

Opinion Aware Blind Image Quality Assessment - A Comparison of three Learning Approaches

Gopika S ^{#1}, D. Malathi^{#2}, J. D. Dorathi Jayaseeli ^{#3}

[#] Department of Computer Science and Engineering, SRM University, Chennai , India.

¹ gopika.s@gmail.com

² malathi.d@ktr.srmuniv.ac.in

³ dorathijayaseeli.jd@ktr.srmuniv.ac.in

Abstract— Human society is living in the age of high speed development and increase of high volume of visual data. It has become a necessity to evaluate the quality of these visual data or images in many applications. This approach of quality assessment focuses on collecting feature vector from Natural Scene Images that reflects the shape parameter and correlation between adjacent pixels. A set of 36 features is collected from each image and training is done on images available in the data set. Then, three learning approaches – General Classification, KNN approach, and Distortion Specific approach are done on this feature vector. The training is done by taking the Mean Opinion Score (MOS) values available at the TID data set. Each method reflected its applicability and accuracy. It has been observed that distortion specific method out performs the other two. Specifically, the area where each one of these methods can be applied is also identified which is of great help to image manipulators.

Keyword- Blind/Non-Reference, Distortion Specific, Normalized Luminance, Quality Assessment

I. INTRODUCTION

The age we live in is a highly digital world. Progresses in technology have helped in images and videos to be captured, stored, shared and viewed easily. All these types of manipulations happening to the image gradually lead to a reduction in quality or content in the image. This led to an increase in development of tools for assessing image quality. Image Quality Assessment (IQA) tools help in rating the way humans perceive the quality of images. It is now a challenge for research community to develop the most accurate quality assessment tool. Investigators in quality assessment of images have endeavored to explore how the existence of these distortions disturbs the viewing experience. The best measure of quality assessment is to exhibit the images to a large sample of human population and to collect their opinion. The average of those opinions will give the average quality measure commonly indicated by subjective assessment. In objective assessment, the evaluation of quality is done using algorithms. The subjective assessment of quality takes lot of time and is impractical. Current IQA methods can be categorized into three based on the presence of the original image: i) complete/full reference IQA [15]; ii) less/reduced reference IQA; and 3) without/no reference/Blind IQA [18] (BIQA). BIQA is applied more in real world.

Human eyes can easily distinguish the distortions or noises in natural images. This is because of the fact that there exist particular structures that separate the unnatural from the natural scenes. Such structures are called Natural Scene Statistics (NSS) [11]. Also, natural images are highly non-random with interdependencies within them [19], [20].

Earlier work in BIQA was based on knowledge of the type of distortion happened to the image and then, as developments happened, the distortion detecting algorithm determined the type of distortion and based on that, quality was assessed [5],[7]. Recently, techniques have been established to directly plot features in images to quality scores without looking into the type of distortion [10]. For example, Saad et al. [5] has developed the methods BLIINDS and BLIINDS-II [12] to degenerate the image features to quality scores.

Quality Assessment Based on DCT: Michele Saad, Alan C. Bovik, Christophe Charrier developed a method called by the name BLIINDS [5] to effectively assess quality in Natural Scene Statistics (NSS). The algorithm relies on Discrete Cosine Transform (DCT) for feature extraction. The local DCT contrast value is calculated and is used for determining the distortion. The histogram pattern of distorted images exhibits a higher peak at '0' and the variance value also gets transformed accordingly. The DCT coefficient Kurtosis value is computed for this. The degree of this peak and tail weight are quantified to get the distortion index.

Anush Krishna Moorthy and Alan C. Bovik [4] in 2011 devised a method for blind image quality assessment using wavelet transform coefficients (DIIVINE). Here a loose wavelet transform is applied on to the image and the scale-space-orientation of the image is noted. They supply a set of statistical features and they are stacked to form a vector. Using this feature vector, the distortion type is determined and the quality score is calculated. Here, a regression model is developed for each distortion category and mapping is done from distortion to quality value..

In 2013, Xinbo Gao and Xuedong Li [1] successfully implemented a method of Image Quality Analysis by multiple Kernel Learning. The authors have used the secondary and tertiary properties of wavelet transform to get the features. Secondary properties of Wavelet transform gives Non-Gaussianity, Clustering and Persistency. Tertiary properties give Exponential Decay Characteristics (EDC). Feature vector was created by the marginal distribution of coefficients. The histogram of wavelet coefficients in different sub bands was calculated and all those were collected to get the long feature vector.

The EDC show that at finer scales the energy decays exponentially. The EDC variations are observed and based on that authors concluded that EDC can replicate the types of distortions. Thus the feature vector included Non-Gaussianity, Local dependency and EDC characteristics.

Another assessment algorithm using Deep Learning was developed recently by Weilong Hou and Xinbo Gao [9], [17]. They evaluated the quality of NSS image by means of collecting linguistic description rules. Normally, subjects are used for providing linguistic descriptions; this method learns the evaluation on quality and gives numerical marks for comparison [16], [13].

Each of the above methods throws light on different approaches to quality prediction. Section II discusses the proposed method of quality assessment of NSS image by feature extraction, Section III gives its comparison with three different feature learning and quality assessment method and section IV gives an analysis of the result. This study attempts to give the accuracy of the method under two conditions; (i) when the distortion type is known and (ii) when it is unknown to the application. Hence real world applications find the proposed method useful by having awareness about the approach to be applied for quality prediction.

II. PROPOSED METHOD-OPINION AWARE BLIND-QUALITY ASSESSMENT

Most of the existing quality assessment methods concentrate on mapping the image domain from spatial to another and doing processing on that. We are processing in spatial domain since we consider that the visual cortex of human beings responds well in the spatial domain.

The normalized luminance value of the image offers a good knowledge about the amount of distortion happened to the image and also the amount of naturalness presented in the image under consideration. In this paper, we examine the quality of an NSS image by inspecting the normalized luminance value [4] and then the pairwise product of these coefficients are modeled to compute the quality score. We then demonstrate how the statistical features obtained from the image can represent quality and that the representation confirms well with human observation of quality. The normalized luminance value of an image can be computed as

$$\bar{I}(i, j) = I(i, j) - \mu(i, j) / \sigma(i, j) + c \tag{1}$$

Where i and j are spatial indices and $i=1, 2, \dots, M$ and $j=1, 2, \dots, N$. The mean, $\mu(i, j)$ is given by

$$\mu(i, j) = \sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} I(i, j) \tag{2}$$

and variance
$$\sigma(i, j) = \sqrt{\sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} I_{k,l}(i, j) - \mu(i, j)^2} \tag{3}$$

where, $w = \{W_{k,l} | k = -K, \dots, K, l = -L, \dots, L\}$ is a 2D circularly-symmetric Gaussian weighting function. In our implementation, K and L are taken as 3.

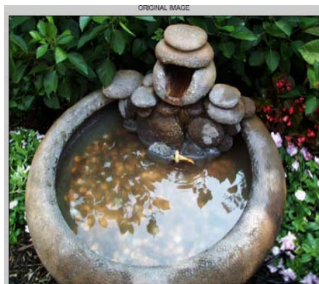


Fig. 1. Undistorted original NSS image



Fig. 2. Two Low Quality distorted version (by high and medium white noise distortion) of Fig. 1.

Careful Observations shows that the luminance values give a unit Gaussian distribution in the lack of any distortion for NSS images [4]. These normalized luminance values are also termed as Mean Subtracted Contrast Normalized values (MSCN). Our experiments are based on the theory that these coefficients have statistical features that are altered by the occurrence of distortion and that enumerating the deviations will help in predicting the distortion occurred in an image and also its quality. To see the variations of MSCN coefficient, the coefficient values were plotted for an undistorted pristine image and on a distorted low quality image. The MSCN coefficients clearly show the distinction between a good quality undistorted natural image and a low quality distorted image as shown in Fig. 1 and Fig. 2. The shape of the curve is perfectly Gaussian for an undistorted NSS image and is exhibiting peak around zero for its distorted version. The shape of the curve varies based on the type of distortion happening to the image i.e. each type of distortion changes the shape of the curve in its characteristic way. Studies have revealed that a Generalized Gaussian Distribution (GGD) can capture characteristics of distortion from the MSCN coefficients [2]. A zero mean GGD is given by

$$f(x; \alpha, \sigma^2) = \frac{\alpha}{2\beta\pi\left(\frac{1}{a}\right)} \exp\left(-\left(\frac{|x|}{\beta}\right)^\alpha\right) \tag{4}$$

$$\beta = \sigma \sqrt{\frac{\pi\left(\frac{1}{\alpha}\right)}{\pi\left(\frac{3}{\alpha}\right)}}$$

Where

and $\pi(\cdot)$ is the gamma function

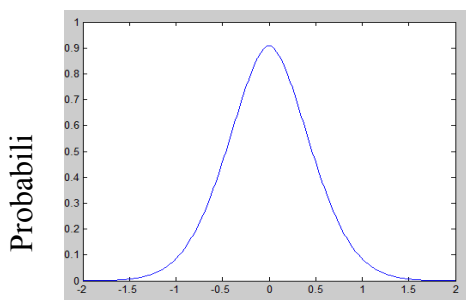


Fig. 3. The distribution of MSCN coefficients of Fig. 1

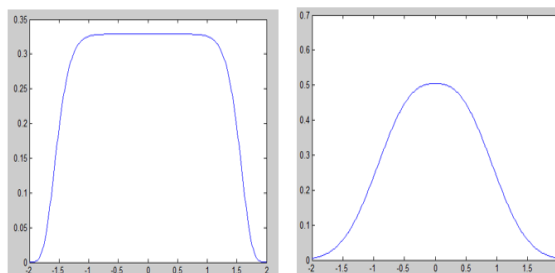


Fig. 4. The distribution of MSCN coefficients of Fig. 2.

Even though Fig. 1 and the second image in Fig. 2 does not show much difference in human perception, the MSCN coefficients clearly show the distortion difference using difference in shape.

The 'shape' of the distribution is controlled by shape parameter, α and variance by σ^2 . The shape parameter α also denotes the rate of decay: the smaller α , the more peaked is the distribution, and the larger α , the flatter is the distribution, so it is also called as the decay rate. For NSS images, there exists a statistical relationship between neighboring pixels and these dependencies get disturbed due to distortion. This is modeled by this by computing the pairwise products of MSCN coefficients in the neighborhood [2] along four orientations – horizontal (H), vertical (V), main-diagonal (D1) and secondary diagonal (D2) [6].

$$H(i,j) = \bar{I}(i,j) \bar{I}(i, j + 1) \tag{5}$$

$$V(i,j) = \bar{I}(i,j) \bar{I}(i + 1,j) \tag{6}$$

$$D1(i, j) = \bar{I}(i, j) \bar{I}(i + 1, j + 1) \tag{7}$$

$$D2(i, j) = \bar{I}(i, j) \bar{I}(i + 1, j - 1) \tag{8}$$

for $i \in \{1, 2 \dots M\}$ and $j \in \{1, 2 \dots N\}$.

The variations of these pairwise products are better represented using the commonly available Asymmetric Generalized Gaussian Distribution function (AGGD). To observe the variation, histograms of paired products along each of four orientations is plotted. This is done for a good quality image and for its distorted versions. The shape parameter and the scale parameters controls the 'shape' and the spread of the distribution on each side of the mode, respectively. The distribution skew is taken as the function of the left and right scale parameters [3]. This becomes a GGD if the skew value is same for left and right. AGGD is usually used to model heavy skewed distributions.



Fig. 5. An undistorted natural image and its MSCN coefficient

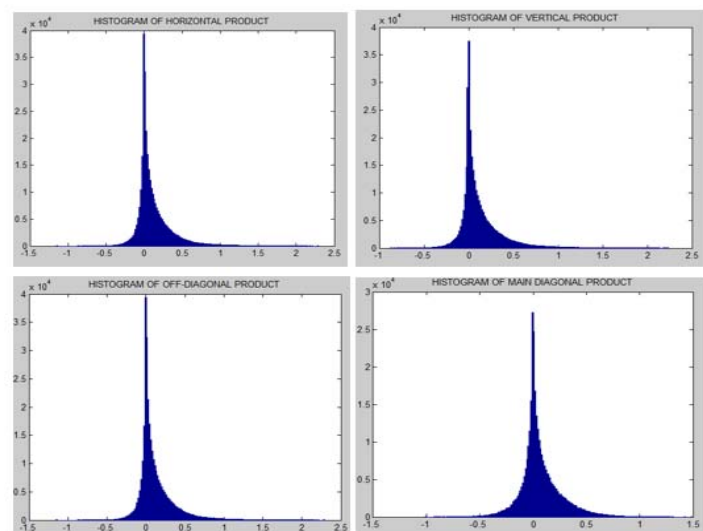


Fig. 6. The horizontal product, vertical product, main diagonal and off diagonal product values of the original image in fig 5.

TABLE I: The Set of Features

Feature ID	Feature Description	Computation Procedure
$f1-f2$	Shape and variance	By fitting Generalized Gaussian Distribution to MSCN Coefficients
$f3-f6$	Shape, mean, left variance, right variance	By fitting AGGD to <i>Horizontal</i> pairwise products
$f7-f10$	Shape, mean, left variance, right variance	By fitting AGGD to <i>Vertical</i> pairwise products
$f11-f14$	Shape, mean, left variance, right variance	By fitting AGGD to <i>Main diagonal</i> pairwise products
$f15-f18$	Shape, mean, left variance, right variance	By fitting AGGD to <i>Off diagonal</i> pairwise products

We extract 16 parameters (4 parameters/ orientation \times 4 orientations) for each paired product, thus giving the next set of features. The same is repeated at two scales - the original image scale, and at down sampled resolution. Increasing the number of scales beyond two did not give much performance increase. Thus 36 features (18 features in two scale) helped in learning and converting feature set to a quality score.

III. LEARNING PROCESS AND QUALITY EVALUATION

We have considered three learning approaches.

- The first approach is the General classification model, which involves constructing a general classification model from the training data consisting of image features and their quality scores. This approach does not take into account the type of distortion in the image.
- The second approach involves a lazy learner prediction scheme involving k nearest neighbours. The quality score of the test image is estimated as the average of quality scores of the k nearest neighbours in the feature space from the training data.
- The third approach is a distortion specific approach.

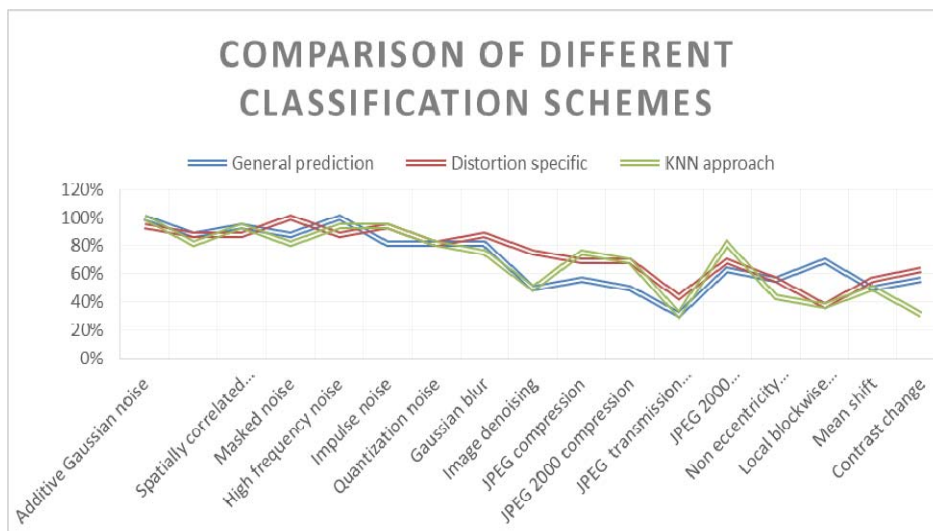


Fig. 7. The comparison between the accuracies of 17 different distortion types available with TID 2008

In this approach, we first use a classification model to yield a vector that indicates the probability of the test image belonging to each distortion type. A regressor is learnt for each distortion type and quality scores are obtained from each regressor [12]. The resultant score is taken to be the dot product of the probability vector and the corresponding regressor scores.

TABLE II. Accuracy of Three Methods

Method	Accuracy
General prediction	70.22%
Distortion specific method	73.89%
Lazy learner approach (KNN)	68.75%

IV. DISCUSSION

The opinion aware image quality assessment system was tested on LIVE II [14] and in TID 2008 [8]. The below illustrated test results are based on TID 2008 dataset. A collection of 272 distorted images from the TID 2008 database was used for training. The test set had a uniform mix of distortions. Table II shows the summary of accuracies of classification under the three adopted schemes. This shows that distortion specific method is good in categorizing a distorted image into low quality or degraded quality set. The detailed values of accuracies of quality prediction are shown in figure 7. The tables show the percentage of accuracy we obtained while comparing our results with the MOS value available at the TID dataset. Here a comparison of three quality evaluation approaches is shown against the 17 different distortion types available in the TID 2008 dataset.

The Mean Opinion Score available at the data set is taken as the subjective basis for comparison of the obtained objective result.

Clearly, the three methods exhibit different accuracies for different distortions. If the type of possible distortion is known, then the best option to predict the quality is to select the most suitable one from the table. i.e. for high frequency noise, General prediction gives the accurate result. For JPEG 2000 transmission error, KNN approach is the best one. If the type of distortion is unknown, which happens mostly, and then distortion specific method is proficient in giving the most accurate opinion. Distortion specific classification method gave a prediction score around 74 and as we consider it as the good performing classification scheme.

V. CONCLUSION

In this paper, we discussed about finding the feature values of an image for opinion-aware quality evaluation method and three ways of interpreting these features to predict the quality of that image. Objective analysis result can be taken into account only by verifying its comparison with subjective assessment. Hence the MOS scores available at the dataset are taken for calculating the accuracy of the method. The results obtained were comparable with human judgment. Also, each of the three methods exposed areas where it can be applied. The efficiency of the method depends greatly on the feature set extracted from the image. Here, we found that all three methods give promising results and among them, distortion specific method acts as the noticeable one for opinion aware, blind Quality Assurance Method.

REFERENCES

- [1] Xinbo Gao, Fei Gao, Dacheng Tao, and Xuelong Li, "Universal Blind Image Quality Assessment Metrics Via Natural Scene Statistics and Multiple Kernel Learning," IEEE Transactions on Neural Networks and Learning Systems, vol. 24, no. 12, pp.2013-2026, July 2013.
- [2] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," IEEE Transactions on Image Processing, vol. 21, no. 12, pp.4695-4078, Dec 2012.
- [3] N.E Lasmari and Y.Stitou, "Multiscale skewed heavy tailed model for texture analysis," in Proc IEEE international Conference on Image Processing, pp. 2281-2284, Nov 2009.
- [4] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," IEEE Transactions on Image Processing, vol. 20, no. 12, pp. 3350–3364, Dec 2011.
- [5] Michele A. Saad, Alan C. Bovik, and Christophe Charrier, "A DCT Statistics-Based Blind Image Quality Index," IEEE Signal Processing Letters, vol. 17, no. 6, pp.583-586, June 2010.
- [6] K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," IEEE Signal Process. Letters, vol. 17, no. 5, pp. 513–516, May 2010.
- [7] H. Tang, N. Joshi, and A. Kapoor, "Learning a blind measure of perceptual image quality," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 305–312, June 2011.
- [8] TID2008 URL:<http://ponomarenko.info/tid2008.htm>
- [9] Weiling Hou, Xinbo Gao, Dacheng Tao and Xuelong Li, "Blind Image Quality Assessment via Deep Learning," IEEE transactions on neural networks and learning systems, vol. 26, no. 6, pp. 1275 – 1286, June 2015.
- [10] L. He, D. Tao, X. Li, and X. Gao, "Sparse representation for blind image quality assessment," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 1146–1153, June 2012.
- [11] Wilson S. Geisler, "Visual Perception and the Statistical Properties of Natural Scenes," Annual review of Psychology, Vol. 59: 167-192, Jan 2008.
- [12] C. Li, A. C. Bovik, and X. Wu, "Blind image quality assessment using a general regression neural network," IEEE Transactions on Neural Networks, vol. 22, no. 5, pp. 793–799, May 2011.
- [13] Tong Li, Junping Zhang and Ye Zhang "Classification of Hyper Spectral image based on Deep belief networks," IEEE International Conference on Image Processing, pp. 5132 – 5136, Oct 2014.
- [14] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, (2006), "LIVE Image Quality Assessment Database Release 2" [Online]. Available: <http://live.ece.utexas.edu/research/quality>
- [15] R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," IEEE Transactions on Image Processing, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
- [16] Anish Mittal, Rajiv Soundararajan, Gautam S. Muralidhar, Alan C. Bovik, and Joydeep Ghosh, "Blind Image Quality Assessment without Training on Human Opinion Scores," 2013.
- [17] G. Hinton, S. Osindero and Y. Teh, 6, "A fast learning algorithm for deep belief nets," Neural Computation, vol. 18, no. 7, pp. 1527–1554, Jul 2006.
- [18] Anish Mittal, Rajiv Soundararajan and Alan C. Bovik, "Making a 'Completely Blind' Image Quality Analyzer," IEEE Signal Processing Letters, vol.20, Issue 3, pp.209-212, March 2013.
- [19] Xinbo Gao, Fei Gao, Dacheng Tao and Xuelong Li, "Universal Blind Image Quality Assessment Metrics Via Natural Scene Statistics and Multiple Kernel Learning", IEEE transactions on neural networks and learning systems, vol. 24, no. 12, Dec 2013.
- [20] D. L. Ruderman, "The statistics of natural images," Network Computation in Neural Systems, vol. 5, no. 4, pp. 517–548, 1996.

AUTHOR PROFILE

Author1: Gopika S is from Kerala, India .She did her B.Tech in Information Technology from Government Engineering College, Trivandrum in 2004. She did M.E in Computer Science from Anna University in 2006 and was a rank holder. Presently she is pursuing Ph.D at SRM University, Chennai. She has worked as Assistant Professor at Rajagiri Engineering college, Kerala and presently she is a faculty member at Kristu Jayanti College, Bangalore. Her area of research work is in Image Processing. She has published papers in International Journals and International Conferences in the area of Image Processing with more than a dozen to count. Email: gopika.s@gmail.com

Author2 : Dr.D. Malathi, was born on 16th March 1967 at Trichy, India. She obtained her Ph.D degree in the year 2010 from Anna University in the area of Artificial Neural Networks under the guidance of Dr. N. Gunasekaran, HOD/Prof (Retd.), SECE at Anna University, Chennai. She received her Master's degree in the year 1998 from University of Madras, India in the area of Computer Science and Engineering. She got her Bachelor's degree from the Institution of Engineers, India in the year 1991. Her research interests include Artificial Neural Networks, Image Processing, Pattern Recognition and Signal Processing. She has joined teaching profession in the year 1994. Currently she is working as Professor, Department of Computer Science and Engineering, S.R.M University, Tamil Nadu, India. She has guided many B.Tech and M.Tech Projects and guiding Ph.D Research Scholars. She has published many papers in National and International Conferences and International Journals. E-mail: mala_kam@yahoo.com, dmalathi@ktr.srmuniv.ac.in. She is a member of IE, IET, ACM, Indian Science Congress and Life Member of ISTE.

Author 3: J. D. Dorathi Jayaseeli was born on 27th May 1970 at Cuddalore, Tamilnadu. She received her B.E. degree in Electronics and Communication Engineering from Madras University, Chennai, in 1991 and obtained her M.E. degree in Computer Science and Engineering from Madras University, Chennai, in 2000 (with University Rank). She is pursuing her Ph.D. degree at SRM University. Presently she is an Assistant Professor (S.G) at SRM University, Chennai, India. She was the Head of the Department, of Information Science Engineering, P.A. College of Engineering, VTU, Mangalore, Karnataka from 2006 to 2014. Her fields of interest are Image Processing, Pattern Recognition and GIS. She has published papers in International Journals. She is a member of IET and Indian Science Congress.