Paper

Benchmarking image codecs by assessment of coded test images: the development of test images and new objective quality metrics

Amal Punchihewa, Donald G. Bailey, and Robert M. Hodgson

Abstract— Objective quality measures are required for benchmarking codec performance. Our aim was to develop a simple, accurate method capable of rapidly measuring the degree of blockiness, edge-blur and ringing due to image compression. Two test images were designed to emphasise these artefacts. The efficacy of the new metrics is demonstrated using a JPEG codec at a range of compression levels.

Keywords— image quality, artefacts, subjective, objective, coding, metric, blockiness, blur, ringing.

1. Introduction

Lossy image and video compression codecs introduce many types of distortions known as artefacts. *The Digital Fact Book* defines artefacts as "particular visible effects, which are a direct result of some technical limitation" [1]. Artefacts are generally not evaluated by traditional methods of signal evaluation. For instance, the visual perception of contouring in a picture cannot be related to signal-to-noise ratio [1].

In multimedia communications, image and video are the dominant components. With limited communication bandwidth and storage capacity in terminal devices, it is necessary to reduce data rates. High levels of compression result in undesirable spurious features and patterns in the reconstructed image; these are the artefacts defined above. Image compression schemes such as JPEG use the techniques of discrete cosine transform (DCT), block processing and quantisation. This may result in blockiness, edge-blur, contouring and ringing artefacts in coded images. The following table summarises these artefacts.

When the original signal is not fully known, quantifying these artefacts is difficult. In particular, it is difficult to isolate the individual components listed in Table 1.

Image codec development, parameter tuning and benchmarking all require availability of more accurate and swift measurements. Subjective assessment can provide an accurate indication of perceptual quality but such methods are very time consuming [3]. Traditional full referenced metrics such as mean square error (MSE) and peak signal to noise ratio (PSNR) do not always correlate well with perceptual quality, and are unable to distinguish between different types of artefacts [3].

Researchers have developed objective quality metrics for different artefacts based on non-referenced or reduced reference techniques [3–5]. They are good for in-service measurements and estimates, as they are not as accurate as full-referenced methods. Bailey *et al*. proposed a non-referenced, objective, quality metrics for blockiness based on edge activity of reconstructed images [4].

Table 1
Summary of common artefacts found in digital image and video systems [2]

Artefact	Description
Blockiness	Distortion of the image characterized by the appearance of an underlying block structure.
Edge-blur	Distortion, characterized by reduced sharpness of edges.
Ringing	Appears as echoes of the hard edges in the picture or a rippling adjacent to step edges.
Contouring	Visibility of bands of intensity over large regions.

If the original image is unknown it is often difficult to determine the presence and extent of artefacts. Therefore the approach in this paper is to use the full referenced method using synthetic images having known spatial distributions of pixel values designed to emphasise the artefacts to be assessed. This study is concentrated primarily on three of the most common coding artefacts, namely blockiness, edge-blur and ringing. A search of the literature did not reveal any full-referenced objective quality metric and accompanying test images for blockiness, ringing or edge-blur.

2. Methodology

The main aim of this full referenced quality assessment approach was to design and synthesise a few test patterns in which the spatial distribution of pixel values will emphasise artefacts due to codec operation. Many image compressors have a control parameter, the quality factor that can be set by the user to adjust the compression ratio. In general the lower the quality factor the higher the compression ratio and the more visible artefacts become. At low compression ratios, the artefacts may not be obvious to the human eye.

2.1. Definition of quality metrics

2.1.1. Blockiness

Blockiness is the distortion of the image characterised by the visibility of the underlying block encoding structure [4]. Some codecs, such as JPEG, divide the image into a number of small blocks which are then processed independently. As there are no constraints applied between adjacent blocks, such processing can result in discontinuity in reconstructed pixel values at block boundaries. The visibility of the block encoding structure depends on the magnitude of the discontinuity in the reconstructed image and can be measured horizontally and vertically as pixel intensity difference at block boundaries.

The proposed blockiness objective quality metric is more suitable for codecs complying with the JPEG standard. The proposed objective quality metric assumes a block size of 8×8, the typical block size in JPEG codecs. JPEG 2000 standard has the provision to divide an image into rectangular blocks of the same size called tiles. Each tile is encoded independently. Tile size is a coding parameter that is explicitly specified [6]. This may result in a blocky appearance however is not considered in this research.

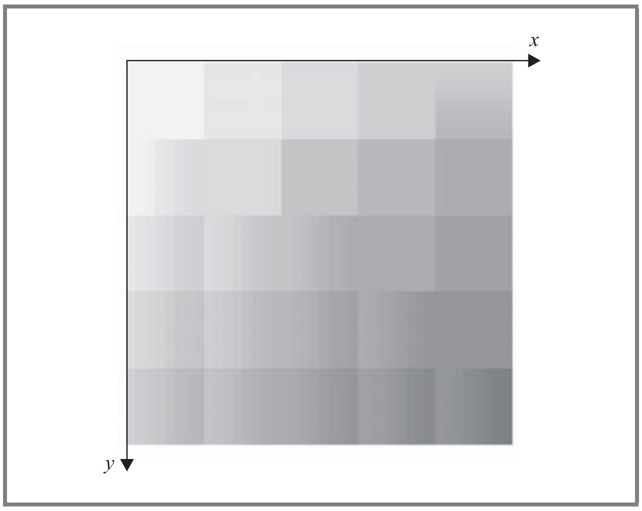


Fig. 1. Example of blockiness resulting from JPEG codec at high compression ratio in the spatial domain.

Blockiness can be expressed as the discontinuity in amplitude per block boundary pixel in the image. The higher the value of the blockiness, the higher the visibility of block structure.

Consider an $M \times N$ image I, reconstructed from a 8×8 block coded image having M rows and N columns. As shown on Fig. 1, both vertical and horizontal edges can be observed at regular pixel intervals of 8 because of the 8×8 block processing. Consider row y, along line y, the horizontal blockiness can be calculated as

$$\sum_x \left| I[x,y] - I[x+1,y] \right| ,$$

where $x = 8, 16, 24, \ldots, (N-8)$. This computation is repeated for all rows from $y = 1$ to M. The total of the vertical blockiness VB can be written as

$$VB = \sum_{y=1}^{M} \sum_x \left| I[x,y] - I[x+1,y] \right| . \tag{1}$$

This results from $\frac{(N-8)}{8} M$ block boundary pixels. Similarly, the horizontal blockiness HB,

$$HB = \sum_{x=1}^{n} \sum_y \left| I[x,y] - I[x,y+1] \right| , \tag{2}$$

results from $\frac{(M-8)}{8} N$ block boundary pixels.

Both the HB and the VB can be combined and normalised by dividing the number of boundary pixels. Hence the blockiness per boundary pixel B can be expressed as

$$\begin{aligned} B &= \frac{HB + VB}{\frac{N-8}{8} M + \frac{M-8}{8} N} \\ &= \frac{4(HB+VB)}{NM - 4(M+N)} . \end{aligned} \tag{3}$$

2.1.2. Edge-blur and ringing

Ringing always occurs at edges and blur generally occurs at edges. Since we are concerned with the blur occurring at an edge, this paper concentrates on the edge-blur rather than a global-blur.

Ringing is an undesirable visible effect around edges. Many codecs transform the pixel values into the frequency domain where the transformed coefficients are then quantised. Quantisation errors resulting from this approach give rise to ringing around sharp discontinuities in the image.

An ideal sharp edge contains components at all frequencies. Any change in the amplitude of any of these components will result in ripples in the image with amplitude corresponding to the error.

As a result of energy compaction in a codec, many of the high frequency components are very small, and get quantised to zero. This loss of high frequency components leads to blur in reconstructed image.

Ringing and edge-blur are defined in Fig. 2. We define the region between the first crossings on each side of the edge transition as the edge-blur region. Outside of this, from the start of the first overshoot on each side, the errors are classified as ringing.

To obtain a measure of edge-blur, consider the shaded area in Fig. 2. The greater the edge-blur, the larger will be the shaded area. By dividing the area by the step height,

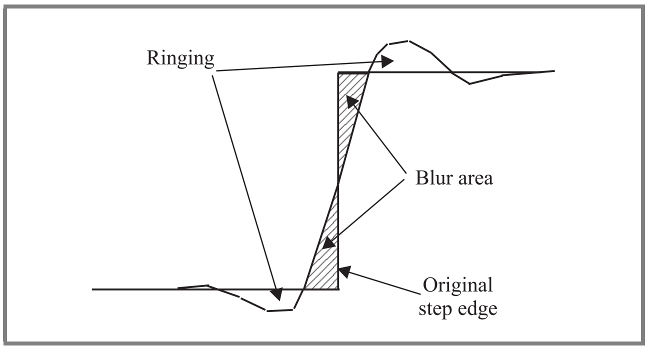


Fig. 2. Ringing and edge-blur at an edge of a one-dimensional signal.

a measure of average edge-blur width can be obtained. In a similar manner, the area between the ringing signal and ideal signal provides a measure of the severity of ringing. With sampled data, an ideal step edge would involve a transition between two pixels, as illustrated by the circles in Fig. 3.

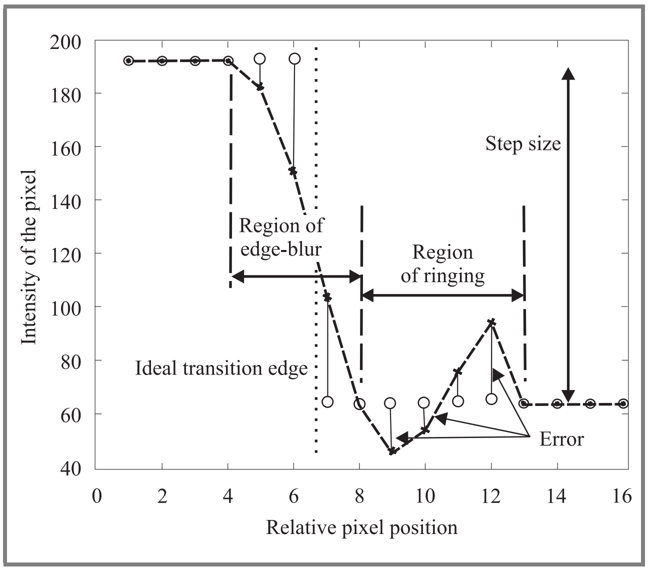


Fig. 3. Edge-blur and ringing for one-dimensional sampled data. Circles represent original pixel value and cross represent reconstructed pixel value.

The crosses in Fig. 3 are the pixel values near the edge of the reconstructed image from a codec. The transition from one intensity to another intensity involves many pixels. The pixel values of reconstructed image outside the region of edge-blur may oscillate around each intensity level of pixels of the original edge.
The edge-blur and ringing are therefore quantified as

$$\text{edge-blur} = \frac{\sum_{\text{blur\_region}} |\text{error}|}{\text{step size}}, \tag{4}$$

$$\text{ringing} = \frac{\sum_{\text{ringing\_region}} |\text{error}|}{\text{step size}}. \tag{5}$$

In 2D images, edges may appear at any orientation. Therefore we consider edge-blur and ringing perpendicular to the edge under consideration. By summing the Eqs. (4) and (5) over whole image and dividing by the number of edge pixels, we can obtain a measure of edge-blur and ringing per edge pixel.

2.2. Design of the test signals

Two simple synthetic test signals have been designed to emphasise visible edge-blur, ringing and blockiness artefacts. The pixel values and the shape of the pattern have been carefully chosen so that the algorithm could detect coding artefacts completely and adequately.

2.2.1. Blockiness

To generate and measure the blockiness artefact, it is necessary to have a test image without edges that results in edges at block boundaries after reconstruction. To produce such edges it is therefore necessary to have an intensity gradient within the test pattern. A simple horizontal or vertical gradient can not distinguish between edges introduced by block processing due to contouring resulting from too few quantisation levels. Therefore an intensity pattern was selected as shown in Fig. 4. The pixel values vary sinusoidally along a diagonal of the image. If pixel intensity varies linearly, the blockiness at certain compression ratios reduces. Nonlinear variation of pixel intensity of the test image (in form of sinusoidal function along a diagonal), stresses the codec at all compression ratios which is required to emphasise the blockiness artefact.

Fig. 4. Original diagonal test image, size: 66 614 bytes, bitmapped.

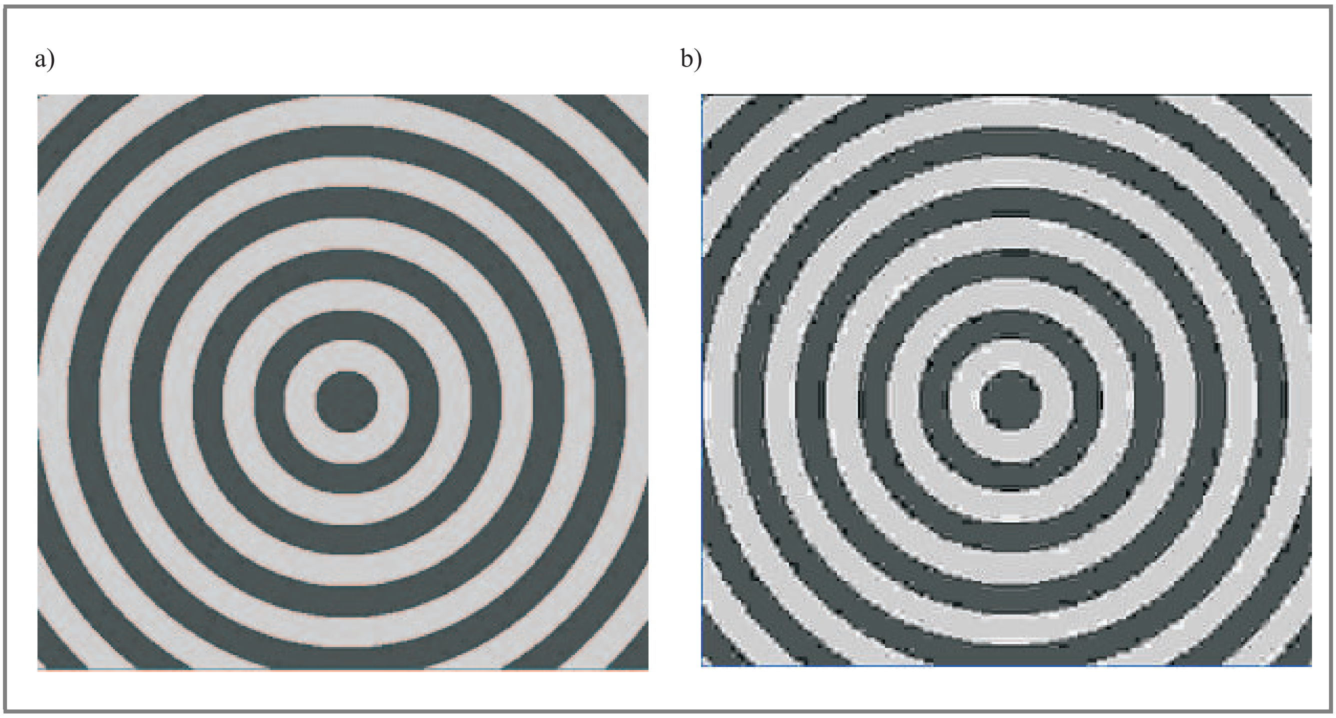


Fig. 5. (a) Original grey scale concentric test image and (b) JPEG reconstructed concentric circles test image with edge-blur and ringing.

Pixel values do not change uniformly within the test image with respect to their neighbours. The blockiness computation algorithm is applied to the error image; that is on the difference between original and reconstructed test image, to prevent the gradient within the original image being measured as blockiness.

2.2.2. Edge-blur and ringing

To test for edge-blur and ringing it is necessary to have step edges within the image. These should include edges of all orientations in order to detect any orientation sensitivity inherent in the codec. A circular pattern contains edges of every orientation. Pixel values of 64 and 192 have been chosen on either side of the boundary, so that after reconstruction there is adequate amplitude margin to allow for ringing in the reconstructed image. To allow for more edges and resulting error pixels, concentric circles have been incorporated (see Fig. 5). The spacing has been chosen as an odd number so that if block processing is used, the edges fall at different places within the blocks.

3. Results

The quality metrics were evaluated by applying them to the test images described in the previous section. The JPEG codec was tested at a range of compression ratios.

3.1. Blockiness

At low compression ratios the blockiness metric is small and increases rapidly with increasing compression ratio as shown in Fig. 6.

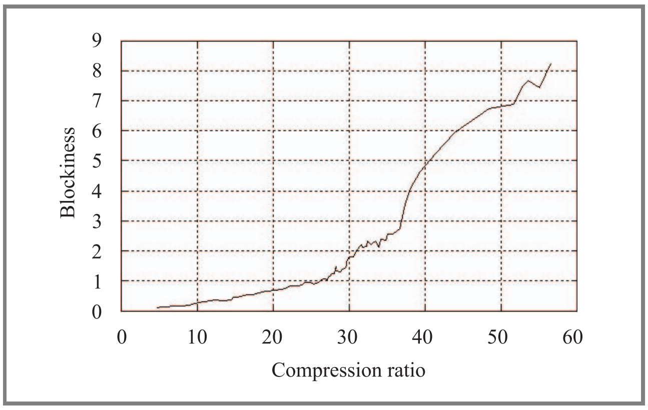


Fig. 6. Blockiness as a function of compression ratio using a JPEG codec on diagonal test image.

It was observed that errors not only occur at block boundaries but in some circumstances in the middle of blocks as well. This occurred at compression ratios of around 30 for this image, resulting in the minor non-monotonic variation seen in the results. This effect was particularly pronounced when a constant gradient image was used because of a threshold effect in quantising the JPEG coefficients.

At some compression levels, errors may actually reduce for higher compression depending on exactly where quantisation levels fall. The sinusoidal variation in the test image means that the different blocks have different gradients, averaging out, and significantly reducing, this effect.

3.2. Edge-blur and ringing

It can be observed that the general trend of ringing and edge-blur is upward with increasing compression ratio (Fig. 7). For the JPEG codec used for the simulations, ringing peaks around compression ratios of 10, 30 and 40.

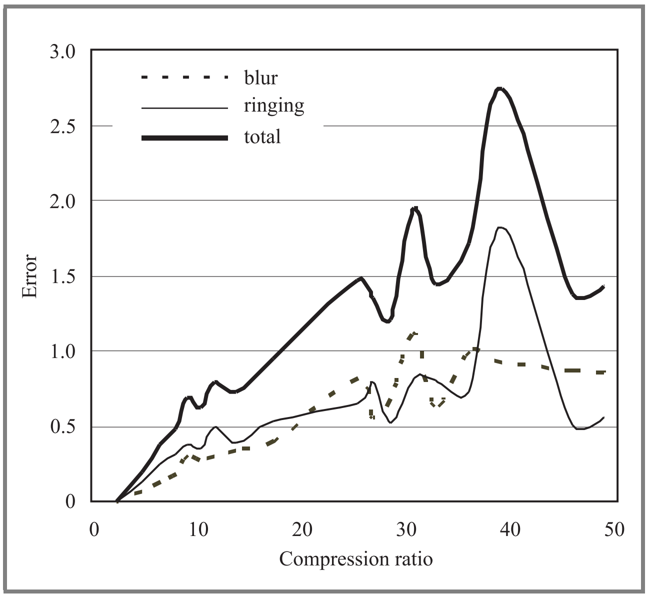


Fig. 7. Edge-blur, ringing metrics and total error as a function of JPEG compression ratio on the concentric circles test image.

These are due to quantisation errors which affect the dc component of the pixel values in reconstructed image around the edge. This has influenced the edge-blur around compression ratios 10 and 30. Edge-blur and ringing decrease above a compression ratio of 40 due to severe quantisation.

4. Conclusions

In this work three new objective quality measures for edge-blur, ringing and blockiness are proposed. The approach is based on known test patterns and measurements of the strength of each in the spatial domain. The quality metrics are good representations of artefacts and are swift in calculation. The proposed measures clearly distinguish between the three artefacts. The diagonal test signals were designed with knowledge of the specific mechanisms and weaknesses inherent in block-based transform coding. However, the concentric circles test image can be used to evaluate blur and ringing produced by any type of codec. The authors intend to perform further research to design test signals for measuring other types of artefacts (global-blur, colour artefacts, contouring) and extending to other types of codecs (JPEG 2000, MPEG, etc.).

References

[1] B. Pank, *The Digital Fact Book*. Berkshire: Quantel Limited, 2002.

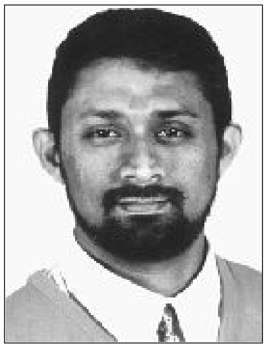
[2] A. Punchihewa and D. G. Bailey, "Artefacts in image and video systems; classification and mitigation", in *Proc. Im. Vis. Comput.*, Auckland, New Zealand, 2002, pp. 197–202.

[3] A. Punchihewa, D. G. Bailey, and R. M. Hodgson, "A survey of coded image and video quality assessment", in *Proc. Im. Vis. Comput.*, Palmerston North, New Zealand, 2003, pp. 326–331.

[4] D. G. Bailey, M. Carli, M. Farias, and S. K. Mitra, "Quality assessment for block-based compressed images and videos with regard to blockiness artefacts", in *Tyrrhenian Int. Worksh. Dig. Commun.*, Capri, Italy, 2002, pp. 237–241.

[5] M. Kusuma and H. Zepernick, "A reduced-reference perceptual quality metric for in-service image quality assessment", in *Proc. Internet, Telecommun. Sig. Proces. WITSP, Conf. & Worksh.*, Coolangatta, Australia, 2003, pp. 72–76.

[6] K. Varma and A. Bell, "JPEG 2000-choices and tradeoffs for encoders", *IEEE Sig. Proces. Mag.*, pp. 70–75, Nov. 2004.



Amal Punchihewa obtained his B.Sc. engineering degree specialising in electronic and telecommunication engineering from the University of Moratuwa, Sri Lanka, in 1986. He received the Master of electronics engineering degree majoring in digital video signal processing from the Technical University of Eindhoven, The Netherlands, in 1991. He served as a computer hardware engineer, as the engineer research and planning at the national TV broadcaster in Sri Lanka, as a faculty member of the Faculty of Engineering at University of Moratuwa. As the head of engineering of the national TV broadcaster in Sri Lanka, he was instrumental in introducing many new technologies and infrastructure developments. He is a Fellow of IEE.
e-mail: g.a.punchihewa@massey.ac.nz
Institute of Information Sciences and Technology
Massey University
Private Bag 11222
Palmerston North, New Zealand

Donald G. Bailey has a B.E. (hons) and Ph.D. in electrical and electronic engineering from University of Canterbury. After spending 2 years applying image analysis techniques to the wool and paper industries within New Zealand, he spent 2 1/2 years as a visiting researcher at the Electrical and Computer Engineering Department at the University of California at Santa Barbara. In 1989, he returned to New Zealand as Director of the Image Analysis Unit at Massey University. In 1998 he moved to the Institute of Information Sciences and Technology, where he is currently senior lecturer and leader of the Image and Signal Processing Research Group.

e-mail: d.g.bailey@massey.ac.nz
Institute of Information Sciences and Technology
Massey University
Private Bag 11222
Palmerston North, New Zealand

Robert M. Hodgson holds a Bachelors degree and Ph.D. in electrical and electronic engineering and was trained in the UK avionics industry. After lecturing in electronic engineering at the University of Hull he joined the University of Canterbury in New Zealand in 1975. He was appointed Professor of information engineering at Massey University in 1988 and shortly after appointed the Head of the Department of Production Technology (Manufacturing and Information Technology). From 1998 to early 2004 he was the Head of Massey's Institute of Information Sciences and Technology. In 2004 he was appointed to be Director of the Massey School of Engineering and Technology.

e-mail: r.m.hodgson@massey.ac.nz
Institute of Information Sciences and Technology
Massey University
Private Bag 11222
Palmerston North, New Zealand