88$ $Mhs = 51.93$

(c) $Ps = 40.12$ $Mhs = 39.13$

Fig.3. An example of applying Real-AdaBoost to assess the quality of JPEG compressed images. (a) the original uncompressed image; (b) some distortion in the compressed image; (c) lots of distortion in the compressed image. Ps is the prediction result; Mhs is the mean human score.



(a) $Ps = 82.73$ $Mhs = 77.47$

(b) $Ps = 50.27$ $Mhs = 50.19$

(c) $Ps = 31.46$ $Mhs = 31.57$

Fig.5. An example of applying Real-AdaBoost to assess the quality of JPEG2000 compressed images by Opt. 1. (a) the original uncompressed image; (b) some distortion in the compressed image; (c) lots of distortion in the compressed image. Ps is the prediction result; Mhs is the mean human score.