A NO-REFERENCE BLOCKING ARTIFACT MEASURE FOR ADAPTIVE VIDEO PROCESSING

R. Muijs and I. Kirenko
Philips Research Laboratories
Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands
Remco.Muijs@ philips.com, Ihor.Kirenko@philips.com

ABSTRACT

Objective image quality metrics offer the prospect of adapting video processing algorithms to the quality of the incoming signal. In the context of image enhancement, such as peaking, contrast enhancement or color mapping, the original image does not necessarily correspond to the subjective optimum. There is, therefore, a compelling need for reliable quality metrics that are based exclusively on the characteristics of the processed images (no-reference). In this paper, we illustrate the design and application of no-reference quality metrics for the case of blocking artifacts that commonly degrade the quality of block-based DCT encoded video. We outline a simple cost-effective method that allows the grid position and its visibility to be determined without the need for access to the coding parameters. This information, in turn, is used to effectively suppress blocking artifacts while preserving the sharpness of object edges.

1. INTRODUCTION

Coarse quantization of DCT coefficients during block-based video encoding, such as MPEG-2 and H.264, may result in the appearance of strong discontinuities at the borders of individual blocks in the decoded video. Subjective experiments have indicated that blockiness is the most annoying coding artifact at low to moderate (<3Mbit/s) bitrates and is highly correlated with the overall perceived quality of MPEG-2 encoded video [1]. Optimal suppression of blocking artifacts requires information about both the position and the visual strength of the discontinuities. Because access to the bit stream is not always feasible for architectural reasons, there is a need for methods that allow this information to be extracted from the decoded image alone.

Existing grid detection methods typically involve the computation of a gradient map (or other high-pass filtering), followed by a thresholding operation to distinguish block discontinuities from natural contours (e.g., [2,3]). In practice, the selection of suitable threshold levels is notoriously difficult and often prevents the algorithms from performing satisfactorily over a wide variety of image content, compression rates and video resolution. Wu and Yuen [4] proposed a Generalized Block Impairment Metric (GBIM) that incorporates perceptual features, such as texture and luminance masking. The GBIM-metric correlates well with subjective data and is widely used in the image processing community to assess the quality of JPEG-encoded images. However, this method does not include a grid detection phase, but explicitly assumes the presence of an 8x8 pixel block grid that starts in the top-left corner of the picture. As such, it is not readily applicable to scaled decoded video that may contain deviating grid dimensions. Alternative methods for detecting and measuring the visibility of the block grid employ a frequency-based representation of the video (e.g., [5]). Due to the additional Fourier, DCT or wavelet transforms involved, the computational cost of these methods is relatively high.

In this paper, we present a simple, cost-effective method for determining the position and the visual strength of the coding block grid. The obtained information is used to drive a straightforward deblocking algorithm that effectively removes the artifacts while retaining the sharpness of object edges.

2. BLOCKING ARTIFACT VISIBILITY

The visual strength of a block edge is predominantly affected by the magnitude of the edge gradient and the spatial activity in the direct vicinity of the block border [4]. Several authors have suggested to make use of information about the human visual system, such as the contrast sensitivity function and texture and luminance masking, to model the visibility of block discontinuities (e.g., [6]). To reduce the numerical cost of these methods, we propose a simple, efficient algorithm based on the principle that block discontinuities can be characterized as edges that stand out from the spatial activity in their vicinity. In other words, the visibility of a block edge is determined by the contrast between the local gradient and the average gradient of the adjacent pixels.

In the following, we discuss the detection of vertical block edges, but identification of horizontal artifacts is accomplished in a similar fashion. Consider an image ***I*** with elements *I(i,j)*, where *i* and *j* denote the line and pixel position, respectively. To express the similarity between the local gradient and its spatial neighbors, we introduce the normalized horizontal gradient $\boldsymbol{D_{H,norm}}$ as the ratio of the absolute gradient and the average gradient calculated over *N* adjacent pixels to the left and to the

$$D_{H,norm}(i,j) = \frac{\left|I(i+1,j) - I(i,j)\right|}{\frac{1}{2N}\sum_{n=-N...N,n\neq 0}\left|I(i+n+1,j) - I(i+n,j)\right|} \cdot (1)$$

right:

Because block edges occur at regular intervals in the horizontal or vertical direction, they can be further highlighted by summing $\boldsymbol{D_{H,norm}}$ over all image lines *nl*:

$$S_H(i) = \sum_{j=1}^{nl} D_{H,norm}(i,j). \qquad (2)$$

The presence of blocking artifacts will result in pronounced maxima in S_H. The above procedure is illustrated for the image *branch* displayed in Figure 1a. Although blocking artifacts are difficult to identify in the original image, the periodic structure of the encoding grid is clearly revealed in the horizontal accumulator S_H shown in Figure 1b. The size and offset of the grid can be readily extracted from this signal by means of conventional histogram analysis of the peak locations.

The visual strength of the blocking artifacts can be determined by averaging S_H over the block edge and intermediate positions. The Blocking Strength (BS) is then defined as:

$$BS = \frac{\bar{S}_H(block)}{\bar{S}_H(non-block)}, \qquad (3)$$

where $\bar{S}_H(block)$ and $\bar{S}_H(non-block)$ denote the average value of S_H at the block edge and intermediate positions, respectively.

The accuracy of the objective blockiness metric *BS* was assessed using the LIVE Image Quality Assessment Database [7], that consists of 169 JPEG encoded images and associated mean quality scores (MQS). Figure 1c displays the relation between the objective metric *BS* and

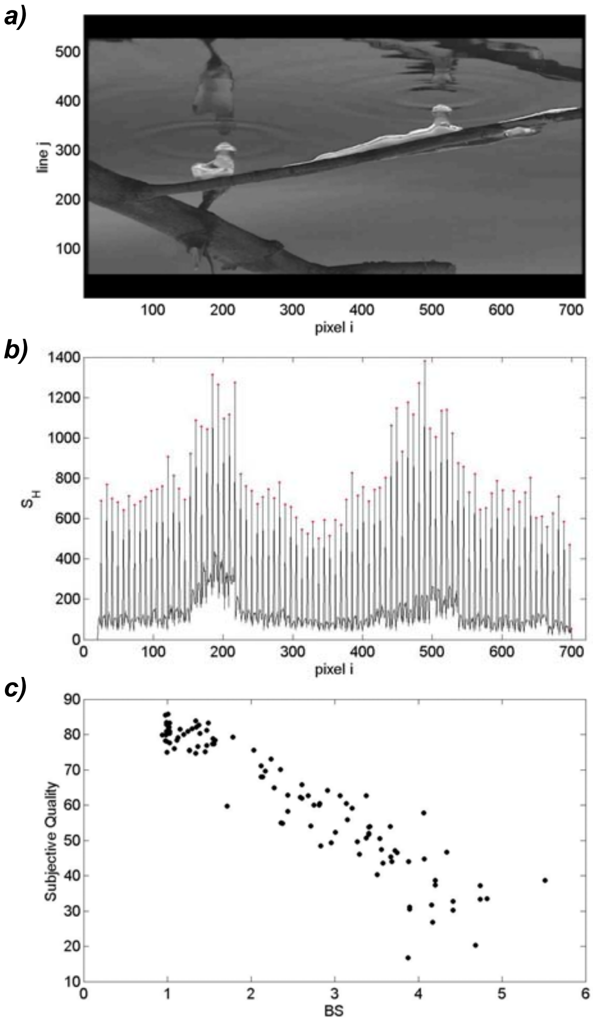

Figure 1: Detection of the position and visibility of the coding grid. Shown are (a) the image *branch* encoded at a bit-rate of 2Mb/s, (b) the horizontal accumulator S_H computed using equations (1) and (2) and (c) the the objective blockiness metric BS vs. the mean subjective quality score for JPEG encoded images from the LIVE database.

the subjective MQS. The Pearson correlation coefficient of these data amounts to 0.92. In spite of the simplicity of the outlined approach, *BS* provides an accurate prediction of the subjective quality.

Although block-based compression algorithms commonly employ grid sizes of 8x8 pixels, spatial scaling may be applied to the video signal after encoding. Consequently, the effective block size in the video signal corresponds to 8 pixels times the scaling factor in the horizontal and vertical directions. Moreover, blocking artifacts are blurred by the scaling operation, which may potentially affect their visibility.

To study the impact of scaling on the visibility of block discontinuities, a small subjective experiment was conducted with four video sequences and four still images (snapshots taken out of the four sequences). The original stimuli were downscaled with scaling factors 1x1, 0.75x1, 0.67x1 and 0.5x0.5, compressed using a regular 8x8 block based MPEG-2 encoder at bitrates of 1, 2, 3 and 4 Mbps and then up-scaled to their original size after decoding. The resulting video material thus contains block sizes of 8x8, 10.67x8, 12x8 and 16x16 pixels, respectively. To keep the amount of induced compression constant, the compression bitrates were scaled by the same factor as the video signal. Twenty naïve subjects scored 136 stimuli (originating from four image contents) on overall quality using a numerical scale from 0 (lowest quality) to 10 (highest quality).

Figure 2a illustrates the effect of scaling on the blockiness metric *BS* [equation (3)] for the image *branch* that is shown in Figure 1a. Although the objective blockiness is highly correlated with the subjective data for each of the individual scaling factors, the correlation across the scaling factors is generally poor. To alleviate this shortcoming, the blockiness metric BS is extended with a correction factor that accounts for the effect of the grid size *d* on the perceived blockiness:

$$BS_{norm} = BS * f(d), \tag{4}$$

where *f(d)* is determined on the basis of linear regression and is defined as:

$$f(d) = 0.38\left(\frac{d}{8}\right) + 0.62. \tag{5}$$

Figures 2b and 2c show the effect of compensating the blockiness metric *BS* for the grid size for the image branch and for all four original image contents, respectively. After correcting the metric for the grid size, the Pearson correlation coefficient of the data shown in Figure 2c is 0.83. Without the correction factor *f(d)*, it amounts to 0.62. The improved consistency of the objective data across scaling factors implies that a meaningful blockiness metric can be accurately determined and applied at arbitrary positions in the video chain.

3. BLOCKING ARTIFACT SUPPRESSION

One of the most efficient and simplest ways to suppress blocking artifacts is by means of adaptive spatial

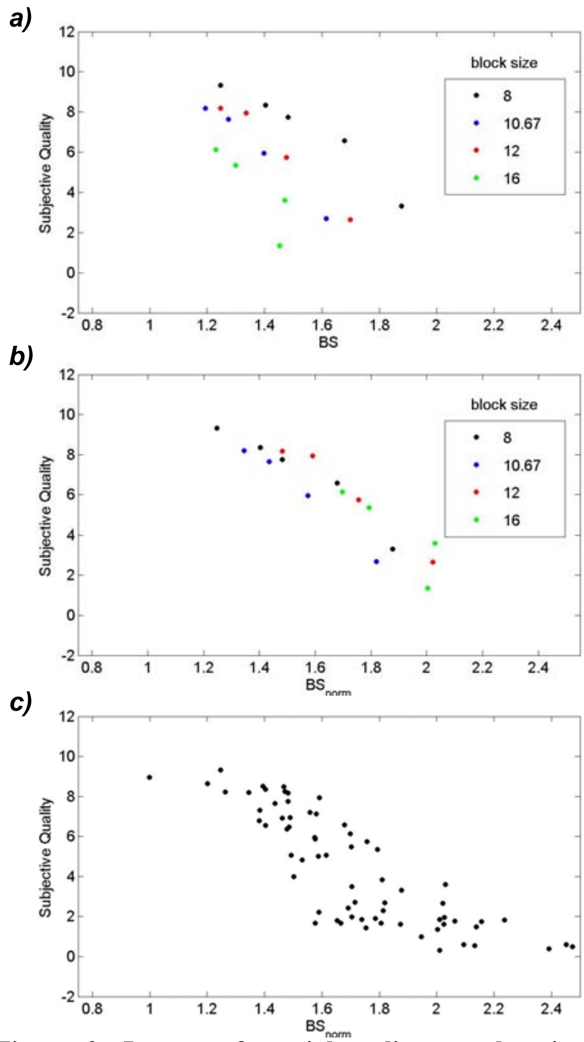

Figure 2: Impact of spatial scaling on the visual strength of blocking artifacts. Shown are the results for the image *branch* (a) before and (b) after correcting BS for the grid size using equation (5). (c) The blockiness metric BS_{norm} vs the subjective data for all 136 stimuli included in the subjective experiment.

low-pass filtering. To optimally preserve the image sharpness, object edges should be distinguished from block discontinuities, such that low-pass filtering is applied only to those pixel positions where block artifacts are visible. Deblocking algorithms thus greatly benefit from the grid detection method and the associated blockiness metric described in the previous section.

Here, we apply one-dimensional low-pass filtering over the detected block edges in the horizontal direction, followed by vertical filtering (of the horizontal edges). Block edges are filtered in raster-scan order throughout the picture. We distinguish two types of low-pass filters, one corresponding to a moderate and the other to a strong blur kernel. The choice between the two filter modes is a)

b)

Figure 3: The image *Akiyo* (a) before and (b) after application of adaptive deblocking.

based on the global blockiness BS_{norm}, as well as the local relative gradient $D_{H,norm}$ as defined in equation (1). As such, this approach provides a simple low-cost solution that allows blocking artifacts to be effectively suppressed while retaining the image sharpness. The performance of this simple deblocking concept is illustrated in Figure 3 for the image *Akiyo*. Figure 3a shows the decoded image, which contains substantial levels of blocking artifacts. The result after grid detection and adaptive deblocking is displayed in Figure 3b. Block artifacts have been substantially suppressed while the sharpness of object edges has been preserved.

4. CONCLUSIONS

Objective image quality metrics are usually developed and combined with the aim of evaluating the performance of individual video processing algorithms or the total video chain. Instead, they can also be used as inherent parts of video processing algorithms, enabling these methods to be adapted to the quality of the incoming video. In this paper, we have presented a simple method with which the position and the visibility of the coding grid can be efficiently determined. The proposed blockiness metric accommodates scaled video and is highly correlated with subjective data. The obtained information was successfully applied to regulate methods aimed at coding artifact suppression, but may also be used for adaptive enhancement video signals in the presence of digital noise.

ACKNOWLEDGMENTS

We are very grateful to F. Oberti, C. Adsimili, P. Hofman A. Dommisse and J. Tegenbosch for their contributions and fruitful discussions.

REFERENCES

[1] C.C. Koh, S.K. Mitra, J.M. Foley, and I.E.J. Heynderickx, "Annoyance of individual artifacts in MPEG-2 compressed video and their relation to overall annoyance", Proceedings of the SPIE, San Jose, CA., Jan. 2005.

[2] A. Drouot, L. Rouvellou, "Method and device for identifying block artifacts in digital Video pictures", Patent WO01/20912A1, Koninklijke Philips Electronics N.V., Mar. 2001.

[3] E. Lesellier and J. Jung, "Robust wavelet-based arbitrary grid detection for MPEG", Proc. IEEE ICIP 2002, pp 417-420, June 2002.

[4] H.R. Wu and M. Yuen, "A generalized block edge impairment metric for video coding", *IEEE Signal Processing Letters*, vol. 4, no. 11, pp. 317-320, Nov. 1997.

[5] B. Janko and J. Raitz, "Harmonic measurement of blockiness in video signals", Patent US 6437812B1, Textronix Inc., Dec. 1999.

[6] S. A. Karunasekera and N. G. Kingsbury, "A Distortion Measure for Blocking Artifacts in Images Based on Human Visual Sensitivity", *IEEE Trans. Image Processing,* vol. 4, no. 6, pp. 713-724, June 1995.

[7] H.R. Sheikh, Z. Wang, L. Cormack and A.C. Bovik, "LIVE Image Quality Assessment Database", http://live.ece.utexas.edu/research/quality.