New No-reference Blocking Artifacts Metric Based on Human Visual System

Ahong Wang¹, Gangyi Jiang¹*, Xu Wang¹
1. Faculty of Information Science and Engineering, Ningbo University, Ningbo-city, 315211, China
yumei2@126.com

Mei Yu¹²
2. National Key Lab of Software New Technology, Nanjing University, Nanjing-city, 210093, China

Abstract— Block-based discrete cosine transform (BDCT) is an important tool in the image and video compression. Blocking artifacts caused by BDCT may be noticeable in the decoded image under low bit-rate conditions. Measuring blocking artifacts is of great significance in the evaluation of image and video image. To ensure reliability of the model outputs, a new no-reference blocking artifacts metric is proposed in this paper, based on human vision sensitivity by the research on a model of luminance and texture masking. Furthermore, blocking artifacts are measured with different weighting coefficient in flat regions and edge regions, respectively. Compared with Wang’s blind measurement of blocking artifacts, experimental results illustrate that the proposed metric has better consistency to the visual perception, and it can reflect the image quality more effectively.

Keywords— image quality assessment; block-based discrete cosine transform; blocking artifacts; masking; visual perception

I. INTRODUCTION

With the widespread development of video communication, many efficient image compression methods have been developed and standardized. Especially, high quality image communication with low bit-rate is widely used for video conferencing, videophone, etc. Block-based discrete cosine transform (BDCT) is the most widely used technique for the compression of both still and moving images [1]. Most of image and video coding standards, such as JPEG and MPEG, are based on BDCT. One of the most noticeable artifacts caused by coding algorithm is an artificial discontinuity across the block boundaries called “blocking artifact”, due to the independent quantization of the transform coefficients. Peak signal to noise ratio (PSNR) and mean square error (MSE) are popular image quality evaluation criterions, but they are ineffective measurements for blocking artifacts that are structural disturbance [2]. It is therefore essential to develop metrics that have good consistency to the visual perception.

Many researches had been done to develop objective image quality metrics, which can be classified into three categories according to dependence on original image, full-reference (FR) metrics, reduced-reference (RR) metrics and no-reference (NR) metrics. Of three categories of metrics, the development of FR metrics is the most maturely, which make use of original images for comparison to distorted images [3, 4, 5]. However, the original images are not always available at the receiving end. Hence, RR metrics and NR metrics, with some statistical characteristics and no knowledge about the original images, are of much greater interest and make more sense in applications. However, it is difficult to design a NR quality evaluation model for the lack of understanding of human visual system (HVS) [6]. Furthermore, it is undeniable that NR metrics are feasible only if we know the type of distortion. Currently, NR quality metrics are the subject by the research community with the emergence of Video Quality Experts Group (VQEG) [3].

A no-reference blocking artifact measurement algorithm generalized block-edge impairment metric (GBIM) [7] was the well-known metric. GBIM uses the intensity changes along the adjacent block boundaries as the blockiness measured by taking into account the luminance masking effects in a reconstructed image. Mean squared difference of slope (MSDS) [8] was a special metric for measuring blocking artifacts, and it calculates the gradients of the horizontal or vertical adjacent blocks. George et al. gave expansion of the MSDS [9] by increasing the distortion measurement of angle directions. Suthaharan proposed PS-BIM standards [10], which makes use of several different regions of the visual perception and generates perceptual weights to evaluate the blocking artifacts. A measuring blocking artifacts algorithm was proposed by taking the edge directional information of the images instead of using the traditional pixel discontinuity along the block boundary [11]. Liu et al. proposed a perceptually relevant no-reference blockiness metric to quantify blocking annoyance based on its local image characteristics [12], which calculates the blockiness metric only at the expected block edges instead of at all pixels in an image and uses a grid detector to ensure that the metric is calculated at the exact position of the block boundaries.

In this paper, a new NR blocking artifacts metric is presented to quantify the quality of distorted images based on various human visual characteristics. Experimental results show that the proposed metric has good consistency to the visual perception, and it can reflect the image quality effectively. The rest of paper is organized as follows. The proposed algorithm is discussed in detail in Section 2. Section 3 provides and discusses experimental results. Finally, conclusions are presented in Section 4.
II. THE PROPOSED BLOCKING ARTIFACTS METRIC

In BDCT image and video coding, one of the most noticeable artifacts is the blocking artifacts. It is the direct result of independent processing of neighboring blocks without considering correlations between these blocks. To evaluate the consistency between subjective quality and visual perception, we propose a new NR blocking artifact metric with taking various human visual characteristics into account. Blocking artifacts include horizontal and vertical edge distortions. The visual sensitivity to these edge artifacts is affected by some parameters. First, distorted images with blocking artifact are divided into flat and edge regions, of which blockiness are measured separately with different weighting coefficients. Second, the visibility of blocking artifacts will be affected by local surrounding background called masking, including luminance masking and texture masking in this model. Fig.1 shows the overall description of the proposed blocking artifact metric.

A. Edge Detection

Human eyes are very sensitive to the blocking artifacts on the flat regions, while they are not sensitive to that on the flat regions based on theory of human visual system, which is referred to as visual masking effect [13]. In order to evaluate image quality more effectively, a weighting model is proposed that the weighting coefficients for flat regions are greater than those of the edge regions. The weighting coefficient is defined as

\[ \omega = \frac{1}{G_{i,j} + T + 1} \]  

Where \( G_{i,j} \) and \( T \) is gradient value at the pixel \( (i,j) \) and the gradient of edge threshold, respectively. \( \omega \) will decrease with \( G_{i,j} \) increasing. Edges are detected with Sobel operator.

The vertical difference image is expressed by

\[ d_{v}(i,j) = |f(i,j) - f(i-1,j)| \]  

where \( f(i,j) \) is test image, and \( d(i,j) \) is vertical difference image.

After considering weighting coefficients for the differences of the flat and edge regions, the weighted vertical difference image is

\[ d_{w}(i,j) = d_{v}(i,j) \times w(i,j) \]  

Horizontal differences can also be easily acquired since human sensitivity to horizontal and vertical orientations have proven to be similar. Fig.2 shows the model to predict vertical difference image.

B. Masking

Through the research on observation of some visual phenomenon combining with psycho-visual experiments, it has been discovered that human vision system has a lot of human visual system characteristics mainly about performance of a variety of visual effects [12], which are related to image information processing directly or indirectly. Therefore, it is essential to take human visual system characteristics into account for the image quality evaluation so that the results of objective evaluation will have better consistency to the visual perception.
In order to make the algorithm reflect the image quality more effectively, masking which will be affected by the local surrounding background is one of the important characteristics. It includes luminance and texture masking that both are highly relevant to the perception of blockiness.

1) luminance masking

The visibility of edge is often affected by the characteristic of its surrounding spatial regions. The edge lying in the middle gray region is the most visible. Fig. 3 shows an example of luminance masking, the difference of brightness level between block and surrounding spatial region is the same. However, the edge in Fig. 3(b) is more noticeable than that in Fig. 3(a), and a distortion is most visible for a surrounding with an averaged luminance value between 70 and 90 (centered approx. at 81) in 8 bits grayscale images [6, 17].

![Fig. 3. An example of luminance masking](image)

The procedure of luminance masking includes two steps; local luminance detection and visibility transform function.

\[
I_{ij} = \frac{1}{25} \sum_{x=1}^{3} \sum_{y=1}^{3} (I(i-3+x,j-3+y) - L(x,y))
\]

\[
VC_{ij} = \begin{cases} 
    0 & 0 \leq I_{ij} \leq 81, \\
    \frac{1-\beta}{14} (81-I_{ij}) + 1 & \text{else,}
\end{cases}
\]

where \(0 < \beta < 1\) (\(\beta = 0.8\) in these experiments) is used to adjust the slope of the linear part of the function, where \(I_{ij}\) is the luminance of the pixel at \((i,j)\), and \(L(x,y)\) is used for calculating the background luminance, which is different in horizontal and vertical directions. \(VC_{ij}\) denotes visibility coefficient of luminance masking.

2) texture masking

Texture masking can hide distortion in high-contrast image regions, and the blocking artifact will also be masked. The edge lying in the textured region is less visible than the edge in the plain background. The procedure of texture masking consists of three steps; texture detection, thresholding and visibility transform function, and represented by

\[
d(i,j) = \begin{cases} 
    0 & \text{if } t(i,j) < Tr, \\
    t(i,j) & \text{else,}
\end{cases}
\]

\[
T(x,y) = \frac{1}{48} \sum_{x=1}^{3} \sum_{y=1}^{3} (I(i-3+x,j-3+y) - T(x,y))
\]

\[
VC_{ij} = \frac{1}{(1 + d(i,j))^\alpha}
\]

where \(Tr\) is a threshold, and set as 0.15 typically. \(\alpha > 1\) is used to adjust to the nonlinearity (\(\alpha = 10\) in the experiments), and \(T(x,y)\) is used for measuring the background activity in the horizontal and vertical directions. \(VC_{ij}\) denotes the visibility coefficient of texture masking.

C. The proposed blocking artifacts metric

Edge detection in this study shows that the weighting coefficients for the flat and edge regions are different, considering that human eyes are more sensitive to blocking artifacts on the flat regions than that on the edge regions. Masking occurs when a stimulus that is visible by itself cannot be detected due to the presence of another. The visibility coefficient \(VC\) is the integration of \(VC_f\) and \(VC_t\).

\[
d_{se}(i,j) = d(i,j) \times VC(i,j)
\]

Then the power spectrum of the image is estimated, which is generated by putting \(d_{se}\) into 1-D using the fast Fourier transform. The final measuring result of the blocking artifacts based on human visual system BAMHVS is obtained by making use of the values at the feature frequencies [14].

III. EXPERIMENTAL RESULTS AND ANALYSES

To test the performance of the proposed algorithm in experiments, JPEG distortion database is used, provided by Laboratory for Image and Video Engineering (LIVE) [15]. In this database, 29 high-resolution images were compressed using JPEG with different compression ratios. There are 233 images including 29 original images, the degradation mean opinion score (DMOS) values of which are also validated in the database.

After obtaining the objective measuring results of the metric, two evaluation criterions are defined to access the consistency of objective and subjective scores [16], Pearson correlation coefficient (CC) for reflecting the accuracy and Spearman rank order correlation coefficient (ROCC) representing monotoncity of objective scores.
The Pearson correlation and Spearman rank-order correlation between BAMHVS and DMOS are 0.9570 and 0.9532 for Group 1 images, respectively. The scatter plots of BAMHVS to Group 1 and Group 2 images in LIVE database show in Fig 4(a).

The results in Fig 4 and Table 1 show that the proposed metric estimates the image quality better than Wang’s metric without considering human visual characteristics [14]. From Table 1, the Pearson CC as well as Spearman ROCC of the proposed metric are both higher than Wang’s.

Therefore, this proposed no-reference blocking artifacts metric based on human visual system can reflect the visual perception of image quality more accurately.

IV. CONCLUSIONS

The BDCT is the widely used technique for compression of both still and moving images, but some traditional image quality evaluation criterions may be ineffective measurements for blocking artifacts. Considering the original image is not always available at the receiving end, it’s therefore essential to develop NR metrics that have good consistency to the visual perception. In order to evaluate image quality more effectively, various human visual characteristics should be considered.

This paper presents an effective no-referenced approach for measuring blocking artifacts in BDCT coding. Compared with Wang’s blind measurement method of blocking artifacts, an additional advantage of the proposed metric is taking various human visual characteristics into account while quantifying the blocking artifact by the research on a model of luminance and texture masking. Furthermore, blocking artifacts are measured with different weighting coefficients on flat and edge regions respectively. The experimental results show that the proposed blocking artifacts metric based on human visual system have a good consistency with subjective scores.

The proposed NR metric is mainly used to evaluate quality of the gray-scale images. In this model, we change color image into gray-scale image before evaluation of its quality. It
is more sensitive to color images for human visual system. Therefore, how to measure quality of color images effectively and how to minimize the complexity of quality evaluation algorithm will be essential for further research.

ACKNOWLEDGMENT

This work was supported by Natural Science Foundation of China (Grant No:60832003, 60872094), the Projects of Chinese Ministry of Education (Grant No:NCET-06-0537, 200816460003), and the National 863 High Technology Research and Development Program of China (Grant No:2009AA01Z327).

REFERENCES


