Blockiness Measurement for MPEG Video

K. T. Tan
School of Electrical and Electronic Engineering,
Singapore Polytechnic, Singapore

and

M. Ghanbari
Department of Electronic Systems Engineering
University of Essex, Colchester, United Kingdom

ABSTRACT

A new method of blockiness measurement for MPEG video is introduced. The amplitude and phase information of the harmonics generated by blocking artefacts in MPEG video may be exploited for estimation of the blockiness. Measurement can be performed without the reference signal, offering the solution for a single-input blockiness detector.

Keywords

Blockiness, Tiling effect, Picture quality, Objective measurement.

1. INTRODUCTION

Blockiness, also known as tiling effect, is a common artefact in MPEG video. A tool for measuring the blockiness is very much desired to facilitate control or monitoring of video quality. Blockiness is normally detected by searching for artificial and abrupt horizontal and vertical luminance transitions in the coded video [1]. By referring to the reference video, the artefacts can be located easily. For some applications, the reference pictures can be assumed available. For instance, when comparing the performance of two MPEG encoders, the reference pictures are normally available. However, for many other applications, such as monitoring of the picture quality over networks, it is too costly. A solution is to transmit sub-reference picture information for quality measurement purposes [2], but accompanying side information with the video signal inevitably reduces the payload efficiency of the transporting system. A more practical way is to use only the compressed video for the measurement process. Thus a robust approach for detection of the blockiness involving only the coded video is urgently needed. In this paper, a novel method achieving that purpose will be introduced, addressing the problem of designing a single-input blockiness detector.

2. HARMONICS ANALYSIS PROCESS

Our blockiness estimation method is based on the concept of harmonics analysis. When transformed into luminance gradient, the discontinuities across DCT blocks in MPEG picture are revealed as pseudo-periodic pattern in spatial domain. When this pattern is transformed (Fourier transformation) into frequency domain, harmonics of fixed frequency are obtained. The strength of these harmonics is proportional to the amount of blockiness in the coded picture, and therefore can be exploited for estimation of the blockiness. Figure 1(a) illustrates the ideal gradient image of a block of 32×32 pixels. It was obtained by taking the absolute value of the gradient image derived using Sobel operators. Figure 1(b) shows a 32×32 block of the gradient image generated from a real image using 3×3 Sobel operators, and its FFT equivalent in Figure 1(c). The harmonics can be clearly observed on the two principle axes.

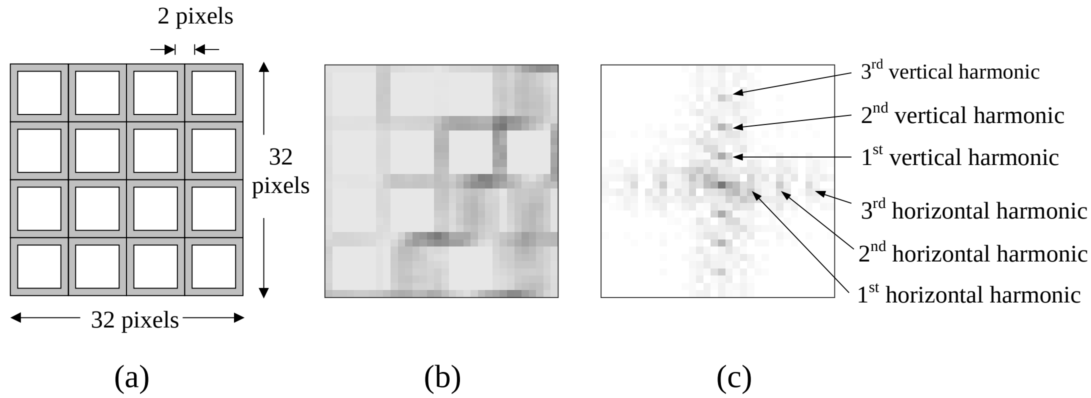


Figure 1. (a) Ideal luminance gradient image of a block of 32 × 32 pixels, (b) an actual 32 × 32-pixel block of luminance gradient image, (c) Fourier Transform of the 32 × 32 pixel luminance gradient image.

To estimate the blockiness, the luminance gradient image is first segmented into blocks of 32×32 pixels. Each block is then Fast Fourier transformed into frequency domain. Following this, we measure the ratio of the sum of harmonics to the sum of all the AC components, in both the horizontal and vertical orientations. The ratios serve as an indication of the severity of the blocking artefacts. The two ratios of interest, R_h and R_v are given by:

$$R_h = \frac{\sum_{m=4,8,12} H_m}{\sum_{m=1}^{15} H_m} \text{ and } R_v = \frac{\sum_{m=4,8,12} V_m}{\sum_{m=1}^{15} V_m} \quad (1)$$

where H_m and V_m are the amplitude of the frequency components on the horizontal and vertical principle axis respectively. Note that H_0=V_0=DC. If any of the ratios is large, it suggests the presence of blockiness.

The blockiness is estimated by first checking whether a 32×32 block contains a significant amount of blockiness. If that is the case, the sum of the harmonics in the horizontal, the vertical, or both orientations, is accumulated. This process can be described as

$$B_t = \sum_{\text{All } i}\left[\alpha \times \sum_{m=4,8,12} H_{m,i}\right] + \left[\beta \times \sum_{m=4,8,12} V_{m,i}\right] \quad (2)$$

where B_t is the accumulated harmonics strength and i is the index of each 32×32 block. $H_{m,i}$ and $V_{m,i}$ are the amplitude of the horizontal and vertical m^{th} frequency component in the FFT spectrum of the i^{th} block respectively. The parameters α and β are simply two multipliers whose value depend on the result of comparing the ratios R_h and R_v against a threshold ζ, such that $\alpha = 1$ if $R_h \geq \zeta$, else $\alpha = 0$; and $\beta = 1$ if $R_v \geq \zeta$, else $\beta = 0$. After testing all blocks in a picture and accumulating the strength of harmonics from those blocks having $R_h \geq \zeta$ and/or $R_v \geq \zeta$, the result of summation B_t represents the quantity of blockiness in that picture. Figure 2 depicts a very blocky picture, and Figure 5(a) illustrates the strength of the blockiness for each 32×32 block, with the darkness of the block proportional to the magnitude of blockiness. Comparing the two figures will reveal that there is a high correlation between the blocky areas in Figure 2 and the amount of blockiness reported in Figure 5(a).

Unfortunately, some contextual details in a picture may produce very strong harmonics resembling the signals resulted from blockiness. For example, the radiator of the car in the picture shown in Figure 2 has produced very strong harmonics, causing the model to misinterpret that area as blocky. Had the reference picture been available, this problem may be alleviated by means of cancelling the common luminance gradient between the reference picture and the compressed picture. To make the model work without the reference picture, a technique distinguishing the radiator pattern from blockiness must be developed.

Figure 2 : A blocky MPEG picture. Serious blockiness can be observed over the human body, but very little occurs in the radiator area of the car.

3. PHASE OF THE HARMONICS

As shown in Figure 2, the contextual details can sometime confuse the model if its luminance gradient happens to generate strong harmonics in the frequency domain. This problem must be overcome to prevent false blockiness detection. Since the DCT block boundaries in MPEG picture are fixed, the harmonics generated by blockiness will have well-defined phase. In contrast, contextual details will have random phase harmonics due to the random position of the details. This suggests the use of phase information of the harmonics for verifying the presence of blockiness.

To form a rectangular signal in the spatial domain, the harmonics need to meet some requirements, both in their amplitudes and phases. According to Fourier analysis, a rectangular wave comprises of cosine waves of different frequencies and amplitudes, at certain phase. The rectangular wave $G(x)$ can be decomposed into sum of cosine waves:

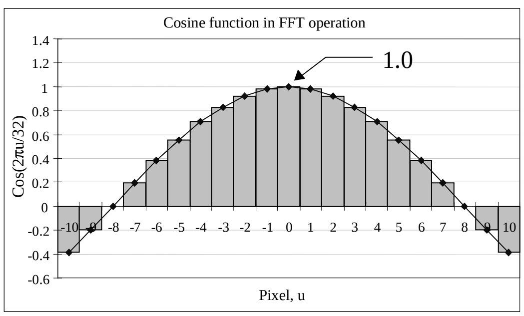
$$G(x) = H_4 Cos 4\omega_0 x + H_8 Cos 8\omega_0 x + H_{12} Cos 12\omega_0 x \quad (3)$$

where ω_0 is the angular frequency of a cosine wave having one complete cycle within a 32-pixel window, and H_4, H_8, and H_{12} are the amplitudes of the first, second and third harmonics. In this case, to form a rectangular wave, the phase of the harmonics must be zero. Therefore if a 32×32 block is contaminated by blocking artefacts, we expect to see the phase of the three harmonics being zero. In other cases, the phase of the harmonics should be random, having uniform probability between 0° and 360°.

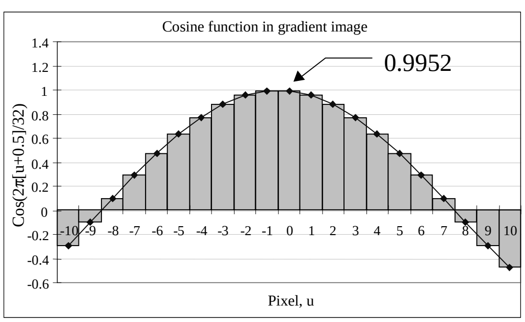
The actual phase angle of the harmonics will not be zero, however. This is due to a slight difference between the gradient image and the FFT operation in defining the fundamental cosine function. In the FFT operation, the definition of the fundamental cosine function (i.e. a cosine function having one complete cycle within the FFT window, with phase offset of 0°) is as shown in Figure 3(a). Here we are assuming that the FFT operation spans over a window of 32 pixels. The sample value of the first pixel of the fundamental cosine function in the FFT window (pixel $u = 0$) follows the value of the cosine function at $u = 0$ (i.e. $Y(u = 0) = Cos(2\pi u \div 32) = 1.0$). However, for the gradient image, there is a small offset in the fundamental cosine function. As illustrated in Figure 3(b), the sample value of the first pixel of the fundamental cosine function in the FFT window (pixel $u=0$) follows the value of the cosine function at $u=0.5$ (i.e. $Y(u=0)=Cos(2\pi[u+0.5]\div 32)=0.9952$). The consequence of this small difference in the definition is a small phase offset when the fundamental cosine function in the gradient image is transformed into the frequency domain by the FFT operation. In other words, if a cosine function $Cos(2\pi u \div 32)$ in the gradient image is transformed by the FFT operation, it will yield

$$\begin{aligned} Cos(2\pi[u+0.5]\div 32) &= Cos(2\pi u \div 32 + [2\pi \times 0.5 \div 32]) \\ &= Cos(2\pi u \div 32 + \phi) \end{aligned} \tag{4}$$

where $\phi = \pi \div 32$ radians. That means the fundamental cosine function from the gradient image is transformed to a cosine function with a phase offset of ϕ radians.



(a)



(b)

Figure 3 Different definition of a complete cycle cosine function over a 32-pixel window, with zero phase offset, (a) Definition in FFT operation, (b) Gradient image.

In Equation (4), we show the phase offset for the fundamental cosine function to be $\pi \div 32$ radians, or 5.625°. For cosine function of other frequency, the phase offset introduced by the FFT process can be estimated as follows. Assume that the cosine function $Cos2\pi Fu$ is to be Fourier transformed. The spatial frequency of the cosine function is $F=\frac{n}{w}$ cycles per pixel, where n is the number of cycles in the FFT window of width w pixels. Due to the difference in definition, the transformed cosine function will become

$$Cos(2\pi F[u+0.5]) = Cos(2\pi Fu + \pi F) \tag{5}$$

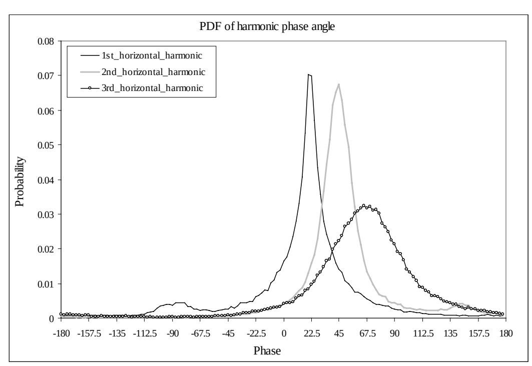
The phase offset is πF radians. Therefore, for a cosine signal of n cycles per FFT window (abbreviated as *cpw* hereafter), the phase offset in degrees will be

$$\phi_n = \frac{\pi F \times 360°}{2\pi} = \frac{n \times 180°}{w} \tag{6}$$

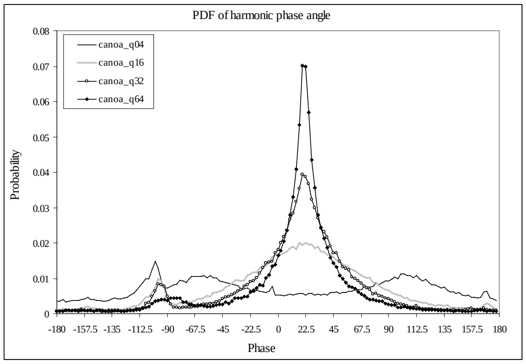
From Equation (6), it is evident that if the window over which the FFT is performed is small, the phase shift can be significant. For example, for $w=32$, the phase shift for the AC component of frequency n cpw will be

$$\phi_n = \frac{n \times 180°}{32} = n \times 5.625° \tag{7}$$

Recall that the harmonics generated by the blockiness are n=4, 8, and 12. The corresponding phase shift for each harmonic will then be 22.5°, 45°, and 67.5° respectively. In other words, if the three harmonics are generated by the blockiness, they should have the phase relationship as stated.



(a)



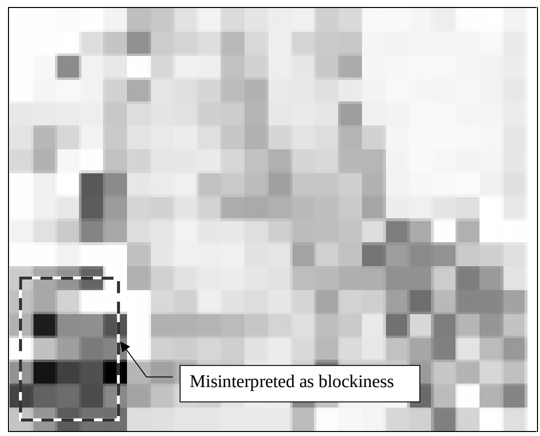
(b)

Figure 4 (a) PDF of phase of first, second and third horizontal harmonics from a blocky picture, (b) PDF of harmonics phase for different amount of blockiness.

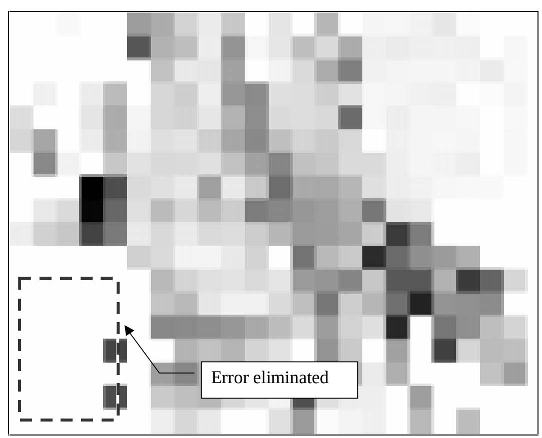
Figure 4(a) shows the PDF of the phase of each horizontal harmonic derived from all the 32×32 blocks of a very blocky picture. The picture was generated using fixed quantization step size of 64. The plot shows explicitly high probability at angle 22.5° for the 1st harmonic, 45° for the 2nd harmonic, and 67.5° for the 3rd harmonic, with the 1st harmonic having the smallest standard deviation. Also interesting is that the blockier the picture is, the smaller the standard deviation in the PDF will be. This is illustrated by Figure 4(b), where the PDFs of the 1st horizontal harmonic's phase of I-frames coded at different quantization step sizes (mquant=4, 16, 32 and 64) are compared. When there is no blockiness (mquant=4) in the coded picture, the PDF is virtually uniform. Sequences having different scenes have been tested and they all more or less show the same property.

The implication is that the phase of the harmonics can be used as one of the criteria in verifying the presence of blockiness in the picture. The harmonics generated by blocking artefacts will have large R_h and/or R_v ratios, and the phase of the harmonics will be close to the desired values. Safeguarding with both the amplitude and phase checking on the harmonics, very limited contextual details may be erroneously interpreted as blockiness. Since the detector can now distinguish between contextual details and blockiness, no reference pictures may be needed.

From Figure 4(a), it is obvious that only the 1st and 2nd harmonics have substantially narrow PDF. We therefore employ only these two harmonics when checking the phase. So for a 32×32 block of the gradient image, blockiness is present if and only if R_h and/or R_v is larger than a threshold, and the phase of the harmonics is within a certain tolerance (we choose ±15°) from the desired value. The advantage of supplementing the harmonic ratio checking with the phase information is demonstrated in Figure 5. Figures 5(a) and 5(b) illustrate the degree of blockiness estimated using the harmonics analysis method, without and with the phase information respectively. With the help of the phase checking, the error over the radiator area has been corrected.



(a)



(b)

Figure 5(a) Blockiness indication, with the brightness of each 32×32 block inversely proportional to the severity of blockiness in that block. Notice that the radiator area has been misinterpreted as blocky. (b) The mistake at the radiator area has been eliminated.

4. PERFORMANCE AND LIMITATIONS

Pearson correlation of 0.88 for I-frames has been obtained between subjective data and the prediction from this model. Since the phase measurement is very sensitive to any spatial offset of the DCT block boundaries, this model may not be optimal for motion compensated frames (P- and B-frames). Figure 6 shows the PDFs of the phase of the 1st horizontal harmonic for I-, P- and B-frames in a coded sequence. The variance in the phase of the harmonic for B-frames is much larger than I- and P-frames. The B-frames thus pose some problem for the phase checking process. If the phase tolerance is set too loose, the problem of misinterpreting contextual details as blockiness will remain. On the other hand, setting the window too tight is also not recommended as this may result in underestimation of blockiness. Therefore this method is not suitable for B-frames. For P-frames, the PDF approximates the I-frames' PDF. Our experiment showed that for P-frames, the Pearson correlation of 0.85 could be achieved. Since it is sensible to sample-check only the I- and P-frames in many applications, this model can be a useful tool for estimating blockiness in MPEG video.

5. CONCLUSIONS

In this paper, we have demonstrated how the harmonics generated by the luminance gradient resulted from discontinuities across DCT blocks may be exploited for measurement of blockiness. The strength of the harmonics with respect to that of other AC components is a good indication of the probability of presence of blocking artefacts. The phase information of the harmonics helps to make the blockiness detection method more robust and overcome the problem of misinterpreting contextual details as blockiness. Using this new method, a single-input blockiness detection model is possible, without the need of the reference pictures
.

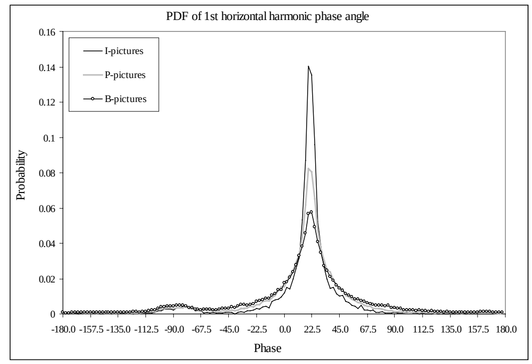


Figure 6 PDFs of the 1st horizontal harmonic phase for I-, P- and B-pictures in an MPEG video.

6. REFERENCES

[1] Lauterjung, J., "Picture quality measurement", IBC-98, 1998, pp. 413-417, 1998.

[2] Baroncini, V. A., "Automatic visual quality control in digital TV services", IBC-98, 1998, pp. 425-430, 1998.