Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 17, 2016

T. Davies
Cisco

March 16, 2016

Quantisation matrices for Thor video coding
draft-davies-netvc-qmtx-00

Abstract

This draft describes a family of default quantisation matrices that may be used to improve perceptual quality when encoding with Thor. Similar quantisation matrix designs may be used in most block-based video and image codecs.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This document describes a family of default quantisation matrices that may be used to improve perceptual quality when encoding with Thor. The quantisation matrices are designed to be near-flat at high quantisation levels and more strongly profiled at low quantisation levels, to avoid ringing artefacts and better shape quantisation error across a whole sequence with varying quantisation levels.

2. Definitions

2.1. Terminology

This document uses the following terms.

QP: quantisation parameter

QM: quantisation matrix

CSF: contrast sensitivity function

BDR: Bjontegaard Delta-Rate

3. Quantisation matrix design

3.1. The function of quantisation matrices

Quantisation matrices work by shaping the residual error after quantisation in the spatial frequency domain, usually the DCT domain. This is done by varying the quantisation factor applied across spatial frequencies in the transform block. Typically a high quantisation factor is applied at high spatial frequencies and a low one at low spatial frequencies.

The aim is roughly to match a Contrast Sensitivity Function for the human visual system. This provides a curve of sensitivity to detail (and therefore coding errors) with spatial frequency. Given known resolutions and assumed viewing distances, a weighting function can be simply defined for all the coefficients in a transform block.

This simple approach is complicated, however, by a number of factors. The first is the CSF is in reality not a simple function of spatial frequency, but depends on factors such as brightness which are imperfectly corrected for by television gammas. There is little that can be done about that in the quantisation matrices themselves, but adjusting QP itself may help.

The second factor is that CSFs are determined experimentally based on models of Just Noticeable Difference (JND) and do not reflect so well the impact of distortions well above this level. Adjustments at high levels of quantisation are needed to reflect this.

Finally, applying quantisation matrices to video is affected by the fact that most frames are predicted and the QM is applied to the residual after prediction. This means that the quantisation error for a block consists of the quantisation error in the reference block, plus any additional error introduced in the current block. These errors will add if they are uncorrelated, but they may well be correlated at high QP.

Despite these difficulties, QMs are widely used and known to work well, and are available in video coding standards such as H264/AVC and H265/HEVC [AVC,HEVC].

3.2 Quantisation matrix design in AVC and HEVC

Quantisation matrices are available in a number of different codecs. The design in AVC and HEVC is to provide default matrices together with the ability to signal bespoke matrices [AVC,HEVC]. These matrices must cover all the different transform block sizes, components (Y, Cb, Cr) and intra and inter frame or block types, with fall-backs defined if bespoke matrices are not provided. Default inter block matrices are flatter than intra matrices, no doubt because of the noise-addition effect described in section 3.1: if they had the same profile as for intra then the overall profile of the combined prediction + residual could be over-shaped.

3.3 Design of quantisation matrices in Thor

Thor provides a set of matrices for each component of 420-sampled video, for each block size and each quantisation parameter. The principles behind the design are as follows:

1) QP dependence. Matrices become flatter as quantisation levels increase

2) Energy preservation for intra. The inverse quantisation matrices for intra blocks are normalised to approximately preserve energy of the residual

3) DC preservation for inter. The inverse quantisation matrices for inter blocks are normalised to preserve the DC level

4) Matrices are also flatter for inter blocks than for intra blocks.

5) Quantisation matrix strength is globally adjustable

The QP dependence takes account of a number of factors. Firstly it reflects that inter blocks typically have higher QPs than the blocks used to predict them. This means that flattening the matrices at higher QP naturally prevents over-shaping the quantisation error.

Secondly, the high-QP flattening process also reflects the fact that errors at this level are very visible even at high spatial frequencies. Strong error-shaping at these QP levels leads to very visible additional ringiness.

SSIM-based metrics [SSIM,MSSSIM,FASTSSIM] indicate that preserving image variances and therefore residual energies is perceptually important. This is feasible for intra where residuals are substantial but in the case of inter it is also important to preserve DC levels since getting these wrong can produce very visible artefacts.

Intra frames tend to have lower QP than inter frames, and this means that QP dependence absorbs most of the requirement for inter matrices to be flatter than intra matrices. However inter matrices are still a little flatter, to take account of the different characteristics of intra and inter blocks within the same frame.

In determining the quantisation matrix, there are 12 possible sets available giving a new set of matrix for each change of approximately 4 in quantisation value. Thor also supports a global adjustment or strength parameter, which offsets the LUT mapping quantisation parameter to quantisation matrix set. This is a value from -32 to 31. A value of -32 will reduce the qp used by 32, increasing the strength of quantisation matrix dramatically. Likewise a value of 31 will eliminate quantisation matrices for all but the smallest QPs.

The effect of the ability to signal strength, and the provision of a range of QP-dependent matrices are intended to remove the need to signal bespoke matrices at all.

3.4 Implementation

Quantisation matrices are applied as multiplicative factors in forward or inverse quantisation processes. In Thor the basic unweighted dequantisation process for a coefficient c with quantisation parameter q is based on two values: scale[q], which depends only on q%6, and shift[q] which depends only on q/6, the block size and the signal dynamic range. scale[q] takes care of quantisation step sizes which fall between powers of 2 and shift[q] takes care of the basic power of 2 part of the quantisation step.

The formula for unweighted dequantisation is then:

```
c -> (c*scale[q] + (1<<(shift[q]-1))) >> shift[q]                  (1)
```

for positive shift[q], otherwise

```
c -> (c*scale[q])<<(-shift[q])                                     (2)
```

To apply a matrix M to a coefficient c[i,j] at position (i,j) within a block, the formulae (1), (2) change to:

```
c[i,j]->(c[i,j]*M[i,j]*scale[q]+(1<<(shift[q]+5)))>>(shift[q]+6)   (3)
```

if shift[q]+6 > 0, otherwise

```
c[i,j]->(c[i,j]*M[i,j]*scale[q])<<(-shift[q]-6)                    (4)
```

otherwise.

Exactly complementary formulae can be derived for the forward quantisation process.

4. Compression performance

Although largely a visual tool, the effectiveness of QMs can be inferred by changes to PSNRHVS [PSNRHVS] and FASTSSIM metrics. FASTSSIM tends to over-estimate gains a little, as it has a bias towards low-pass filtering. Overall BDR results for the Low-Delay B (LDB) and High-Delay B GOP 16 configuration (HDB16) are as follows (QPs 22, 27, 32, 37):

Config	PSNR	PSNRHVS	FASTSSIM
LDB	+1.1%	-3.3%	-9.0%
HDB	+2.2%	-2.6%	-11.6%

These were computed on the same test sequences as in IRFVC.

FASTSSIM and PSNRHVS gains are typically larger, and PSNR losses smaller, for higher resolution material.

5. Informative References

[AVC] ITU-T Recommendation H.264, "Advanced video coding for generic audiovisual services", March 2010.

[HEVC] ITU-T Recommendation H.265, "High efficiency video coding", April 2013.

[FASTSSIM] Chen, M. and A. Bovik, "Fast structural similarity index algorithm", 2010, <http://live.ece.utexas.edu/publications/2011/chen_rtip_2011.pdf>.

[MSSSIM] Wang, Z., Simoncelli, E., and A. Bovik, "Multi-Scale Structural Similarity for Image Quality Assessment", n.d., <http://www.cns.nyu.edu/~zwang/files/papers/msssim.pdf>.

[PSNRHVS] Egiazarian, K., Astola, J., Ponomarenko, N., Lukin, V., Battisti, F., and M. Carli, "A New Full-Reference Quality Metrics Based on HVS", 2002.

[SSIM] Wang, Z., Bovik, A., Sheikh, H., and E. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity", 2004, <http://www.cns.nyu.edu/pub/eero/wang03-reprint.pdf>.

[IRFVC] Davies, T. "Interpolated reference frames for video coding", IETF draft https://www.ietf.org/id/draft-davies-netvc-irfvc-00.txt

Author's Address:

Thomas Davies
Cisco
Feltham
UK

Email: thdavies@cisco.com