



International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

www.iasir.net

No-Reference Perceptual Blockiness Estimation Method for JPEG Coded Images

Manika Rani Dey*, Hajera Siddiqua, Rayeen Sultana and Z. M. Parvez Sazzad
Department of Applied Physics, Electronics and Communication Engineering
University of Dhaka
Dhaka-1000, Bangladesh

*Department of Information & Communication Technology
Mawlana Bhasani Science & Technology University
Santosh, Tangail-1902, Bangladesh.

Abstract: In this paper, we propose an efficient no-reference perceptual blockiness estimation method based on local features and segmentation for JPEG coded images that can automatically quantify perceptual blocking artifacts. We believe that perceptual blockiness of any image is strongly dependent on local features such as edge, flat and texture. Therefore, edge, flat, and texture based blockiness are evaluated in this method. Experimental results on LIVE database show that proposed method has sufficient correlation with subjective evaluation scores.

Keywords: Perceptual Blockiness; JPEG; MOS; Segmentation; Texture.

I. Introduction

In image/video processing algorithms blocking artifact is one of the most annoying perceptual degradation caused by the block-based Discrete Cosine Transform (B-DCT) coding technique, like JPEG coding under low bit-rate conditions. Measuring perceptual blockiness plays an important role in many image processing applications. Digital image processing algorithms create several of image distortions which results visual quality degradation of an image. Lossy compression techniques are used to reduce bandwidth which degrades the quality due to the blockiness across the block boundary during quantization process. Consequently, blockiness measurement is required to evaluate the perceptual quality degradation, to monitor the encoding process and to remove artifact at the receiving end using de-blocking algorithms.

Natural scenes contain nonlinear dependencies and perceived image distortions are strongly dependant on local features [1], [2], [3]. As a result, perceived image distortion will be different with the same compression ratios on different scenes contain. Therefore, a new approach for the design of blockiness measures is proposed in this paper based on perceived differences of the local features of edge, flat and texture. Most of the proposed no-reference (NR) image quality assessments of block based Discrete Cosine Transform (DCT) coded images, like JPEG images, are used to quantify the major artifacts of blocking and blur in spatial domain [4]. The most common artifact in digital compressed images is blockiness [5], [6], which appears as small square block all over the image. Blockiness arises when a Discrete Cosine Transform (DCT) based compression algorithm uses a high compression ratio [7]. In DCT, data are presented as a sum of cosine functions of various magnitudes and frequencies. The DCT quantized an image by combining a large amount of image data into a small number of coefficients. In a DCT-based coder, an image is divided into several blocks of 8×8 pixels. Each of the blocks is then encoded, from the top left corner to keep only the important information of an image. As it processes forward from the upper left hand corner, the blocks are encoded with fewer and fewer bits and finally, the whole image is DCT quantized. During compression, the DCT coefficients of zeros are discarded without affecting the quality of the image. The compression rate and the quality of an image depend on the level of quantization of the DCT coefficients. From a technical point of view, blockiness occurs due to the discontinuity at block boundaries which is generated during block-based quantization of DCT coder. During compression the DCT transform to blocks of $M \times N$ (row $\times$ column) pixels; as a consequence, horizontal and vertical lines at the edges of the DCT boundaries (i.e., blockiness) are exhibited in compressed images. The presence of a periodic 8×8 edge structure in an image is called blocking artifact.

In this paper we intend to incorporate a new approach of perceived differences of the local features of images, such as edge, flat and texture, which will be more relevant measure of perceptual blockiness. Therefore perceived blockiness measurement method is proposed for JPEG images that produce the comparable result with those of subjective scores. We report that the performance of the method is sufficient and reliable.

II. Subjective Experiments

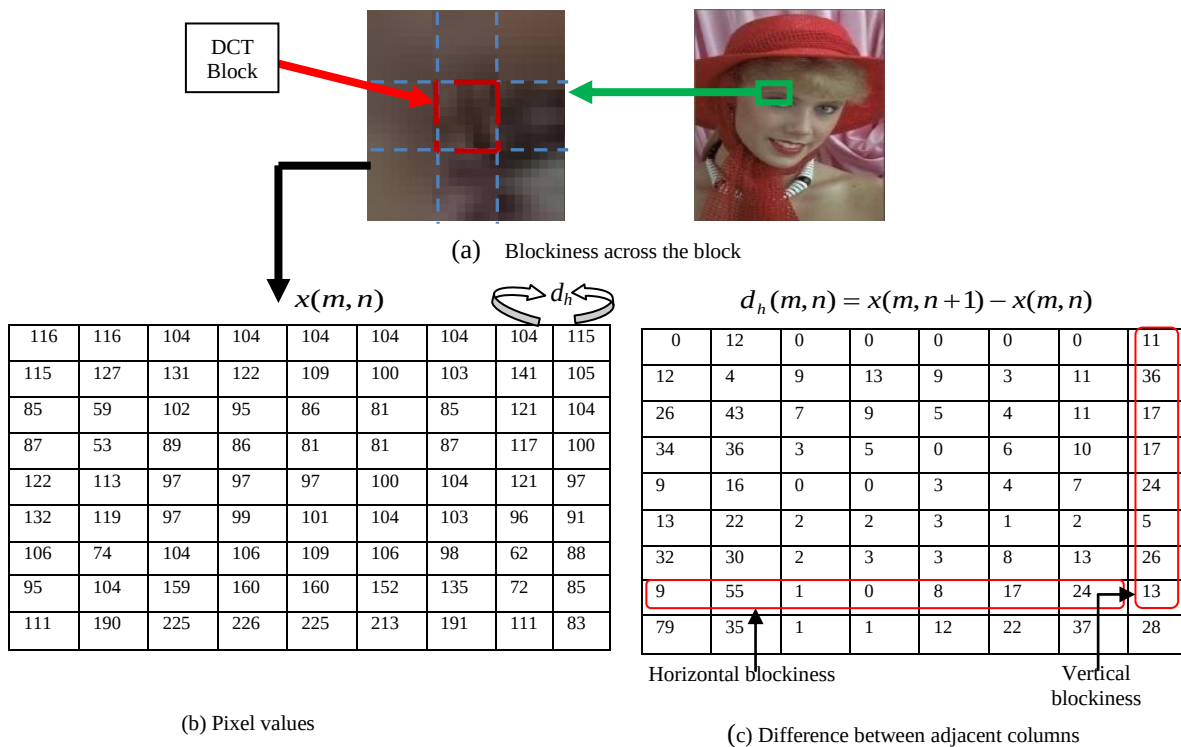
The most reliable way of assessing the quality of perceptual blockiness of a digital image/video is subjective evaluation, because human beings are the ultimate receivers in most applications. Such subjective quality is gauged by conducting a large scale human study where sizeable quantities of human observers are shown a series of visual stimuli whose quality they are asked to rate on a particular scale [8]. The mean score of these stimuli (after accounting for outlier subjects) is termed the mean opinion score (MOS) which represents the perceived quality of the stimuli. Such human assessment of quality is referred to as subjective quality assessment.

Twenty-nine high-resolution 24-bits/pixel RGB color images (typically 768×512) were compressed using JPEG with different compression ratios to yield a database of 233 images. The study was conducted in two sessions, with the original images included in both. Study 1 contained images img1.bmp to img116.bmp and study 2 contained images img117.bmp to img233.bmp [9]. The bit rates chosen such that the resulting distribution of quality scores for the compressed images were roughly uniform over the entire range. Each observer was shown the images randomly. Observers were asked to provide their perception of quality on a continuous linear scale that was divided into five equal regions marked with adjectives “Bad”, “Poor”, “Fair”, “Good” and “Excellent”. The scale was then converted into 1-100 linearly. The testing was done in two sessions with about half of the images in each session. No viewing distance restrictions were imposed, display device configurations were identical and ambient illumination levels were normal indoor illumination. Subjects were asked to comfortably view the images and make their judgments. Details of the experiment are discussed in [9].

III. Proposed Blockiness Prediction Model

In this method, firstly the average absolute difference between adjacent pixels within the block boundary is measured for a given image, secondly, a block based segmentation algorithm is used to classify edge, flat and texture blocks in the image, subsequently, each value of blockiness for edge, flat and texture blocks are averaged and updated with some parameters that are optimized by particle swarm optimization (PSO) algorithm. The proposed mathematical extracted features are calculated in spatial domain. The extracted feature measure is blockiness within the block (8×8) of the image. The feature is calculated horizontally and then vertically. For all calculations, only luminance component of color image is considered for simplicity.

Figure 1 Estimation of blockiness.



Horizontal blockiness is calculated by the average differences across the block boundaries using the following equation [4]:

$$B_h = \frac{1}{M([N/8]-1)} \sum_{i=1}^M \sum_{j=1}^{[N/8]-1} |d_h(i,8j)| \quad (1)$$

where it is denoted 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:

$$d_h(m, n) = x(m, n+1) - x(m, n) \quad (2)$$

Similarly, the vertical blockiness of B_v is calculated. Therefore, the overall blockiness B per image is given by:

$$B = \frac{B_h + B_v}{2} \quad (3)$$

The blockiness estimation approach is shown in Figure 1. Subsequently, a block based segmentation algorithm is applied to classify each block (8×8) of an image into either edge, flat, or texture blocks, and then calculated the blockiness of these blocks. The average blockiness value of the edge, flat, and texture blocks are calculated by

$$B_e = \frac{1}{N_e} \sum_{n=1}^{N_e} B_{en}, B_f = \frac{1}{N_f} \sum_{n=1}^{N_f} B_{fn}, B_t = \frac{1}{N_t} \sum_{n=1}^{N_t} B_{tn} \quad (4)$$

where B_e , B_f and B_t are respectively the average blockiness value of edge, flat, and texture blocks of the image. And also N_e , N_f , and N_t are respectively the number of edge, flat, and texture blocks of the image.

The features are combined by the following equation:

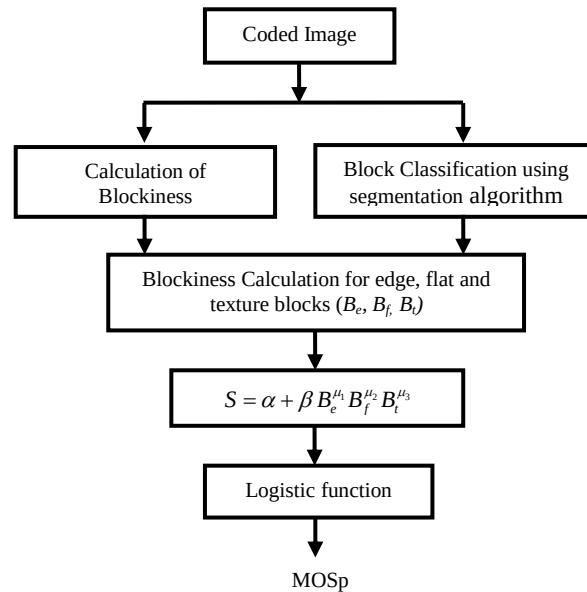
$$S = \alpha + \beta B_e^{\mu_1} B_f^{\mu_2} B_t^{\mu_3} \quad (5)$$

where α , β , μ_1 , μ_2 , and μ_3 are the method's parameters. The method parameters are optimized by the PSO algorithm with the subjective test data. The method has not taken into account the nonlinearity between the human perception and the physical feature; therefore we consider a logistic function as the nonlinear property. Finally, the obtained MOS prediction score (MOS_p) is derived from the following equation.

$$MOS_p = \frac{99}{1 + \exp[-1.0217(S - 50)]} + 1 \quad (6)$$

The block diagram of the proposed method is shown in Figure 2

Figure 2 Proposed perceptual blockiness estimation model.



IV. Block Classification and Segmentation

Since perceived image distortions are strongly dependent on the local features of edge, flat and texture areas, we need to consider different weighting on these three areas for perceived blockiness detection. In order to classify

edge, flat, and texture area of an image we consider a block based segmentation algorithm which was proposed in [2]. Suppose that an $M \times N$ rectangular image is divided into 8×8 non overlapping small square blocks. Let n_e , n_f and n_t are respectively the number of edge, flat, and texture pixels per (8×8) block within the image.

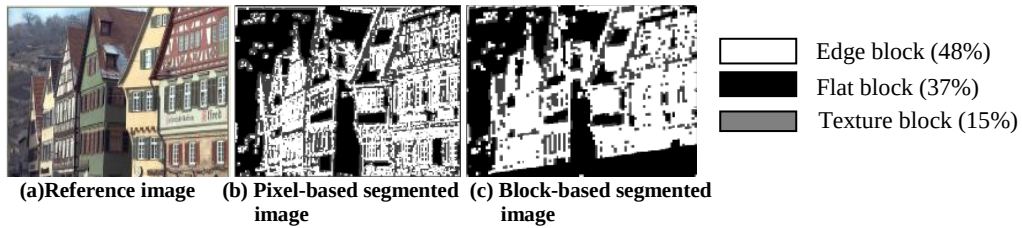
$$Sum = n_e + n_f + n_t \quad (7)$$

where “Sum” is the total number of pixels per block. The following is the block-based segmentation algorithm:

if $(\frac{n_e}{Sum} > th_e)$ then the block is “edge” else if $(\frac{n_f}{Sum - n_e} > th_f)$ then the block is “flat” else the block is “texture”

where “ th_e ”, and “ th_f ” are the algorithmic thresholds. In this segmentation algorithm, a block is considered as “edge block” if number of edge pixels are around thirty one percent of all pixels of the block. On the other hand, if number of flat pixels is around thirty four percent pixels of all flat and texture pixels of a block we consider the block is “flat block”. Otherwise the block is defined as “texture block”. Figure 3 is showing the percentage of edge, flat and texture blocks of a block segmented image.

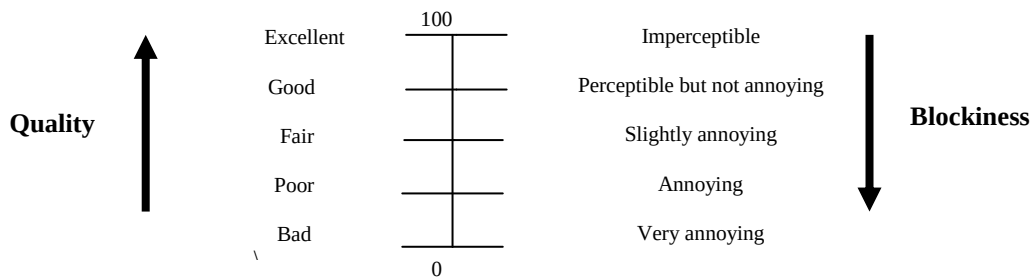
Figure 3 Reference image and its pixel and block based segmented images.



V. Results and Performance Evaluation

Target of our work is to detect perceptual blockiness automatically. However, the databases that are used to verify the performances of the proposed blockiness detection are subjective image quality assessment type which is evaluated by mean opinion score (MOS). It is recognized that image quality assessment indirectly represents perceived blockiness specifically in JPEG coded images. As increasing perceptual blockiness of an image degrades perceptual quality of the image, so the relationship between them is inversely proportional (see Figure 4). Therefore, in order to verify our perceived blockiness detection method proposed method uses the subjective scores (MOSSs) of image quality assessment.

Figure 4 Perceived image quality scale with respect to perceptual blockiness.



We consider the followings evaluation performance of our proposed method:

- Threshold estimation of the block-based segmentation algorithm.
- Performance evaluation according to VQEG.
- Comparison between PSNR and the proposed method
- Comparison between mathematical blockiness and the proposed perceived blockiness.

A. Threshold estimation of the block based segmentation algorithm

We have considered LIVE databases to evaluate and verify the best suitable algorithmic thresholds (“ th_e ” and “ th_f ”) for perceived blockiness detection. In order to use the database we divide the database into two parts for

training and testing. There is no overlapping between training and testing. The first part which we have considered as training database consists of 116 images. And the second part consists of 117 images that we considered as testing database. As verification, we can compare the performances of the algorithm using different sets of thresholds' values. Since the target of the segmentation is used to improve the perceived blockiness detection (i.e. inverse of image quality prediction), we have estimated all model parameters (α , β , μ^1 , μ^2 and μ^3) for each set of threshold values separately by the PSO algorithm with the subjective test data. Here, we have calculated the average absolute error (AAE) and maximum error (MAX) between MOS and MOSp for all training and testing images that are shown in Tables I and II. The optimum thresholds are estimated by comparing the performance of the algorithm with the different sets of thresholds values based on the AAE, and MAX. Therefore, in all cases (see Tables I and II), the performance of the segmentation algorithm is sufficient with the thresholds values of " th_e " = 0.32 and " th_f " = 0.34. The obtained method parameters by the PSO optimization algorithm for the thresholds values of " th_e " = 0.32 and " th_f " = 0.34 of the training database (MOS scale, 1-100) are shown in Table III.

Table I Thresholds performances based on AAE and MAX (training).

th_e	th_f	AAE	MAX
0.23	0.26	11.17	30.58
0.26	0.29	10.70	26.88
0.29	0.32	8.32	22.84
0.32	0.34	8.08	21.59
0.35	0.36	8.78	32.85

Table II Thresholds performances based on AAE and MAX (testing).

th_e	th_f	AAE	MAX
0.23	0.26	10.19	30.58
0.26	0.29	9.95	26.88
0.29	0.32	8.61	22.84
0.32	0.34	8.18	21.59
0.35	0.36	9.45	32.86

Table III Model parameters and weighting factors for MOS scale, 1-100 (Texas' LIVE database).

$\alpha = 44.64$	$\beta = 22.52$	
$\mu_1 = 0.28$	$\mu_2 = 0.040$	$\mu_3 = -0.86$

Table IV Performance evaluation on Texas database.

Images	Type	Scale	CC	AAE	RMSE	SROCC
Training	NR	1-100	0.93	8.07	9.83	0.96
Testing	NR	1-100	0.92	8.18	10.1	0.95

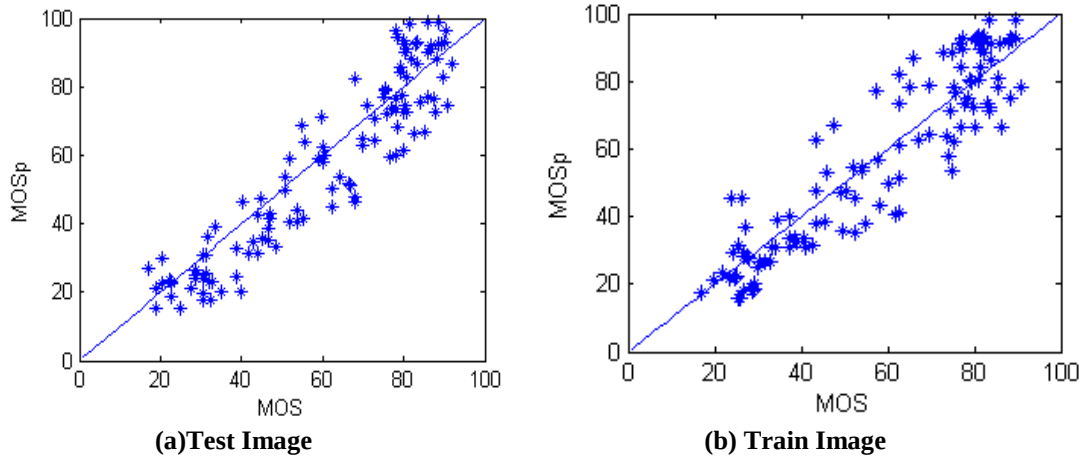
B. Performance evaluation according to VQEG

Mainly four evaluation metrics pearson correlation coefficient (CC), average absolute error (AAE), root mean square error (RMSE), and Spearman rank order correlation coefficient (SROCC) between MOS and MOSp were recommended by VQEG to evaluate the prediction. In order to verify the permanence of our method, we consider the Texas'Live database of JPEG. The database is divided into two parts for training and testing. The method's parameters are obtained for the quality scales (scale, 1-100) by using the PSO algorithm for all of the training images. Same parameters are chosen for the rest of the images of database for testing. The evaluation results are summarized in Table IV. Table IV shows that the performance of the proposed method are quite sufficient for every one of the evaluation metrics. It has also been observed from

Table IV that the proposed method provides sufficient prediction accuracy (higher CC, lower AAE, and RMSE), and sufficient prediction monotonicity (higher SROCC).

The MOS versus MOSp of our proposed method for training and testing images are respectively shown in Figures 5(a) and (b). The Figures indicate our prediction (MOSp) is sufficient with the subjective test scores (MOS) for both training and testing. Therefore, it can be said that the method performs consistently well on unknown dataset.

Figure 5 Proposed model performances.



C. Comparison between PSNR and the proposed model

Most of the objective metrics consider the statistical or mathematical measurement for finding the image artifacts. The mean squared error (MSE) [10] and the peak signal-to-noise ratio (PSNR) [11] are the most widely used pixel-based image quality metrics. These techniques are simple and fast, but widely criticized for not correlating well with human visual perception and require reference images [12]. PSNR is a simple pixel-based comparison method whereas MSE is designed on statistical features for finding differences between reference and original images. They do not consider the relationship between pixels. Although MSE or PSNR could be taken as a quality metrics but these are not consistent with the HVS because of considering every pixel within an image with equal priority. In addition, no information of structure, contrast, visibility, etc. are considered in these methods. These metrics consider the power of the error signal, but not how it affects the image. In reality pixels at different position create various effects on the HVS. Since image quality is truly represented by subjective evaluations, these metrics rarely work accurately on the quality judgment. MSE is the differences between corresponding pixels of the reference and the distorted images can be defined as:

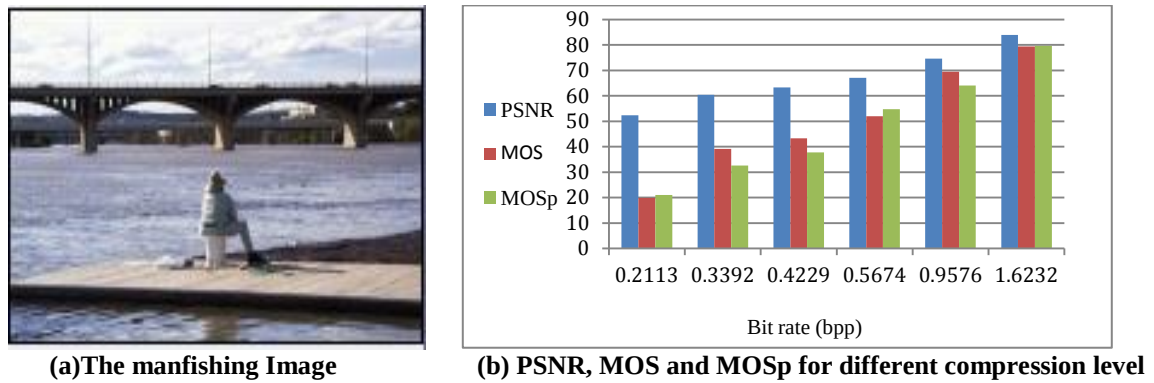
$$MSE = \frac{1}{M \times N} \sum_{m=1, n=1}^{M, N} |I(m, n) - I_d(m, n)|^2 \quad (8)$$

where $M \times N$ is image size. $I(m, n)$ and $I_d(m, n)$ represent pixels of reference and distorted images, respectively. PSNR maps the MSE in a logarithmic way which is defined as:

$$PSNR = 10 \log_{10} \frac{MAX_I}{MSE} \quad (9)$$

where MAX_I is the maximum value that a pixel can have.

We ran several experiments to analyze the prediction performance of PSNR over MOS. In order to show the performance of PSNR and our prediction (MOSp) over MOS for different levels of compression of the manfishing image is given in Figure 6. The figure shows increasing trends for MOS, MOSp and PSNR scores respectively for low levels of compressions (i.e., with increasing bit rate) for the manfishing image. From the figure we can see that PSNR is increasing very slowly where MOS and MOSp are increasing very rapidly with decreasing the compression level. So, it has been proved that PSNR, a widely used mathematical approach, does not truly reflect human perception or the subjective score, whereas our prediction model (MOSp) reflects the subjective score (MOS) quite accurately.

Figure 6 The manfishing image and its PSNR ,MOS and MOSp for different compression level.

D. Comparison between mathematical blockiness and the proposed perceived blockiness

To evaluate our blockiness detection method we calculate the mathematical blockiness which is actually the difference at the block boundary between reference image and blockiness or distorted images. In order to calculate the mathematical blockiness we first calculate the horizontal blockiness that can be defined as:

$$B_{mh} = \frac{1}{M([N/8]-1)} \sum_{i=1}^M \sum_{j=1}^{[N/8]-1} |d_{mh}(i, 8j)| \quad (10)$$

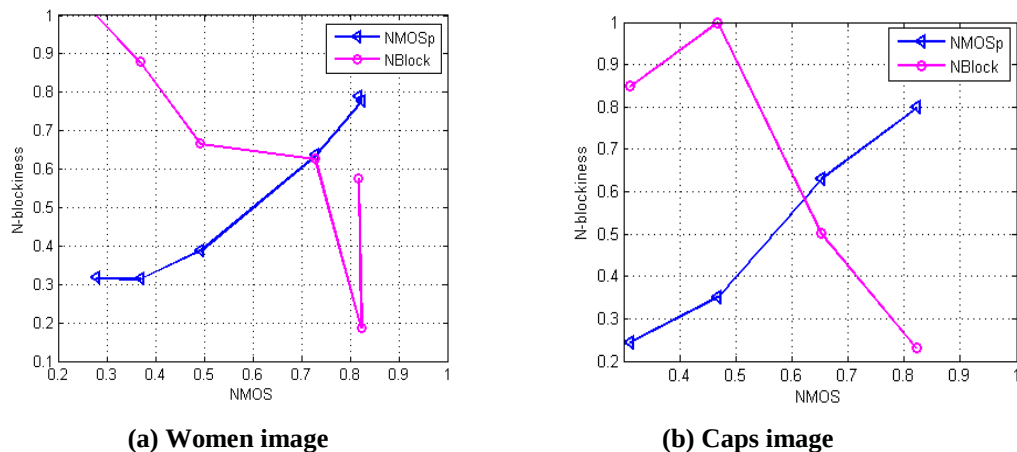
$$d_{mh}(m, n) = X(m, n+1) - X_d(m, n) \quad (11)$$

where $X(m, n)$ and $X_d(m, n)$ represent the reference and distorted images, respectively. Similarly we calculate the vertical real blockiness B_{mv} . Finally we calculate the overall mathematical blockiness as:

$$B_m = \frac{B_{mh} + B_{mv}}{2} \quad (12)$$

This mathematical blockiness calculation technique is simple and fast, but widely criticized for not correlating well with human visual perception. As increasing image quality decreases its blockiness. Therefore the relationship between the perceived blockiness and human perceived image quality (MOS) is expected to be inversely proportional. The relationship between MOS and our inverse prediction of perceptual blockiness assessment (i.e., MOSp) is expected to be linear.

We conducted several experiments to analyze the performance of mathematical blockiness (Nblock) and our prediction (NMOSp) in normalized form over MOS. The performances of NBlock and NMOSp over NMOS for two different images e.g., Women, and Caps are given in Figures 7(a) and (b). From the figures, it is clear that mathematical blockiness (NBlock) is not truly inversely proportional with human perception (NMOS). On the other hand the figures show that our prediction (NMOSp) is almost linear with NMOS. So it can be observed that our prediction of perceptual blockiness detection is much better than conventional mathematical blockiness estimation technique for different image processing to maintain image quality.

Figure 7 Normalized MOS versus blockiness.

VI. Conclusion

When images are highly compressed using B-DCT transforms, the decompressed images contain bothersome blocking artifacts. Blockiness is a special kind of image feature in the sense that human eyes can easily feel it without observation of the original images. This implies that blockiness can be and should be detected and measured blindly. This paper presented a no- reference perceived blockiness estimation algorithm based on local features and segmentation. The proposed approach is promising in terms of computational efficiency, prediction accuracy and practical reliability for real-time applications. The main contribution of this work is to develop a NR perceived blockiness detection method, which can be used not only for image quality assessment but also for real time blockiness estimation. Major artifact in compressed images is blockiness that is addressed in this paper and the test results of our proposed method illustrates the sufficient consistency with human visual perception. The evaluation also shows that our method is very compatible with human assessed scores for all different types of images and various levels of JPEG compression. This model can be embedded into JPEG coder to reduce perceptual blockiness. In future, the work can be extended to identify another major artifact blurring for JPEG2000 coded images.

VII. References

- [1] ITU-T Recommendation J.144, "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference," March, 2004.
- [2] Y. Horita, M. Sato, Y. Kawayoke, Z. M. Parvez Sazzad, and K. Shibta, "Image quality evaluation model based on local features and segmentation," in Proc. IEEE ICIP, Atlanta, USA, October, 2006.
- [3] H. R. Sheikh, A. C. Bovik, and L. Cormack, "No reference quality assessment using natural scene statistics: JPEG2000," IEEE Trans. on Image Process., vol. 14, no. 11, pp. 1918-1927, Nov., 2005.
- [4] Zhou Wang, Hamid R. Sheikh and Alan C. Bovik, "No-Reference Perceptual Quality Assessment of JPEG Compressed Images." IEEE ICIP New York, USA, September 22-25, 2002.
- [5] Anil K Jain, "Fundamentals of Digital Image Processing," Prentice Hall, NJ, USA, October 1988.
- [6] Alan Wee-Chung Liew, and Hong Yan, "Blocking artifacts suppression in block-coded images using over complete wavelet representation," IEEE Trans. on Circuits and Systems for Video Technology (TCSV), vol. 14, no. 4, pp. 450-461, April, 2004.
- [7] K. R. Rao, and C. Patrick C. Y. Yip, "Discrete Cosine Transform: Algorithms, Advantages, Applications," Academic Press Professional Inc., San Diego, CA, USA, September, 1990.
- [8] ITU-R BT.500-11: Methodology for the subjective assessment of the quality of television pictures.
- [9] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. Live image quality assessment database, 2005.
- [10] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," IEEE Trans. on Image Process., vol. 14, no. 2, pp. 2117-2128, Dec. 2005.
- [11] Y. Horita, Y. Kawai, Y. Minami, and T. Murai, "Quality evaluation model of coded stereoscopic color image," in Proc. of SPIE: Visual Communications and Image Processing (VCIP), vol. 4067, pp. 389-398, Perth, Australia, June 2000.
- [12] Z. M. Parvez Sazzad, "No-Reference Image Quality Assessments for JPEG and JPEG2000 Coded Images," PhD thesis, University of Toyama, Department of Systems Science and Engineering, March, 20.