A no-reference blocking artifact metric for B-DCT video

YANG Fu-zheng†1, WAN Shuai2, CHANG Yi-lin1, LUO Zhong3
(1National Key Lab on ISN, Xidian University, Xi’an 710071, China)
(2Multimedia and Vision Research Laboratory, Electronic Engineering Department Queen Mary, University of London, London E1 4NS, UK)
(3Huawei Technology Limited Cooperation, Shenzhen 518057, China)
†E-mail: fzhyang@mail.xidian.edu.cn
Received Nov. 30, 2005; revision accepted Feb. 17, 2006

Abstract: A new no-reference blocking artifact metric for B-DCT compression video is presented in this paper. We first present a new definition of blocking artifact and a new method for measuring perceptive blocking artifact based on HVS taking into account the luminance masking and activity masking characteristic. Then, we propose a new concept of blocking artifact cluster and the algorithm for clustering blocking artifacts. Considering eye movement and fixation, we select several clusters with most serious blocking artifacts and utilize the average of their blocking artifacts to assess the total blocking artifact of B-DCT reconstructed video. Experimental results illustrating the performance of the proposed method are presented and evaluated.

Key words: Blocking artifact cluster, Blocking artifact metric, HVS, Video quality assessment
doi: 10.1631/jzus.2006.AS0095

INTRODUCTION

The block-based discrete cosine transform (B-DCT) scheme is a fundamental component of many image and video compression standards including JPEG, H.263, MPEG-1, MPEG-2, MPEG-4, H.264, and others, used in a wide range of applications. The coding algorithm exploits the local spatial redundancies of input video to achieve a low bit rate, which may cause visible coding distortions in reconstructed sequences due to the lost coding nature, such as blocking artifact, ringing, mosquito effect, MC mismatch, blurring, and color bleeding (Yuen and Wu, 1998). Among B-DCT digital-video coding distortions, blocking artifact is of particular importance. As shown in (Karunasekera and Kingsbury, 1995), blocking artifact and its propagation through reconstructed video sequences, are the most significant of all coding artifacts. The blocking artifact thus represents a major type of distortion in the reconstructed video compressed with B-DCT algorithm, which will be particularly annoying in the case of high compression ratio.

A great deal of research has been carried out measuring blocking artifact. Karunasekera and Kingsbury introduced a distortion measure of blocking artifact based on human visual system (HVS) (Wu, 1995). And a perceptual blocking distortion metric is introduced which is based on more complex HVS models (Yu et al., 2002). These quantitative measures require both the original and reconstructed images. However, in the absence of the original image, the above distortion measures cannot be used to evaluate the quality of a reconstructed image, which makes them impractical for real time transmission applications. People have long been investigating no-reference blocking artifact metrics (Vlachos, 2000; Gao et al., 2002; Triantafyllidis et al., 2002). However, HVS should be taken into account in the measurement of blocking artifact. A no-reference measure of blocking artifact is also proposed with some inadequate models of HVS (Wu and Yuen, 1997).

In this paper, a no-reference blocking artifact metric is presented which does not need reference
image at all. We present a new definition of blocking artifact and a new method for measuring the perceptive blocking artifact based on HVS. Considering eye movement and fixation, we also introduce a new concept of blocking artifact cluster, and propose a no-reference blocking artifact metric for B-DCT reconstructed video based on it.

The rest of this paper is organized as follows. In Section 2, the perceptive blocking artifact measure is proposed. Section 3 presents the new concept of blocking artifact cluster and the new blocking artifact metric for B-DCT reconstructed video. Experimental results given in Section 4 evaluate the performance of the metric. Finally, conclusions are drawn in Section 5.

PERCEPTIVE BLOCKING ARTIFACT

Since blocks of pixels are treated as single entities and coded separately, correlation among spatially adjacent blocks is not taken into account in coding, which results in block boundaries being visible when the decoded image is reconstructed. The blocking artifact is commonly defined as the discontinuities at the boundaries. However, the term can be generalized to mean an overall difference between adjacent blocks. In experiments, we observed that the block is especially annoying to human eyes when it is discontinuous with its spatially adjacent blocks at more than one boundary. Here, we call it mosaic block. Therefore, we define the blocking artifact of a block as the total discontinuity of its four boundaries with the adjacent blocks in this paper. Then the blocking artifact \( d_{ij} \) of block \( B_{ij} \) is defined by

\[
d_{ij} = \frac{1}{4} \sum_{l=1}^{4} d_{ij,l},
\]

where \( d_{ij,l} \) \((l=1, \ldots , 4)\) represent the discontinuities at \( B_{ij} \)’s four boundaries with spatially adjacent blocks \( B_{ij-1} \) (left), \( B_{ij+1} \) (right), \( B_{i-1,j} \) (upper) and \( B_{i+1,j} \) (lower) respectively. \( d_{ij,l} \) is called “edge artifact” in this paper.

The human observer is the end user of most image information. Therefore, human perception should be taken into account in the measure of blocking artifact. The models of HVS generally consist of opposite-colors space, contrast sensitivity, multi-resolution architecture, masking, pattern adaptation, pooling and cognitive processes (Winkler, 1999). However, the human visual system is extremely complex, and many of its properties are not well understood even today.

In this paper, straightforward but effective HVS models are used to get the perceptive blocking artifact measure. Because the multichannel process is very complex and the contrast gain control at interchannel is not very clear, most attention is focused on background luminance (brightness) masking and activity masking in this paper. Human vision has lower acuity in the color pathway. Winkler (2000) conducted research on the influence of the choice of color space on the performance of models. It has been shown that it is possible for the vision models to work on luminance component only without a dramatic degradation in prediction accuracy. Therefore, the metric introduced in this paper uses only the luminance component. Take the perceptive edge artifact \( d_{1,ij} \) for example, we will introduce in detail how to obtain the edge artifact with consideration of HVS in the following.

Original edge artifact

Let us consider an image \( F \) (size of \( L \times C \)) which consists of \( K \) blocks, each of size \( N \times N \). A pixel luminance value at \((n,m)\), \(0 \leq n,m \leq N-1\), in the block \( B_{ij} \) \((0 \leq i \leq L/N, 0 \leq j \leq C/N)\) is denoted by \( f_i(n,m) \).

The original edge artifact \( ds_{1,ij} \) of \( B_{ij} \) at the left boundary with \( B_{ij-1} \) is defined by

\[
ds_{1,ij} = \left\{ \begin{array}{ll}
\frac{1}{N} \sum_{n=0}^{N-1} f_{ij}(n,0) - \sum_{n=0}^{N-1} f_{ij-1}(n,N-1), & j>0, \\
0, & j=0.
\end{array} \right.
\]

Luminance masking

The response of the human visual system depends much less on the absolute luminance than on the relation of its local variations from the surrounding luminance. Unfortunately, a common definition of contrast suitable for all stimuli does not exist. We use Weber contrast here that can be expressed as

\[
C_W = \frac{\Delta L}{L}.
\]