International Journal of Emerging Engineering Research and Technology
Volume 2, Issue 2, May 2014, PP 192-197

Spatial Domain No Reference Image Quality Assessment of JPEG Compressed Images

Ajinkya M. Pund[1], Charudatta V.Kulkarni[2]
[1]Department of Electronics & Tele-Communication, MIT College of Engineering, Pune, India
[2]Department of Electronics & Tele-Communication, MIT College of Engineering, Pune, India

Abstract: *Human observers can easily assess the quality of a distorted image without examining the original image as a reference. By contrast, designing objective No-Reference (NR) quality measurement algorithms is a very difficult task. Currently, NR quality assessment is feasible only when prior knowledge about the types of image distortion is available. This paper aims to develop NR quality measurement algorithm for JPEG compressed images. For this purpose, standard LIVE image database is used. We try to show that Peak Signal to-Noise Ratio (PSNR), which requires the reference images, is a poor indicator of subjective quality. Therefore, tuning an NR measurement model towards PSNR is not an appropriate approach in designing NR quality metrics. Furthermore, we tries to develop a computational and memory efficient NR quality assessment models for JPEG compressed images. Subjective test results are used to train the model, which achieves good quality prediction performance.*

Keywords: *No reference (NR), peak signal to noise ratio (PSNR), quality assessment, subjective test.*

1. Introduction

It is an ever increasing requirement to send more multimedia data over tighter bandwidth which has been driven to develop advanced compression technology. Due to the advanced development of different image compression techniques and processing systems, there is a very big concern about the levels of image quality both for providers and users in many image processing applications from compression to printing. Images are subject to distortions during acquisition, compression, transmission, processing, and reproduction. To maintain, control, and enhance the quality of images, it is important for image acquisition, management, communication, and processing systems to be able to identify and quantify image quality degradations. The development of effective automatic image quality assessment systems is a necessary goal for this purpose. Yet, until recently, the field of image quality assessment has remained in a nascent state awaiting new models of human vision and of natural image structure and statistics before meaningful progress could be made.

Since human beings are the ultimate receivers in most image-processing applications, the most reliable way of assessing the quality of an image is by subjective evaluation. Indeed, the mean opinion score (MOS), a subjective quality measure requiring the services of a number of human observers, has been long regarded as the best method of image quality measurement. However, the MOS method is expensive, and it is usually too slow to be useful in real-world applications. The goal of *objective* image quality assessment research is to design computational models that can predict perceived image quality accurately and automatically. We use the term *predict* here, since the numerical measures of quality that an algorithm provides are useless unless they correlate well with human subjectivity. In other words, the algorithm should predict the quality of an image that an average human observer will report. Clearly, the successful development of such objective image quality measures has great potential in a wide range of application environments [1].

First, they can be used to *monitor* image quality in quality control systems. For example, an image acquisition system can use a quality metric to monitor and automatically adjust itself to obtain the best quality image data. A network video server can examine the quality of the digital video transmitted on the network to control and allocate streaming resources. In light of the recent gigantic growth of Internet video sources, this application is quite important.

Second, they can be employed to *benchmark* image processing systems and algorithms. For instance, if a number of image denoising and restoration algorithms are available to enhance the quality of images captured using digital cameras, then a quality metric can be deployed to determine which of them provides the best quality results.

Third, they can be embedded into image-processing and transmission systems to *optimize* the systems and the parameter settings. For example, in a visual communication system, an image quality measure can assist in the optimal design of the pre filtering and bit assignment algorithms at the encoder and of optimal reconstruction, error concealment, and post filtering algorithms at the decoder.

In the design and selection of image quality assessment methods, there is often a trade-off between accuracy and complexity, depending on the application scenario. For example, if there were an objective system that could completely simulate all relevant aspects of the human visual system (HVS), including its built-in knowledge of the environment, then it should be able to supply precise predictions of image quality. However, our knowledge of the HVS and our models of the environment remain limited in their sophistication. As we increase our knowledge in these domains, then it is to be expected that image quality assessment systems that come very close to human performance will be developed.

However, it is possible that future quality assessment systems that include such knowledge-based sophistications might require complex implementations, making them cumbersome for inclusion in image-processing algorithms and systems. Yet, it is also possible that elegant solutions will be found that provide superior performance with simple and easily implemented processing steps. Indeed, in later chapters we will describe some systems of this type that provide superior performance relative to previous technologies.

Historically, methods for image quality assessment have mostly been based on simple mathematical measures such as the *mean squared error* (MSE). This is largely because of a lack of knowledge regarding both the HVS and the structure and statistics of natural images. It is also owing to the analytic and computational simplicity of these measures, which makes them convenient in the context of design optimization.

However, the *predictive* performance of such systems relative to subjective human quality assessment has generally been quite poor. Indeed, while these methods for quality assessment have found considerable use as analytic metrics for theoretical algorithm design, they have long been considered as rather weak for assessing the quality of real images, processed or otherwise. Indeed, the field of image quality assessment, until the last decade, remained in a largely moribund state. Owing to a lack of driving forces in the form of new models for human visual perception of images, or of image formation, natural image structure, and natural scene statistics, research into image quality assessment was nearly non-existent, a sort of Rodney Dangerfield of vision and image engineering.

2. Classification of Objective Image Quality Measurement Techniques

Any image quality assessment method that aspires to perform at a level comparable to the average human observer must certainly overcome the drawbacks of the MSE and other *lp* norms. In fact, many image quality measures have been proposed over the past few decades. Although it is difficult to classify all of these methods into "crisp" categories, we believe that a rough classification can help sort out the fundamental ideas and facilitate real-world application as well as future research. In general, three types of knowledge can be used for the design of image quality measure: knowledge about the "original image," knowledge about the distortion process, and knowledge about the HVS. Accordingly, our classification is based on three different criteria.

2.1 The Full Reference Image Quality Assessment

The first criterion to classify objective image quality measures is the availability of an "original image," which is considered to be distortion-free or perfect quality, and may be used as a reference in evaluating a distorted image. Most of the proposed objective quality measures in the literature assume that the undistorted reference image exists and is fully available. Although "image quality" is frequently used for historical reasons, the more precise term for this type of metric would be image similarity or fidelity measurement, or *full reference* (FR) image quality assessment [1].

The vast majority of IQA algorithms are so-called full reference algorithms, which take as input both a distorted image and a reference image and yield as output an estimate of the quality of the distorted image relative to the reference. The simplest approach to full-reference (FR) IQA is to measure local pixel wise differences and then to collapse these local measurements into a scalar which represents the overall quality difference, for example, the mean-squared error (MSE) or peak signal-to-noise ratio (PSNR), often measured in different domains. More complete FR IQA algorithms have employed a wide variety of approaches ranging from estimating quality based on models of the HVS to estimating quality based on image structure to estimating quality by using various statistical and information-theoretic-based approaches and many other techniques.

2.2 Reduced Reference Image Quality Assessment

In the second type of image quality assessment method, the reference image is not fully available. Instead, certain features are extracted from the reference image and employed by the quality assessment system as side information to help evaluate the quality of the distorted image. This is referred to as *reduced-reference* (RR) image quality assessment. The idea of RR quality assessment was first proposed as a means to track the degree of visual quality degradation of video data transmitted through complex communication networks.

Reduced-reference (RR) IQA methods provide a solution for cases in which the reference image is not fully accessible. Methods of this type generally operate by extracting a minimal set of parameters from the reference image, parameters which are later used with the distorted image to estimate quality. An important question in RR research is how to determine effective parameters for the IQA task [3].

2.3 No Reference Image Quality Assessment

In many practical applications, an image quality assessment system does not have access to the reference images. Therefore, it is desirable to develop measurement approaches that can evaluate image quality blindly. Blind or *no-reference* (NR) image quality assessment turns out to be a very difficult task, although human observers usually can effectively and reliably assess the quality of distorted images without using any reference at all. The reason for this is probably that the human brain holds a lot of knowledge about what images should, or should not, look like [2].

Before the widespread use of compression systems, most NR distortion measurements were aimed at measuring the distortions inherent in acquisition or display systems, such as the blur introduced by optics of the capture or display devices, sensor noise, etc. But since the evolution of high-quality image and video acquisition and display devices, and the widespread use of digital images and videos, the emphasis has shifted toward distortions that are introduced during the stages of compression or transmission. Indeed, the experiments conducted by VQEG in Phase-I and Phase-II of the testing consisted of videos distorted mainly with compression artifacts and transmission errors [3].

3. QUALITY ASSESSMENT OF JPEG COMPRESSED IMAGES

Block-based image and video compression usually involve block partitioning of the image prior to subsequent processing steps. For example, in JPEG compression, the image being encoded is first partitioned into 8 × 8 blocks, and then a local discrete cosine transform (DCT) is applied to the pixels in each block. Each DCT coefficient in each block is independently quantized prior to an entropy coding procedure that further eliminates redundancies. At low bit rates, the most prominent types of artifacts that are created by such block based image compression algorithms are *inter block blurring* (blurring within blocks) and *blocking artifacts* across block boundaries. The blurring effect is due to the loss of high frequencies during quantization. Since natural images typically have much lower energy at high frequencies, the high-frequency DCT coefficients have lower magnitudes, and the process of quantization tends to zero these coefficients [4]. Consequently, the decoded image loses high-frequency components, which might be visually noticeable, and which is seen as blurring within the blocks. Blocking artifacts manifest as visually apparent discontinuities across block boundaries. This is a consequence of the independent quantization of each block. As a result, the decoded image may exhibit regularly spaced horizontal and vertical edges.

3.1 Spatial Domain Feature Extraction

One effective way to examine both the blurring and blocking effects is to transform the signal into the frequency domain [6]. We denote the test image signal as $x(m, n)$ for

$m \in [1, M]$ and $n \in [1, N]$, and calculate a differencing signal along each horizontal line:

$$d_h(i, j) = x(i, j) - x(i, j-1) \text{ for } j \in [1, N-1]. \tag{1}$$

A disadvantage of the frequency domain method is the involvement of the Fast Fourier Transform (FFT), which has to be calculated many times for each image, and is therefore expensive. FFT also requires more storage space because it cannot be computed locally. Spatial domain technique is designed for JPEG compressed images and based on a straightforward feature extraction process, where the features directly reflect local blockiness and activity of the image being evaluated [6]. The features are calculated horizontally

and then vertically. First, the blockiness is estimated as the average differences across block boundaries:

$$D_h = \frac{1}{M\left(\left\lfloor \frac{N}{B} \right\rfloor - 1\right)} \sum_{i=1}^{M} \sum_{j=1}^{\left\lfloor \frac{N}{B} \right\rfloor - 1} \left| d_h\left(i, B_j\right) \right| \tag{2}$$

Where, $\lfloor a \rfloor$ denotes the largest integer smaller than a. B is block size for jpeg compression (generally B=8) and the horizontal block boundaries occur between B_jth and B_{j+1}th pixels in each horizontal line.

Second, we estimate the activity of the image signal. Although blurring is difficult to be evaluated without the reference image, it causes the reduction of signal activity, and combining the blockiness and activity measures gives more insight into the relative blur in the image. The activity is measured using two factors [6]. The first is the average absolute difference between in-block image samples:

$$A_h = \frac{1}{B-1}\left[\frac{B}{M(N-1)} \sum_{i=1}^{M} \sum_{j=1}^{N-1} \left| d_h(i,j) \right| - D_n\right] \tag{3}$$

The second activity measure is the zero-crossing (ZC) rate. We define for $n \in [1, N\text{-}2]$,

$$z_h(i,j) = \begin{cases} 1 & \text{there is a horizontal } Z \text{ at } d_h(i,j) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$Z_h = \frac{1}{M(N-2)} \sum_{i=1}^{M} \sum_{j=1}^{N-2} z_h\,(i,j) \tag{5}$$

Using similar methods, the vertical features of *Dv*, *Av*, and *Zv* can be computed. The overall features are simply the average of the corresponding horizontal and vertical features:

$$D = \frac{D_h + D_v}{2} \tag{6}$$

$$A = \frac{A_h + A_v}{2} \tag{7}$$

$$Z = \frac{Z_h + Z_v}{2} \tag{8}$$

There are many different ways to combine the features to constitute a quality assessment model. One method we find that gives good prediction performance is given by

$$S = \alpha + \beta D^{\gamma_1} A^{\gamma_2} Z^{\gamma_3} \tag{9}$$

Where α, $\beta, \gamma_1, \gamma_2$ and γ_3 are the model parameters that must be estimated with the subjective test data. The nonlinear regression routine "nlinfit" in the Matlab statistics toolbox is used to find the best parameters for equation (9). It is important that the model is not over trained.

3.2 Algorithm Flow

Complete design flow of above algorithm is given in Fig.1 differencing signal and average difference provides details of blockiness in image. Average absolute difference and zero crossing rate provides details of blockiness in image.

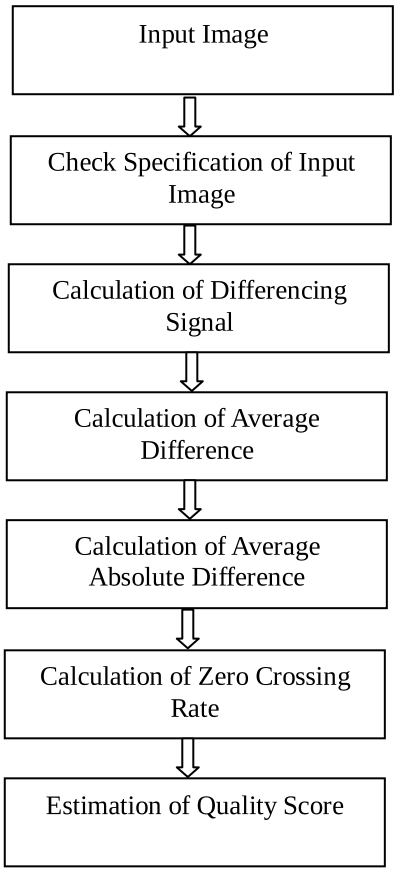

Figure 1. *Design Flow*

4. Live Image Database and Experimental Results

In the Texas' database, the subjective test was conducted on 8 bits/pixel gray level images. There are 120 test images in the database. Thirty of them are original images, which are randomly divided into two groups with 15 images in each group [6]. The two groups of images are shown in Figs. 2 and 3, respectively. The rest of the test images are JPEG-compressed. The quality factors are selected randomly between 5 and 100, and the resulting bit rates range from 0.2 to 1.7 bits/pixel. Fifty three subjects were shown the database; most of them were college students. The subjects were asked to assign each image a quality score between 1 and 10 (10 represents the best quality and 1 the worst). The 53 scores of each image were averaged to a final Mean Opinion Score (MOS) of the image.

Figure 2. *Group I Images*

Figure 3. *Group II Images*

Model prediction results using both group images is given in Figure 4 shown below.

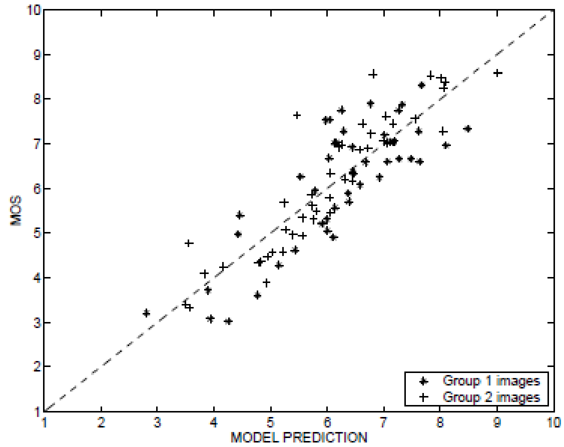

Figure 4. *Model prediction using both groups of images for as training images.*

Figure 5 shows prediction results for group I images as training images.

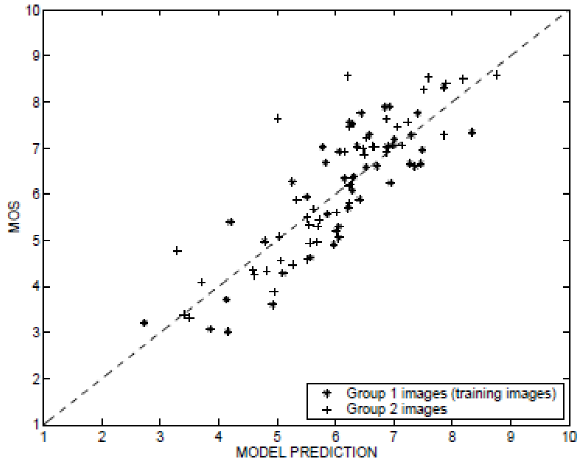

Figure 5. *Model prediction using group I images as training images*

5. Conclusion

We demonstrated spatial domain no reference image quality assessment of jpeg compressed images. Features described in this paper effectively captures artifacts introduced by jpeg image compression. The method is computationally efficient since no complicated transforms are required and is memory efficient too as it doesn't requires storing images for computation.

References

[1] Z. Wang and A. C. Bovik, "Why Is Image Quality Assessment So Difficult?" IEEE Signal Processing Letters, 2002.

[2] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," IEEE Signal Processing Letters, vol. 17, no. 2, pp. 587–599, May 2010

[3] R. Ferzli and L. Karam, “A No-Reference Objective Image Sharpness Metric Based on the Notion of Just Noticeable Blur (JNB),” IEEE Transactions on Image Processing, vol. 18, no. 4, p. 717, 2009.

[4] H. R. Sheikh and A. C. Bovik, “Image information and visual quality,” IEEE Transactions on Image Processing, vol. 15, no. 2, pp. 430–444, 2006.

[5] H. R. Sheikh, A. C. Bovik, and L. K. Cormack, “No-reference quality assessment using natural scene statistics: JPEG 2000,” IEEE Transactions on Image Processing, vol. 14, no. 11, pp. 1918–1927, 2005.

[6] Z. Wang, H. R. Sheikh, and A. C. Bovik, “No-reference perceptual quality assessment of jpeg compressed images,” Proc. of IEEE Int. Conf. on Image Processing, vol. 1, pp. 477–480, 2002.

Author’s Biography

Ajinkya Manoharao Pund received the B.E. degree in Electronics and Tele-communication with from the Amravati University, Amravati, India in June 2008. He is pursuing M.E. in VLSI And Embedded Systems from University of Pune, Pune , India