New Objective Visual Quality Assessment of JPEG Compressed images

Mohammad K. Ibrahim and M. M. Al-Burikan

ibrahimm@ccse.kfupm.edu.sa, st962108@ccse.kfupm.edu.sa

ABSTRACT

The significant advances in image compression have not been matched by the same level of advancement in the objective assessment of visual quality. The most commonly used criteria is the Peak Signal-to-Noise-Ratio (PSNR). However, it has been very well known that the visual evaluation with PSNR does not correlate well with the results obtained with subjective test carried out by humans. In this paper, it is first shown that a local PSNR measure is a much better estimate of image distortion than a global PSNR measure. In addition, a new graphical distortion measure is introduced which takes into consideration the visibility threshold of human visual perception. The basic idea is to adapt the concept of non-uniform quantization used in DPCM coding to distinguish between the perceptually distinct transitions in luminance. The comparison of the new measure with the PSNR and subjective test is also presented using JPEG compressed images. It is shown that this graphical measure of distortion is more informative than a single value measure such as the PSNR.

1. INTRODUCTION

Objective measures of visual information and distortion in still images is crucial in the design and implementation of compression algorithms and systems. We would like to distinguish between the different types of measures:

1 measure of visual information, which can be used in the evaluation of the amount of visual information prior to compression.

2 measure of distortion in visual information which relates to the loss in the visual information

3 image activity measure which is simply an indication of how detailed and complex the image is. This is only a pointer to the amount of visual information but not a direct measure of this information.

In this paper we are concerned with the second type of measures. The applications of objective measures of distortion are:

1. they can be used to evaluate and test compression algorithms and image coders.
2. they can be used in rate control mechanisms that are now part of any image/video compression standard. (It is significant to note that for rate control mechanism, the visual quality assessment must not incur a significant overhead to the codec complexity. Hence the complexity of any objective measure is also an issue that need to be considered.)
3. to select pre-processing parameters to reduce the effect of artifacts due to compression.
4. the design of intelligent compression systems that match the type of compression technique to the image visual information

The overall majority of distortion measures use a single number (scalar) to point out the amount of degradation. The most commonly used scalar DMs are based on the Mean Square Error (MSE) that has been used extensively in the evaluation of image compression algorithms. Many researchers have reported the limitations of the MSE which does not correlate well with subjective tests. The main reason given for this is that the MSE does not take into consideration the characteristics of the Human Vision System (HVS) [1].

Several scalar DM have been proposed that incorporate a model of the HVS that have been reported to have a better correlation with subjective tests than the MSE [2]. They are generally based on a model of the Human visibility threshold. Evaluation of the human visibility threshold, and hence visual masking, has been carried out in both the spatial-domain as well as the frequency-domain. The two domains are complementary in nature [4]. There is no general agreement among psychophysicists about which domain is more suitable. Known psychological facts can be interpreted to support partially either of the two domains.

The work in [3] is based on the frequency-domain analysis of vision threshold. It is based on the hypothesis of independent channels which considers that masking occurs when two stimuli are closely coupled in the spectral domain, i.e. they must have similar spatial frequency and temporal frequency. This implies that stimulus can only mask another if both of them belong to the same channel [3]. Research has shown however that the adjacent bands affect each other.

Furthermore, these measures still use a single number to measure degradation, and are sensitive to parameters that control functions such as image segmentation [3]. In [1], it was pointed out that the human observer makes a decision on the (1) amount of distortion, (2) the type of distortion, and (3) the distribution of error. It is clear that no matter how sophisticated a scalar DM is, a single number will not be adequate to represent these decisions, and that more than a single number is needed [1].

Two such methods that use more than one value to represent distortion are the Probability Distribution (Histogram Analysis) of the compression error and Hosaka plots [1,2]. The values computed in these methods are usually represented as graphs. The performance of these two methods were compared in [2]. As indicated in [2], the Hosaka plots are sensitive to the variance threshold and the initial block size and that the choice of these parameters depend on the compression ratio, spatial frequency, and the type of impairment. As with regard to the Histogram Analysis (HA), the authors in [2] indicate that the HA of the compression error, which is obtained by subtracting the original image from the decompressed one, is not useful to identify the type of degradation.

The approach adopted in this paper is to study the effect of lossy compression on the impairments caused to image features. The most basic image feature is the spatial transition in luminance between a pixel and its neighbors. In this paper, the range of values of the spatial-transition of luminance is split into visually distinguishable (VD) bands based on the spatial masking of the HVS. This is discussed in more details in sections 3 &4. But first the evaluation of local histograms is assessed first.

2. LOCAL MEAN SQUARE ERROR

The most commonly used quantative measure is the PSNR

$$PSNR = 10 \log \frac{256^2}{\sum_{n=1}^{N}\sum_{m=1}^{N}(I(n,m) - \hat{I}(n,m))^2}$$

where I(n,m) is the original image and $\hat{I}(n,m)$ is a modified version. (It is worth pointing at this stage that the denominator of the PSNR is the Mean Square Error (MSE)). It has been known for some time that the PSNR does not correlate well with the perceptual quality of images as humans perceive it. In fact, in many cases, high PSNR is obtained for picture quality with annoying impairments and sometimes with objectionable quality. For this reason, evaluation of compression algorithms relies more on subjective tests as they give a more accurate measure of visual impairments caused by errors due to compression [4].

There are two important points that need to be noted with the regard to the PSNR:

1 it is a spatially-global measure of image distortion which is contrary to the fact that the human observers are also very sensitive to spatially-local distortion,
2 it is a measure based on a single value, which combines the effects of different types of distortion that can be observed by the human observer,
3 it does not take into account the visibility threshold of the human visual system despite the fact that humans are the intended customers.

To study the limitations of the PSNR as a global and measure, the PSNR values are studied for different regions of the same image. The distortion used in this study is that caused by JPEG compression with different compression rates. In this test two images are used, Lenna and Cameraman, which each is divided into 4x4 sub images. Figure 1 show the MSE values for each sub-image for the Cameraman and Lenna images respectively. The values clearly show that reducing the bit rates affects the distortion of the different sub-images in different ways. The distortion of some sub-images is degraded more than others. This result clearly indicates that a single global value is not adequate as a measure of distortion.

One interesting conclusion which this result supports is that some sub-images can tolerate higher compression rates than others. In other words different compression rates can be applied to different sub-images rather than using a unified compression rate for the whole image, which is usually limited by the worst affected areas. This also raises an important question. Although, some image exhibit higher MSE than others, but does that lead one to conclude that these sub-images exhibit higher perceptual distortion? This is addressed next.

3. VISIBILTY THRESHOLD AND PERCEPTIBAL DISTORION

The key to the measure of perceptible image quality is the visibility threshold of the human visual system which is used to mask the variations caused by compression. Clearly, as long as the impairments are kept below the visibility threshold, they should not be considered as impairments and consequently should not contribute to any measure of distortion. [2,3].

Many studies have been carried out on the visibility threshold of perturbation in the spatial domain in the area of visual psychophysics of the Human Visual

System (HVS). The findings of these studies have been used extensively in the design of non-uniform quantizers for DPCM coding [4]. It is known that the visibility threshold increases at supra-threshold luminance changes. Many psycho-visual experiments have been carried out to study the effect of spatial transition in luminance on the visibility threshold of perturbation in the spatial domain. It has become a known fact from these studies that there is a reduction in the visibility of perturbation near a sudden change in the luminance, such as edges. The findings of these studies have been used extensively in the design of non-uniform quantizers for DPCM coding [4]. Such quantizers are used in a new measure of distortion as shown in the next section.

4. MEASURE OF PERCEPTUAL DISTORION

It is clear that there is a need for a method to measure the errors in the image caused by compression with the condition that the resulting measurement should be *only* sensitive to impairments that are *visually perceived* by humans.

In this paper, this measurement is obtained as follows:

1. Divide the original image, I(n,m), and the reconstructed image, $\hat{I}(n,m)$, into K non-overlapping blocks, $I_k(n,m)$ and $\hat{I}_k(n,m)$, k=1,..,K, respectively of size L-by-L,
2. Find the mean of each block of the original image, $\bar{I}_k$, k=1,..,K,
3. Subtract the *mean* values of the blocks of the *original* image from corresponding block of both the original image and the reconstructed image, viz.

$$D_k(n,m) \;=\; I_k(n,m) \;-\; \bar{I}_k$$
$$\hat{D}_k(n,m) \;=\; \hat{I}_k(n,m) \;-\; \bar{I}_k$$

   It is significant to note that the mean of the original block is subtracted from the blocks of the reconstructed image and not its own mean.

4. Quantize $D_k(n,m)$ and $\hat{D}_k(n,m)$ using a non-uniform quantizers such as that in Table 1.
5. Calculate the energy of the perceptible noise for each block,

$$e_k(n,m) \;=\; (\, Q[D_k(n,m)] - Q[\hat{D}_k(n.m)]\,)^2$$

6. Calculate the error for each quantization level as follows,

$$E(l) \;=\; \sum_{k=1}^{L} \sum_{n,m \in B(l)} e_k(n,m)$$

   where *B(l)* is the set of co-ordinates (n,m) where $Q[D_k(n,m)]$ has the value of the *lth* quantization level.

It should be noted that the distortion measure E(l) is not a single value but an array of values, one for each quantization level. Also, this new measure takes into consideration the visibility threshold of the HVS. The effect of using the non-uniform quantizers are as follows. Although the residuals can have all possible values from –128 to 128 (for 8 bits/pixel in the original image), not all the values of the residuals can be treated as *distinct* from the viewpoint of the *Human Vision Systems*. This fact is the basis of the design of quantizers for DPCM image coders. As an example, the quantizer [4] in Table 1 demonstrates clearly how several values of the true differential values of an image are grouped together into one level. The effect of the quantizer is to group the differential values that can be distinguished visually due to their visibility threshold. This is the basis used to split the luminance-transition into perceptually distinct bands.

5. SIMULATION RESULTS AND CONCLUSION

The results shown in this section are again obtained using the image Lenna. Also, JPEG compression was used as the source of distortion. To obtain different levels of distortion different levels of compression was used. Table 2 show the degradation in the values of E(l) is different for different quantization levels. This result clearly indicate that different luminance transition are affected in different ways and hence the use of a single value to measure the distortion for all the different bands of the luminance transition is not adequate.

7. REFERENCES

[1] Eskicioglu, A. M, "Quality measurement for monochrome compressed images in the past 25 years" IEEE Conference on Image Processing, 2000, pp1907-1910.

[2] Eskicioglu, A. M., and Fisher, P. S., "Image quality measures and their performance", IEEE Transactions on Communications, Vol 43, No. 12, 1995, pp2959-2965.

[3] Hermiston, K.J., Booth, D. M., "Image quality measurement using integer wavelet transformations", IEEE Conference on Image Processing, 1999, pp293-297.

[4] A. N. Netravali and B. G. Haskell, Digital Pictures: Representation and Compression, Plenum Press, New York, 1988.

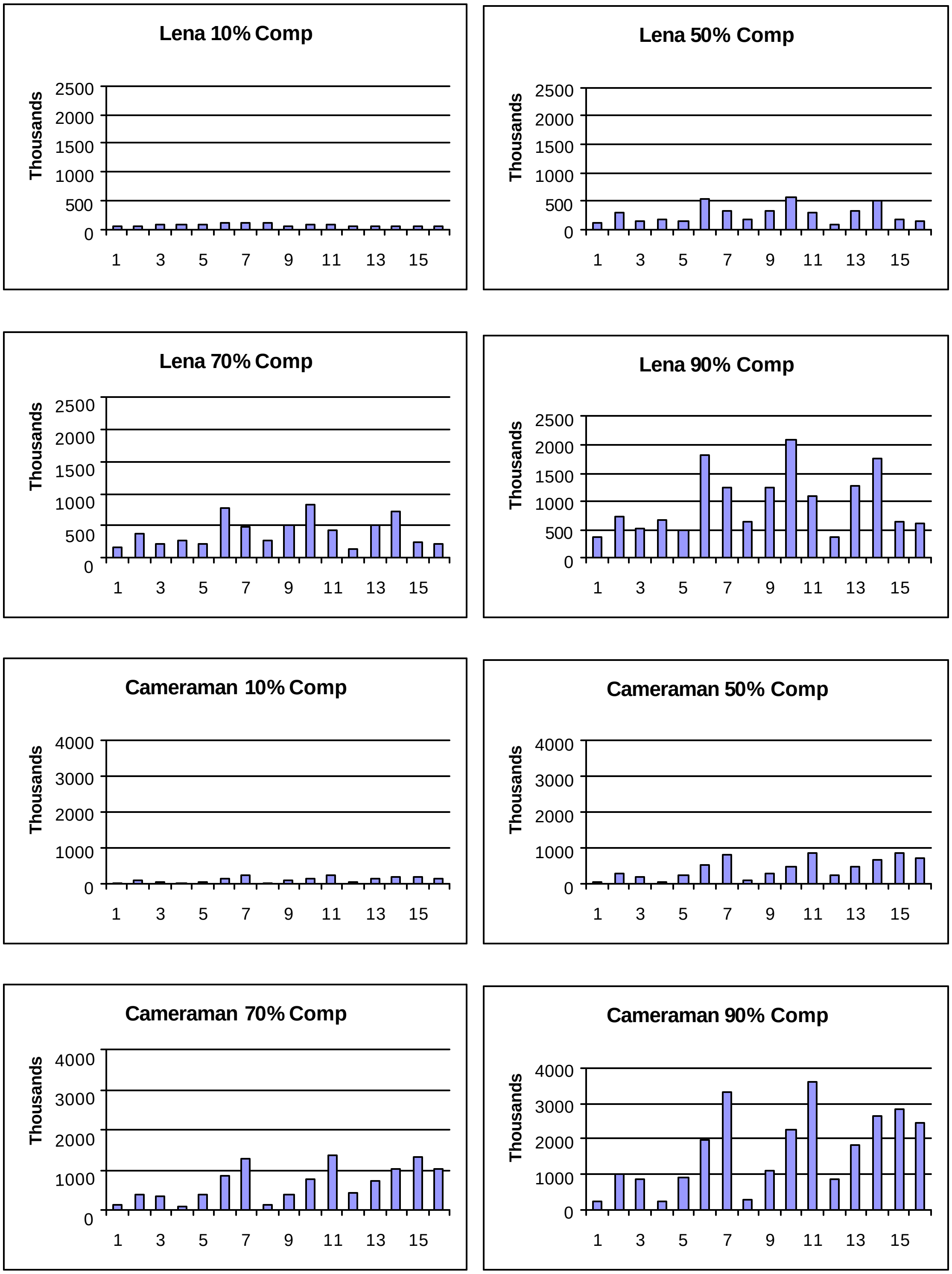

Figure 1: Local histograms for different compression rates for Lenna and Cameraman images.

Table 1 A symmetric non-uniform Quantizer [4] showing the range of spatial transition in luminance for positive values.

Quantization Level	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Quantization Step	-1 1	2 4	5 7	8 10	11 14	15 19	20 25	26 33	34 43	44 56	57 73	74 94	95 120	121 256

Table 2: $[H_{n:1}(l)/H_{2:1}(l)]$ for the different compression ratios for each of the 27 Quantization levels.

Quantization Level	**Comp. Ratio 4:1**	**Comp. Ratio 8:1**	**Comp. Ratio 12:1**	**Comp. Ratio 20:1**
-14	1	1	1	1
-13	5.3	9.2	14	25
-12	6	11	22	40
-11	7	15	23	48
-10	7.55	13	17	37
-9	6.9	11	14	26
-8	5.25	7.5	9.6	15
-7	4.3	5.6	6.9	11
-6	3	3.7	4	6
-5	2.25	2.6	2.7	3.8
-4	1.84	2	2.1	2.9
-3	1.6	1.7	1.7	2.4
-2	1.4	1.4	1.4	2
1	1.27	1.3	1.4	2.5
2	1.3	1.5	1.7	33
3	1.6	2	2.3	44
4	1.9	2.3	3	5.5
5	2.3	2.8	3.6	77
6	2.7	3.5	4.7	99
7	3.6	5	6.7	13
8	5	7.5	10	20
9	6.2	10	14	31
10	7	12	19	40
11	7.4	14	21	43
12	5.7	11	18	40
13	5	10	16	41
14	4.5	8	14	49
Ratio of $[MSE_{n:1}/MSE_{2:1}]$	**3**	**4**	**5**	**8**