Benchmarking Video Codecs for Blockiness Compression Artefact

A. Punchihewa, J Armstrong

Institute of Information Sciences and Technology, Massey University, Palmerston North, 4410, New Zealand.

Email: g.a.punchihewa@massey.ac.nz

Abstract

This paper presents a method to benchmark video codecs objectively for the blockiness artefact. A static test pattern previously designed to highlight blockiness artefact of image codecs was applied to video codecs under test to evaluate the performance when intra-frame coded mode is used. The performance of the video codec in inter-frame coded mode is evaluated using an image sequence generated using a similar function used to generate the still test pattern. A software codec having two implementations, namely MPEG-2 and MPEG-1 was configured at different bit rates. At low bit rates, the codec tested as MPEG-1 performs better than the MPEG-2 with respect to the blockiness distortion. The test sequences and the blockiness distortion measure algorithm are useful tools to evaluate codecs for blockiness artefact. Different types of video codecs or different implementations of similar type of video codec can be benchmarked for performance against blockiness distortion using the tools developed.

Keywords: Video codecs, blockiness artefact, benchmarking, test patterns, compression artefacts, MPEG-1, MPEG-2

1 Introduction

Advances in computing and communication technologies have led to common use of digital media content such as pictures and video. However, digital images and video are still demanding in terms of storage space and transmission bandwidth. Lossy compression is essential to manage the raw image and video data before they are stored or transmitted. The image and video codecs use lossy compression techniques such as quantisation which will introduce distortions known as artefacts. Blockiness is a common and dominant form of compression artefact in decoded images and video. References [1 to 6] address image artefacts and objective evaluation of image coding artefacts. Therefore, this paper concentrates on video artefacts only.

Video codecs such as MPEG uses JPEG like coding to compress key frames or I frames. I frames are the intra-frame coded frames that are generated at regular frame interval known as group of pictures (GOP). There are two other types of coded frames known as forward predicted frames and backward predicted frames. They are based on frame differences in successive frames and either forward predicted motion vectors or bi-directionally predicted motion vectors are used. The longer the GOP, the greater the reduction in bit rate as the process takes the temporal redundancies into account. This research intends to investigate the impact of GOP on the blockiness distortion over individual frames.

2 Methodology

2.1 Motivation

Many researchers have developed blockiness artefact mitigation techniques which can be applied as pre-processing or post-processing to a standard codec. Codecs such as MPEG-4 has optional features to pre-process video to mitigate blockiness artefacts. Usually, such pre or post-processing technique trades-off one form of artefact to another. Gao et al. proposed a de-blocking algorithm [7] as a post-processing algorithm. The technique proposed is called semaphore extraction which attempts to reduce the blur due to post-processing of image boundaries as block boundaries. It is difficult to evaluate the effectiveness of such techniques proposed as the subjective methods are expensive and inconsistent though they are the ultimate measure. An objective measure based on a known reference would provide an accurate and repeatable measure for blockiness artefact.

The video codecs such as MPEG defines and standardise the syntax for the decoder only. Hence many implementations are possible for an encoder. Objective distortion measure tools for artefact will provide an accurate and convenient method to evaluate codecs for comparison. Tan et al. proposed an algorithm [8] to detect and measure blockiness based on the amplitude and phase information of the harmonics generated by the blockiness artefact in
MPEG-2 coded video. The phase information of the harmonics was used to eliminate image boundaries being detected and measured as block boundaries.

2.2 Design of Test Patterns and Metric

Blockiness is the false block structure that would appear in reconstructed video. If a test pattern is generated with no perceivable edges in uncompressed sequence, then each frame would enables the evaluation of blockiness accurately in the coded sequence.

2.2.1 Design of test Pattern

Design of a test pattern as a static pattern with no boundaries is described in detail in the reference [4]. The same test pattern was animated over a virtual image space and then captured at the size of the frame required for the codec. The test sequence generator algorithm reads test image data from the capture window having image size equal to the frame size) and then is stored as an uncompressed test sequence. When codec under test to be stressed, test sequence is read from the storage to the codec input. A frame from the test pattern/sequence is shown in figure 1.

Figure 1: A frame prior to compression from the test pattern/sequence used to evaluate blockiness artefact due the codec under test.

It can be observed that in the above frame there are no visible boundaries. The spatial distribution of the intensity of each of the frame in the test sequence has been defined as a raised cosine. The intensity of each of the pixel is defined so that, in the direction of x-axis or in the direction of y-axis, the intensity gradients are less than 2% of the maximum value of the intensity; in this case 8 bits were used hence on a scale of 255.

2.2.2 Design of blockiness algorithm

This research is an extension of work done previously for still image codecs. Some of the findings are published in references [2 to 6]. This work re-uses the same blockiness algorithm presented in reference [4] to evaluate the level of blockiness in the reconstructed test patterns/sequences. The figure 2 shows a reconstructed frame from the test sequence. The block boundaries are clearly visible and original frame did not have any of those boundaries.

Figure 2: A decoded frame from the test pattern/sequence used to evaluate blockiness artefact after compression.

The figure 3 shows distribution of blockiness measure over the complete frame. The modified algorithm calculates the vertical and horizontal intensity gradients in the reconstructed test pattern, then average of blockiness for each pixel. When the gradient values are greater than pre-defined threshold of 2% of the maximum representable original value, which is based on an informal subjective test, that blockiness value is classified and counted separately towards the calculation of global average for the complete frame considered.

Figure 3: The histogram of blockiness of each of pixels for the MPEG-1 codec.
3 Results and Discussion

Experiments were carried out to evaluate the blockiness introduced by each codec MPEG-1 and MPEG-2 when intra-frame coded is used. Figure 4 and figure 5 show the variation of average blockiness over the complete frame when intra-framed codec under different quality factors. In figure 3 it can be observed that the individual blockiness or gradients vary between 1 and 7. Gradient of 4 and above does not make any perceivable distortion as they are small gradient though large in numbers in the total population.

![Figure 4: Average blockiness vs. Quality factor for MPEG-1 codec.](image1)

![Figure 5: Average blockiness vs. Quality factor for MPEG-2 codec.](image2)

3.1 The Impact of GOP on average blockiness

The test pattern/sequence was compressed using many GOP structures such as I, I, …, or I, B, P, B or I, P I, P, so on. Figure 6 and figure 7 are representative plot of average blockiness for GOP of I, B, P, B. This GOP structure was applied to both MPEG-1 and MPEG-2 codecs. The average blockiness was measured on each frame for the sequence. The three types of frames, namely, Intra-frame coded frame I, forward predicted frames P and the bi-directional predicted frames B showed a similar pattern over the complete test sequence. The highest average blockiness resulted from I frames, then P frames and B frames had the lowest value. This is due to the fact that P frames are generated using I frames and B frames are generated using P and I frames resulting averaging or low pass filtering.

![Figure 6: Average blockiness for MPEG-1 codec with GOP of IBPB.](image3)

![Figure 7: Average blockiness for MPEG-2 codec with GOP of IBPB.](image4)

In general MPEG-1 resulted in more average blockiness compare to MPEG-2. I frames always had the highest and fairly constant level of blockiness.
3.2 MPEG-1 vs. MPEG-2

When video codecs are to be benchmarked, they need to be compared at a given bit rate. Experiments were carried out to check the level of average blockiness induced by each MPEG-1 and the MPEG-2 codecs under different bit rates but for the same test sequence. Figure 8 shows the results of the experiment. Around 1750 kbps, performance curves cross each other.

Figure 8: A comparison between MPEG-1 and MPEG-2 codecs for average blockiness on the same test sequence.

For bit rates below 1750kbps, the MPEG-1 performs better. For high bit rate applications MPEG-2 introduce less average blockiness. In this case with the given test sequence, if the video bit rate is higher than 1750kbps, MPEG-2 codec outperforms MPEG-1 against the blockiness distortion. MPEG-1 was designed for CD-ROM applications. Hence it has not optimised to handle bit rates above 2Mbps where as MPEG-2 being used for broadcast application, it can handle well the bit rates above 2Mbps.

4 Conclusions

The test patterns developed for still image codecs can be easily extended for video codecs evaluations. For high bit rates, MPEG-2 outperforms MPEG-1 codec. The blockiness algorithm and test sequences are useful tools for codec evaluations against blockiness artefact distortion. Further improvements can be done to the test sequence by identifying the limitations of the animated test sequence. The blockiness algorithm and the test sequence can be used to benchmark the different implementations of the same type of codecs such as MPEG-2 as well.

5 Future work

Future work involves comparing similar type of codecs. Recently acquired state of the art codec for broadcast applications, namely a MPEG-4 codec, will be benchmarked against the available MPEG-2 codec.

References