Available online at www.sciencedirect.com
SCIENCE DIRECT
Signal Processing: *Image Communication* 19 (2004) 173–184
SIGNAL PROCESSING: IMAGE COMMUNICATION
www.elsevier.com/locate/image

No reference PSNR estimation for compressed pictures

Deepak S. Turaga*, Yingwei Chen, Jorge Caviedes

Wireless Communications and Networking, Philips Research, Briarcliff Manor, NY 10510, USA

Abstract

Many user-end applications require an estimate of the quality of coded video or images without having access to the original, i.e. a no-reference quality metric. Furthermore, in many such applications the compressed video bitstream is also not available. This paper describes methods for using the statistical properties of intra coded video data to estimate the quantization error caused by compression without accessing either the original pictures or the bitstream. We derive closed form expressions for the quantization error in coding schemes based on the discrete cosine transform and block based coding. A commonly used quality metric, the peak signal to noise ratio (PSNR) is subsequently computed from the estimated quantization error. Since quantization error is the most significant loss incurred during typical coding schemes, the estimated PSNR, or any PSNR-based quality metric may be used to gauge the overall quality of the pictures.

Keywords: PSNR estimation; No-reference quality measurement

1. Introduction

Video and image[1] quality metrics are essential to evaluating the performance of coding and processing algorithms. Often subjective testing with a group of individuals is used to determine the perceived quality of pictures. Such subjective testing is the most accurate in terms of human perception of quality. The methodology for subjective testing has been standardized by ITU [8], as a means to make the tests reproducible and verifiable. However, these tests are expensive in terms of time and the strict setting required.

Objective quality metrics are an alternative to subjective testing. Although they do require accurate subjective base data for training and validating the metric, the subjective testing involved is only done during design and development of the metric. Among the most commonly used objective quality metrics is the peak signal to noise ratio (PSNR). It provides a quality measurement based on the squared error between the original and the processed pictures. Although PSNR has been known to be unreliable especially for enhancement functions, it has been widely used to assess picture quality resulting from compression. Much work has been done in modeling the human visual system (HVS) to better approximate subjective metrics. Work in this domain includes

*Corresponding author. Sony Electronics Inc., Media Processing Division, 3300 Zanker Road, MD: SJ3B10, San Jose, CA 95134, USA. Tel.: +1-408-955-4663; fax: +1-408-955-5724.

E-mail addresses: deepak.turaga@am.sony.com (D.S. Turaga), yingwei.chen@philips.com (Y. Chen), jorge.caviedes@philips.com (J. Caviedes).

[1] In this paper we group together video and images under the generic label pictures.

0923-5965/$ - see front matter
doi:10.1016/j.image.2003.09.001

that by Lubin and Fibush [7] and by Lambrecht and Verscheure [5]. The work by Wolf and Pinson [11] uses these HVS characteristics to introduce a new metric. Another such metric is introduced by Miyahara et al. in [9]. These metrics involve a comparison between the test pictures and the original, or features extracted from the original pictures. However, there are many cases when we need to measure the picture quality in the absence of information about the original, such as during in-service testing or at the user-end.

In this paper we focus on estimating PSNR without reference or access to the compressed bitstream. Some no-reference quality metrics have been introduced previously for discrete cosine transform (DCT) and block based compression schemes. These include measures for the blockiness as introduced by Karunasekera and Kingsbury [3] and ringing artifacts as described by Yuen and Wu [12]. Some work on combining these ringing and blocking metrics to measure quality is proposed by Caviedes and Jung [1]. Despite their effectiveness in characterizing DCT-specific compression artifacts, measures of blockiness and ringing cover only some aspects of quality, and therefore need to be combined with other metrics to measure the overall quality. Furthermore, they are relevant only if the compressed pictures exhibit visible blockiness or ringing artifacts whereas we would like to be able to assess the quality of compressed pictures containing any type of coding distortion.

In this paper we focus on measuring the quality of video compressed with schemes that use DCT and block based coding such as MPEG-1, 2, and 4 and H.261 and H.263, etc. We first estimate the quality for the Intra coded frames and use it to estimate the quality across the sequence, since the quality is likely to be consistent across different types of frames.

We exploit knowledge of the statistical properties of the quantized DCT data to estimate the quantization error. It is well known from literature that the DCT coefficients for video sequences obey a Laplacian probability distribution. Some more details may be obtained from the work by Smoot and Rowe [10]. Quantization schemes used in typical video coding applications are also fairly well defined. One of the typically used quantization schemes in coding algorithms is the one used in the MPEG-2 Test Model 5 (TM5). Knee [4] has examined using statistical properties of data to estimate the quantization error from the MPEG stream. However, due to the lack of reliable estimates of these statistical parameters, and the use of one distribution to categorize all the DCT coefficients, he concludes that the scheme is infeasible. In this paper we allow for separate distributions for each of the 64 DCT coefficients (frequency bands) in an 8×8 block, and describe schemes to estimate these distribution parameters accurately from the quantized data. We may then use these to estimate the quantization error and hence, the PSNR. We also include a brief discussion on using this estimate to obtain more perceptually relevant quality metrics.

This paper is organized as follows. We first provide a brief overview of the coding schemes we consider in this paper in Section 2. We describe the estimation of the quantization parameters in Section 3 and the DCT coefficient distribution parameters in Section 4. We then describe the estimation of the quantization error and PSNR in Section 5. We include some experimental results in Section 6 and conclude in Section 7.

2. Overview of coding scheme

In this paper we estimate PSNR for pictures coded using DCT and block based coding. One such popular coding scheme is the MPEG-2 standard. MPEG-2 has three different kinds of pictures, intra (I), predicted (P) and bi-directionally predicted (B). I frames are encoded using block based DCT (with non-overlapping 8×8 blocks) followed by quantization and entropy coding of the coefficients. P and B pictures use motion estimation and compensation, following which the residue blocks are encoded using DCT, quantization and entropy coding. On the average, P and B pictures require fewer bits to code than I pictures due to the use of prediction. However whenever there is a scene change or a significant change in the video content, the use of prediction is not suitable, and hence I pictures need to be inserted at scene boundaries. Also, P and B pictures are extremely susceptible to error propagation due to the motion compensation. Hence traditional coding schemes introduce I pictures at regular intervals in the coded sequence. All pictures between successive I pictures (including the leading I picture) are labeled a group of pictures (GOP).

We focus on extracting the statistical parameters for the decoded DCT coefficients for the I pictures and using those to estimate the quantization error. We can then compute a measure of quality such as PSNR from this quantization error. In this paper we do not estimate the quantization error for P and B pictures, because motion compensation is not block-aligned, and hence the DCT coefficients of these pictures do not exhibit the clustering structure as observed in I pictures. However, since the coded video quality is likely to be consistent across the different types of frames, we can use the I picture PSNR as a measure of the quality of the sequence.

3. Estimation of compression and quantization parameters

What makes quantization parameter extraction possible is that quantized DCT coefficients cluster around dominant reconstruction levels (DRL) for each DCT frequency-band. Therefore by analyzing the clustering structures of the quantized DCT coefficients we can recover these parameters. As an illustration, we show the histogram of the quantized first AC coefficient of an I picture from the Basket video sequence, coded at 6.5 Mbps, in Fig. 1.

As can be seen from Fig. 1, quantized coefficients cluster around DRLs. However, since our PSNR estimator can access only the decompressed video in the spatial domain, the clustering of quantized DCT coefficients can be observed only for I pictures. For P and B pictures, because motion compensation is not block-aligned, the DCT coefficients do not exhibit the clustering structure as observed in I pictures. The different statistical characteristics of DCT coefficients of I or P coded pictures are illustrated in Figs. 2a–c. Note the horizontal axis (coefficient magnitude) is much smaller than in Fig. 1.

As can be seen from the above figures, I pictures possess unique statistical properties. Therefore, our quantization parameter extraction begins with I picture detection. After an I picture is detected, we extract the quantization matrix used during encoding. Finally, we recover the DCT type (field or frame based DCT) and macroblock level quantization step size.

3.1. Intra picture detection

To detect I pictures, we first find the intra DC precision used during quantization. An intra DC precision of less than 11 (the full precision in

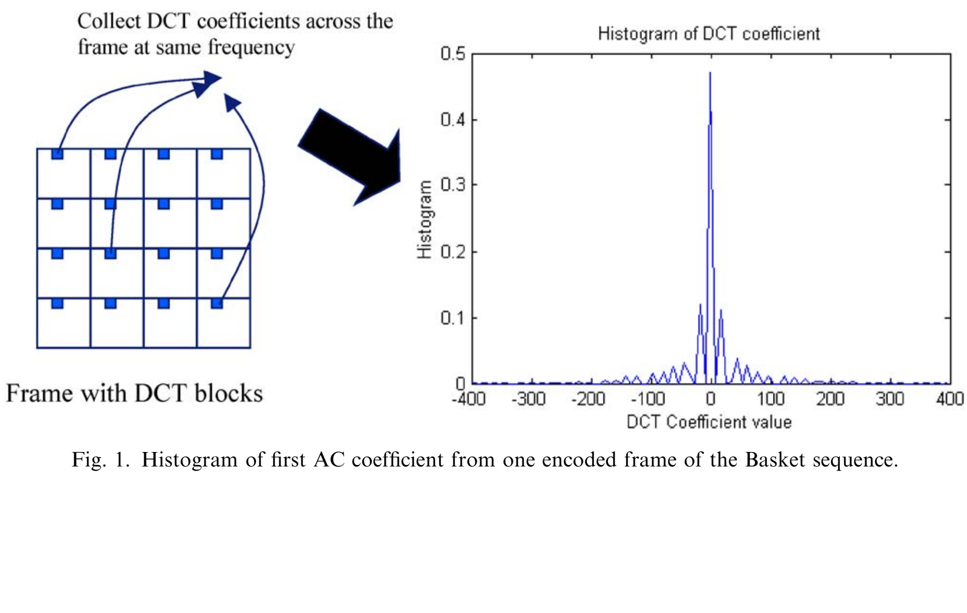

Fig. 1. Histogram of first AC coefficient from one encoded frame of the Basket sequence.

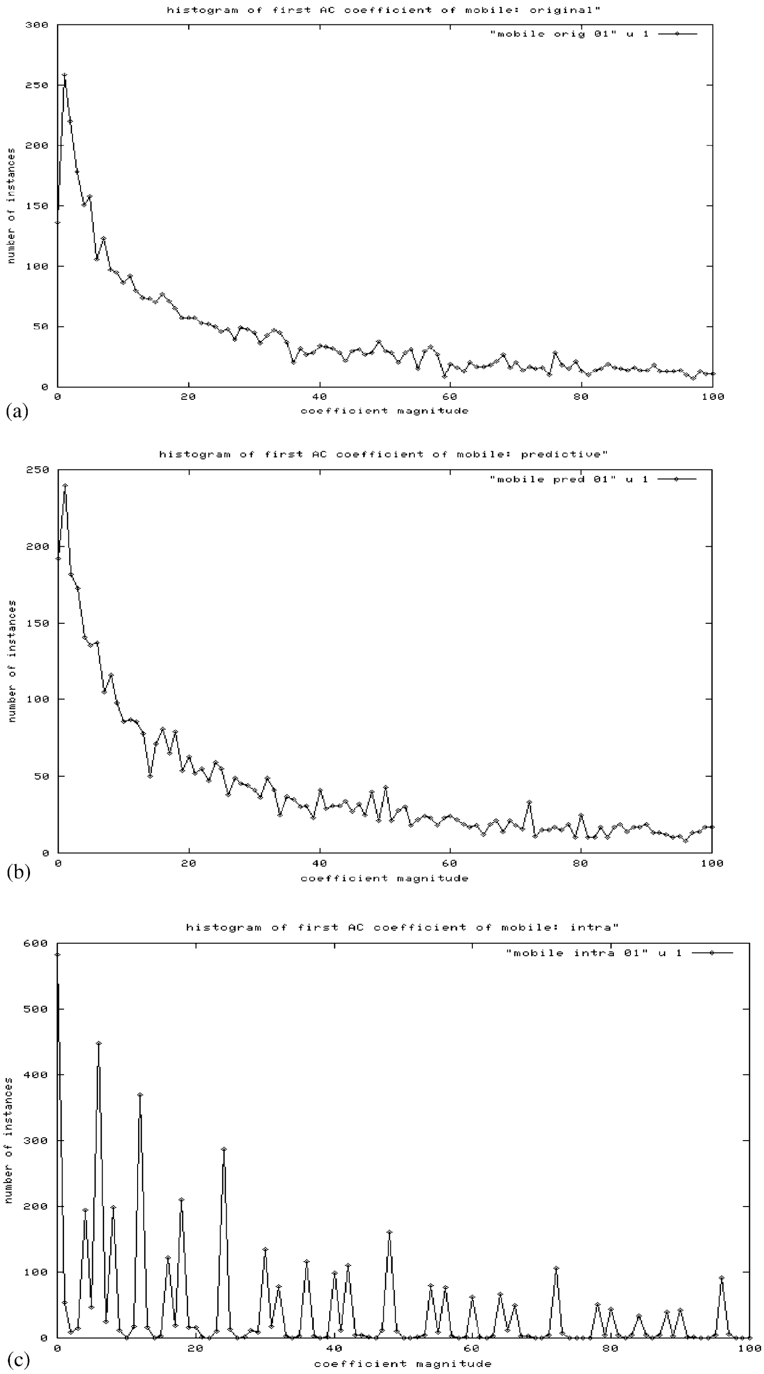

Fig. 2. (a) Histogram of the first AC coefficient of an original frame of mobile. (b) Histogram of the first AC coefficient from a P frame of mobile. (c) Histogram of the first AC coefficient from an I frame of mobile.

MPEG-2) indicates that the picture was intra coded. However, even if the intra DC precision is found to be 11, it is still possible that the picture was intra coded because an intra DC precision of 11 is allowed by MPEG-2. Therefore, if intra DC precision is detected to be 11, then the first AC DCT coefficients are examined to see if the average quantization scale for the whole frame exceeds a low threshold that indicates no quantization or nearly lossless quantization. If yes, then the picture is determined to have been intracoded. Otherwise, the picture is determined to be a P or B picture and the rest of PSNR estimation is skipped until an I picture is found. The procedure for I picture detection is illustrated in Fig. 3.

The GOP size is determined as the distance between two consecutive detected intra-coded pictures.

3.2. *Average quantization scale determination*

The average quantization scale Q is determined to be the quantization step size used on the first AC coefficient, or $\Delta_{0,1}$. In doing this, we fix the quantization matrix weight for the first AC coefficient at 16. Even if a different weight than 16 is used during compression, the macroblock level quantization scale will be adjusted accordingly such that the quantization step size for each coefficient in each macroblock is still determined correctly. In other words,

$$Q = \frac{16\Delta_{0,1}}{W_{0,1}} = \Delta_{0,1}.$$

To find $\Delta_{0,1}$, we use a procedure similar to that used in pitch detection in speech processing. We examine the auto-correlation function of the histogram of the first AC coefficient and find the DRL that yields the largest auto-correlation function value.

3.3. *Quantization matrix extraction*

Once the average quantization scale $\Delta_{0,1}$, or Q is determined, we extract each quantization weight in the quantization matrix. Recall that

$$\Delta_{i,j} = \frac{W_{i,j}Q}{16}.$$

Hence

$$W_{i,j} = \frac{\Delta_{i,j}16}{Q}.$$

Therefore, the quantization matrix extraction problem is reduced to finding the DRL for each coefficient (or DCT frequency band). This is performed as described in average quantization scale determination.

3.4. *Macroblock quantization scale and DCT type extraction*

For MPEG-2 compressed video, both field and frame based DCTs are performed on each macroblock. To isolate the macroblock level quantization scale from the quantization equation, the DCT coefficients are normalized by the quantization matrix weight.

$$\bar{C}_{i,j} = \frac{C_{i,j} \times 16}{W_{i,j}}.$$

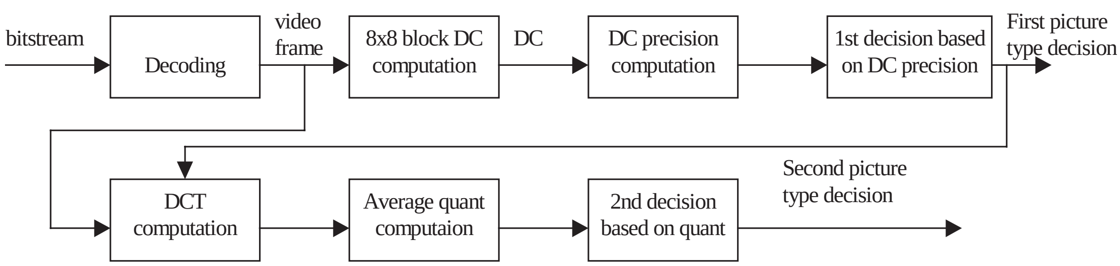

Fig. 3. Intra picture detection.

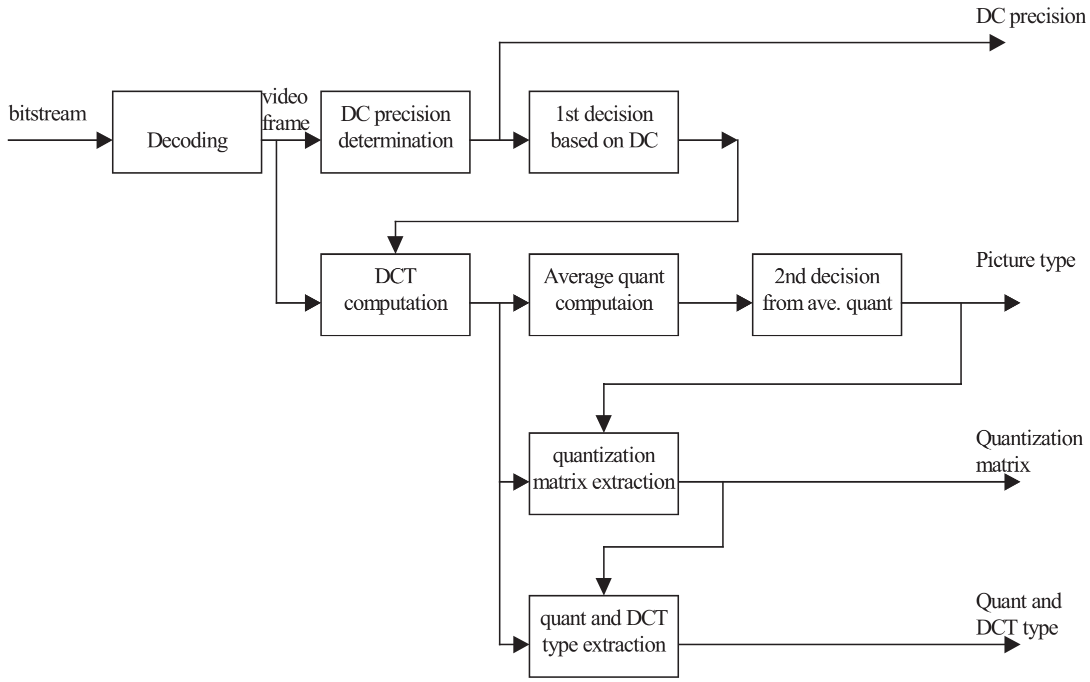

Fig. 4. System diagram for coding parameter extraction.

Here $C_{i,j}$ is the (i,j)th AC coefficient in the current macroblock and $\bar{C}_{i,j}$ is the weighted or normalized AC coefficient.

The next step is to find the DCT type and quantization step size q for each macroblock. This can be done by finding the largest common divider for all normalized AC coefficients in the macroblock, or just for a selected set of coefficients, such as the first 4 AC coefficients in zig-zag scanning order. This is performed on both frame DCT data and field DCT data. The DCT type and quantization scale combination that yields the smaller quantization error is determined to be the compression parameters. The overall coding parameter extraction system is illustrated in Fig. 4.

4. Estimation of statistical parameters

It is well known from literature that the DCT coefficients for I pictures obey a Laplacian probability distribution. The Laplacian probability density, $f(x)$, for each AC coefficient may be written as

$$f(x) = \frac{1}{2\lambda_i} e^{-|x|/\lambda_i}, \tag{1}$$

where λ_i is the rate parameter of the distribution, with different λ_i corresponding to the 63 different AC coefficients in a block.

As outlined in the introduction, in order to compute the quantization distortion, we need to estimate statistical properties of the original data. However, since we only have the quantized data, we need to relate the parameters of the original data distribution with the properties of the quantized data. LoPresto et al. [6] propose a maximum likelihood approach to estimate the variance of a generalized Gaussian distribution from quantized data. Their approach however is complex and needs an iterative solution. Alternately, we provide a simple scheme to estimate these parameters. We use the MPEG-2 TM5 quantization shown in Fig. 5.

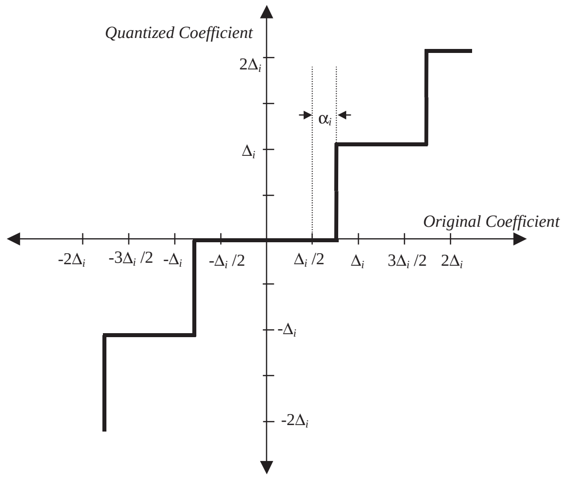

Fig. 5. MPEG-2 quantization scheme.

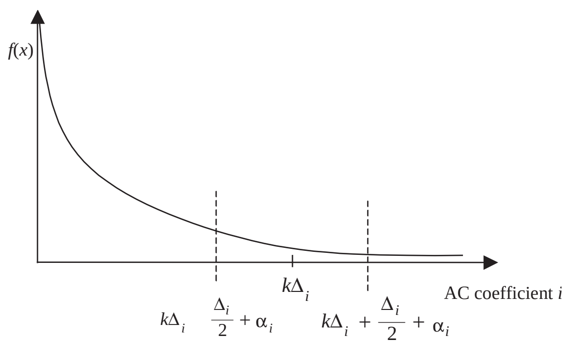

Fig. 6. Estimation of squared error for one sample interval.

In a simple quantization scheme, values in interval $[k\Delta_i - \Delta_i/2, k\Delta_i + \Delta_i/2]$ are truncated to $k\Delta_i$. However, as may be seen from Fig. 6, the MPEG-2 TM5 quantization scheme involves shifting the reconstruction windows by α_i for higher coding efficiency. The parameter α_i depends on Q and W_i, and may be obtained from the MPEG-2 TM-5.

As mentioned before, we assume that each original AC coefficient belongs to a Laplacian distribution with parameter λ_i. Since the quantized coefficients X_i^Q have discrete values, we may compute the discrete probability density function for these quantized coefficients, and this is shown in Eq. (2). For ease of notation, we have dropped the subscript i corresponding to the different AC coefficients. However, we compute these probabilities separately for each AC coefficient using their corresponding Δ_i and λ_i.

$$P(|X^Q| = k\Delta) = \begin{cases} \frac{1}{2\lambda}\Big[\int_{k\Delta-(\Delta/2)+\alpha}^{k\Delta+(\Delta/2)+\alpha} e^{-x/\lambda}\,dx \\ \quad + \int_{-k\Delta-(\Delta/2)-\alpha}^{-k\Delta+(\Delta/2)-\alpha} e^{x/\lambda}\,dx\Big]; & k>0, \\ \frac{1}{2\lambda}\int_{-(\Delta/2)-\alpha}^{(\Delta/2)+\alpha} e^{-|x|/\lambda}\,dx; & k=0. \end{cases} \tag{2}$$

This may be simplified into the following

$$P(|X^Q| = k\Delta) = \begin{cases} e^{-\alpha/\lambda}[e^{-(k\Delta-\Delta/2)/\lambda} - e^{-(k\Delta+\Delta/2)/\lambda}]; & k>0, \\ [1 - e^{-(\Delta/2)-\alpha/\lambda}]; & k=0. \end{cases} \tag{3}$$

Once we have these discrete probabilities, we can compute the second moment S_i^Q of each AC coefficient of the quantized data as follows:

$$S_i^Q = \sum_{k=0}^{\infty} k^2 \Delta_i^2 P(|X_i^Q| = k\Delta_i). \tag{4}$$

An assumption made in the above equation is that the reconstruction levels do not have an upper bound, which is typically not the case. However, this assumption is reasonable as the probability of the coefficients taking large values decays exponentially. This decay outweighs the increasing squared terms leading to a small approximation error. We may substitute the pdf in Eq. (3) into Eq. (4) and simplify the expression to obtain the following relationship. The derivation is shown in Appendix A.

$$S_i^Q = \Delta_i^2 e^{-\alpha_i/\lambda_i} e^{-3\Delta_i/2\lambda_i}\left[\frac{1+e^{-2\Delta_i/\lambda_i}}{(1-e^{-\Delta_i/\lambda_i})^2}\right]. \tag{5}$$

Now that we have derived the theoretical second moment for each AC coefficient quantized using the MPEG-2 scheme, we can compare it with the second moment of the received data to obtain the data statistical properties. The second moment $\hat{S}_i^Q$, of the received quantized data, for each AC coefficient may be computed as

$$\hat{S}_i^Q = \frac{1}{N}\sum_{j=1}^{N} (X_{i,j}^Q)^2, \tag{6}$$

where N is the number of blocks in the frame.

A comparison of Eqs. (5) and (6) leads us to an estimate for the parameter λ_i for each AC coefficient of the original data. This estimate is very accurate when there are a sufficient number of non-zero quantized coefficients. This is typically true for low frequency DCT coefficients. However due to the large W_i and the lack of information in higher frequencies, these estimates become inaccurate for the high frequency DCT coefficients. A comparison of Eqs. (5) and (6) leads to an over-estimation of the parameter. Instead, it is better to compare Eq. (6) to the second moment of the original un-quantized data, $2\lambda_i^2$ to estimate this parameter. Empirically, we have determined that λ_i for the first 24 DCT coefficients (in the zigzag scan order) is estimated accurately using Eqs. (5) and (6), across different sequences. We estimate the parameters for the other high frequency coefficients by comparing Eq. (6) with $2\lambda_i^2$.

5. Quantization error estimation

Once we have the quantization parameters and the parameters of the distribution of the original data, we can estimate the average AC quantization error incurred over the frame. As may be seen in Fig. 6, the coefficients in the range $[k\Delta_i - \Delta_i/2 + \alpha_i, k\Delta_i + \Delta_i/2 + \alpha_i]$ are truncated to $k\Delta_i$.

The average squared quantization error, $\varepsilon_{i,k}^2$, for AC coefficient i in this interval k, may be written as

$$\varepsilon_{i,k}^2 = \frac{1/2\lambda_i \int_{k\Delta_i-(\Delta_i/2)+\alpha_i}^{k\Delta_i+(\Delta_i/2)+\alpha_i} (x-k\Delta_i)^2 \mathrm{e}^{-x/\lambda_i}\,\mathrm{d}x}{P(k\Delta_i-(\Delta_i/2)+\alpha_i < x < k\Delta_i+(\Delta_i/2)+\alpha_i)}. \tag{7}$$

The expression in Eq. (7), corresponds to the case when $k>0$. It is similarly very easy to write expressions for $k \leqslant 0$. Hence, the overall average squared distortion ε_i^2 for each AC coefficient may be obtained as follows:

$$\varepsilon_i^2 = \sum_{k=-\infty}^{\infty} \varepsilon_{i,k}^2 P\left(k\Delta_i - \frac{\Delta_i}{2} + \alpha_i < x < k\Delta_i + \frac{\Delta_i}{2} + \alpha_i\right). \tag{8}$$

By substituting from Eq. (7) to Eq. (8), we may obtain Eq. (9). As before, in order to simplify the notation we drop the subscript i corresponding to the different AC coefficients.

$$\begin{aligned}\varepsilon^2 = &\frac{1}{2\lambda}\sum_{k=1}^{\infty}\int_{-k\Delta-(\Delta/2)-\alpha}^{-k\Delta+(\Delta/2)-\alpha} (x+k\Delta)^2\mathrm{e}^{x/\lambda}\,\mathrm{d}x \\ &+ \frac{1}{2\lambda}\sum_{k=1}^{\infty}\int_{k\Delta-(\Delta/2)+\alpha}^{k\Delta+(\Delta/2)+\alpha} (x-k\Delta)^2\mathrm{e}^{-x/\lambda}\,\mathrm{d}x \\ &+ \frac{1}{2\lambda}\int_{-\Delta/2-\alpha}^{\Delta/2+\alpha} x^2\mathrm{e}^{-|x|/\lambda}\,\mathrm{d}x.\end{aligned} \tag{9}$$

Due to the symmetry of the distribution, the first and the second terms are identical. We solve for this ε^2 and obtain a closed form expression. The details of this derivation are included in Appendix B.

$$\varepsilon^2 = 2\lambda^2 - \frac{2\lambda\Delta\mathrm{e}^{-\alpha/\lambda}\mathrm{e}^{-\Delta/2\lambda}}{(1-\mathrm{e}^{-\Delta/\lambda})}\left[\frac{\alpha}{\lambda}+1\right]. \tag{10}$$

This distortion is computed for each AC coefficient separately.

Besides the AC error, we also need to estimate the DC coefficient error. In MPEG-2 video, intra DC precision (IDP) controls the quantization coarseness of DC coefficients. This value ranges from 8 to 11 bits, with 11 bits corresponding to no quantization error. In order to estimate the IDP we use the method described in [2]. Once the IDP is determined, the average DC quantization error is determined as follows:

$$\varepsilon_{\mathrm{DC}} = \begin{cases} 2^{10-\mathrm{IDP}}; & \mathrm{IDP} < 11, \\ 0; & \mathrm{IDP} = 11. \end{cases} \tag{11}$$

We can combine these estimates to obtain the average squared distortion D over the frame and hence, the PSNR:

$$D = \frac{1}{64}\left[\varepsilon_{DC}^2 + \sum_{i=1}^{63} \varepsilon_i^2\right]$$
$$\text{and PSNR} = 10\log_{10}\left(\frac{255^2}{D}\right). \quad (12)$$

We may summarize the entire scheme as follows:

1. Estimate Δ_i and λ_i for each AC coefficient.
2. Use Eqs. (10) and (11) to compute the squared AC and DC quantization error.
3. Combine these estimated errors to obtain D and the PSNR, as in Eq. (12).

6. Experimental results

We use three test sequences, Basket, Thelma and Doll. These sequences are 720×576 at 30 Hz, and range from high to moderate spatial details. Sample frames from two of these sequences, Basket and Doll are shown in Fig. 7.

We intra code frames from these sequences using the MPEG quantization weights and a fixed Q. We use an intermediate range $[8, \ldots, 40]$ for Q as this includes values typically used during coding. Results for these sequences across different Q are shown in Table 1.

Each PSNR entry is averaged across three Intra coded frames from the corresponding sequence. In all cases the estimated PSNR lies within 3% or 1 dB of the actual PSNR. It is clear that the performance of the scheme is consistent across these different sequences. We observe that the PSNR is consistently under-estimated for high Q step size values. This is because at large Q values the number of non-zero coefficients is very small, leading to inaccurate estimates of the statistical parameters, and consequently, an under-estimation of the quantization error.


Fig. 7. Snapshots from Basket (left) and Doll.

Table 1
PSNR estimation results

Sequence	PSNR (dB)	$Q = 8$	$Q = 16$	$Q = 24$	$Q = 32$	$Q = 40$
Basket	Actual	36.52	32.99	31.13	29.80	28.78
	Est.	37.51	33.30	31.24	29.58	28.37
Thelma	Actual	37.57	35.46	34.63	33.94	33.38
	Est.	38.42	35.77	34.56	33.56	32.39
Doll	Actual	39.06	35.36	33.52	32.30	31.46
	Est.	39.51	35.56	33.75	32.12	30.61

7. Conclusion

In this paper we have described a scheme to use statistical properties of the picture data to estimate the quantization error. After estimating the quantization error, we can compute simple metrics of quality like the PSNR or may compute more perceptually relevant metrics using models of human perception. We have implemented the scheme to estimate the PSNR of I pictures in coded video sequences and find that the estimates are within 3% or 1 dB across different sequences and quantization step sizes. Since Intra coding of video frames is very similar to image coding schemes like JPEG, the same scheme may be used to evaluate compressed image quality.

Future work includes examining the performance under varying quantization step sizes as well as varying quantization schemes. Importantly, although, we have demonstrated the process using DCT based coding schemes, the ideas of extracting statistical properties from coded data and estimating quantization error may be extended to non-DCT based coding schemes. These ideas may also be easily extended to region-based quality measures. Such region-based error estimates may be combined using perceptual information for better measures of quality. For instance, distortions are less visible in heavily textured areas. Besides this region-based combination, we may also use perceptual weighting of the different frequency bands or pooling to obtain a more relevant quality metric. This region-based scheme may be extended to object-based quality estimation schemes also.

Appendix A

We may now find the first and second moments (respectively m_Q and S_Q) of the data. For ease of notation, we drop the index i.

$$m_Q = \sum_{k=0}^{\infty} k\Delta P(|X_Q| = k\Delta) = \Delta e^{-\alpha/\lambda} \times \left[\sum_{k=1}^{\infty} k e^{-(k\Delta-\Delta/2)/\lambda} - \sum_{k=1}^{\infty} k e^{-(k\Delta+\Delta/2)/\lambda}\right],$$

$$m_Q = \Delta e^{-\alpha/\lambda}(e^{\Delta/2\lambda} - e^{-\Delta/2\lambda})\left[\sum_{k=1}^{\infty} k e^{-(k\Delta)/\lambda}\right].$$

We know that

$$\sum_{k=0}^{\infty} e^{ky} = \frac{1}{1-e^{y}} \quad \text{given that } |e^{y}| < 1.$$

Taking derivatives w.r.t. y on both sides, we obtain that $\sum_{k=1}^{\infty} k e^{(k-1)y} = e^{y}/(1-e^{y})^2 \Rightarrow \sum_{k=1}^{\infty} k e^{ky} = e^{2y}/(1-e^{y})^2$.

Hence

$$m_Q = \Delta e^{-\alpha/\lambda}(e^{\Delta/2\lambda} - e^{-\Delta/2\lambda})\left[\frac{e^{-2\Delta/\lambda}}{(1-e^{-\Delta/\lambda})^2}\right] = \frac{\Delta e^{-\alpha/\lambda} e^{-3\Delta/2\lambda}}{(1-e^{-\Delta/\lambda})}.$$

Continuing further, we can find the second moment S_Q of the data.

$$S_Q = \sum_{k=0}^{\infty} k^2 \Delta P(|X_Q| = k\Delta) = \Delta^2 e^{-\alpha/\lambda} \times \left[\sum_{k=1}^{\infty} k^2 e^{-(k\Delta-\Delta/2)/\lambda} - \sum_{k=1}^{\infty} k^2 e^{-(k\Delta+\Delta/2)/\lambda}\right]$$

$$S_Q = \Delta^2 e^{-\alpha/\lambda}(e^{\Delta/2\lambda} - e^{-\Delta/2\lambda})\left[\sum_{k=1}^{\infty} k^2 e^{-(k\Delta)/\lambda}\right]. \tag{A.1}$$

We know

$$\sum_{k=1}^{\infty} k e^{(k-1)y} = \frac{e^{y}}{(1-e^{y})^2}.$$

Taking derivative w.r.t. y we get

$$\sum_{k=2}^{\infty} k(k-1) e^{(k-2)y} = \frac{e^{y}}{(1-e^{y})^2} + \frac{2e^{2y}}{(1-e^{y})^3} = \frac{e^{y}(1+e^{y})}{(1-e^{y})^3} \Rightarrow \sum_{k=2}^{\infty} k(k-1) e^{ky} = \frac{e^{3y}(1+e^{y})}{(1-e^{y})^3}.$$

Since

$$\sum_{k=1}^{\infty} k^2 e^{ky} = \sum_{k=2}^{\infty} k(k-1) e^{ky} + \sum_{k=1}^{\infty} k e^{ky},$$

we may substitute this into Eq. (A.1) to obtain the following:

$$S_Q = \Delta^2 e^{-\alpha/\lambda}(e^{\Delta/2\lambda} - e^{-\Delta/2\lambda})\left[\sum_{k=1}^{\infty} k^2 e^{-(k\Delta)/\lambda}\right]$$
$$= \Delta^2 e^{-\alpha/\lambda}(e^{\Delta/2\lambda} - e^{-\Delta/2\lambda}) \times \left[\sum_{k=2}^{\infty} k(k-1)e^{-(k\Delta)/\lambda} + \sum_{k=1}^{\infty} k e^{-(k\Delta)/\lambda}\right]$$
$$\Rightarrow S_Q = \Delta^2 e^{-\alpha/\lambda}(e^{\Delta/2\lambda} - e^{-\Delta/2\lambda}) \times \left[\frac{e^{-3\Delta/\lambda}(1+e^{-\Delta/\lambda})}{(1-e^{-\Delta/\lambda})^3} + \frac{e^{-2\Delta/\lambda}}{(1-e^{-\Delta/\lambda})^2}\right]$$
$$\Rightarrow S_Q = \Delta^2 e^{-\alpha/\lambda} e^{\Delta/2\lambda}(1-e^{-\Delta/\lambda})\left[\frac{e^{-4\Delta/\lambda} + e^{-2\Delta/\lambda}}{(1-e^{-\Delta/\lambda})^3}\right]$$
$$= \Delta^2 e^{\Delta/2\lambda} e^{-2\Delta/\lambda}\left[\frac{1+e^{-2\Delta/\lambda}}{(1-e^{-\Delta/\lambda})^2}\right]$$
$$\Rightarrow S_Q = \Delta^2 e^{-\alpha/\lambda} e^{-3\Delta/2\lambda}\left[\frac{1+e^{-2\Delta/\lambda}}{(1-e^{-\Delta/\lambda})^2}\right].$$

Appendix B

In this appendix we derive an expression for the squared quantization error. This error may be computed as in Eq. (9). As before, we drop the subscript i for ease of notation.

$$\varepsilon^2 = \frac{1}{2\lambda}\sum_{k=1}^{\infty}\int_{-k\Delta-(\Delta/2)-\alpha}^{-k\Delta+(\Delta/2)-\alpha}(x+k\Delta)^2 e^{x/\lambda}\,dx + \frac{1}{2\lambda}\sum_{k=1}^{\infty}\int_{k\Delta-(\Delta/2)+\alpha}^{k\Delta+(\Delta/2)+\alpha}(x-k\Delta)^2 e^{-x/\lambda}\,dx + \frac{1}{2\lambda}\int_{-\Delta/2-\alpha}^{\Delta/2+\alpha} x^2 e^{-|x|/\lambda}\,dx.$$

Due to symmetry the first and second terms may be combined into one term. Hence

$$\varepsilon^2 = \underbrace{\frac{1}{\lambda}\sum_{k=1}^{\infty}\int_{k\Delta-(\Delta/2)+\alpha}^{k\Delta+(\Delta/2)+\alpha}(x-k\Delta)^2 e^{-x/\lambda}\,dx}_{\varepsilon_1^2} + \underbrace{\frac{1}{2\lambda}\int_{-\Delta/2-\alpha}^{\Delta/2+\alpha} x^2 e^{-|x|/\lambda}\,dx}_{\varepsilon_2^2}.$$

Replacing $y = x - k\Delta - \alpha$, we may rewrite ε_1^2 as

$$\varepsilon_1^2 = \frac{1}{\lambda}\sum_{k=1}^{\infty} e^{-(k\Delta+\alpha)/\lambda}\int_{-\Delta/2}^{\Delta/2}(y+\alpha)^2 e^{-y/\lambda}\,dy$$
$$= \frac{e^{-(\Delta+\alpha)/\lambda}}{(1-e^{-\Delta/\lambda})}\frac{1}{\lambda}\int_{-\Delta/2}^{\Delta/2}(y+\alpha)^2 e^{-y/\lambda}\,dy$$
$$\Rightarrow \varepsilon_1^2 = \frac{\lambda^2 e^{-(\Delta+\alpha)/\lambda}}{(1-e^{-\Delta/\lambda})}\left\{e^{\Delta/2\lambda}\left[\left(\frac{\Delta-2\alpha}{2\lambda}\right)^2 - \left(\frac{\Delta-2\alpha}{\lambda}\right) + 2\right] - e^{-\Delta/2\lambda}\left[\left(\frac{-\Delta-2\alpha}{2\lambda}\right)^2 - \left(\frac{-\Delta-2\alpha}{\lambda}\right) + 2\right]\right\}. \quad \text{(B.1)}$$

Similarly we may solve for the second term to obtain

$$\varepsilon_2^2 = \lambda^2\left\{2 - e^{-(\Delta+2\alpha)/2\lambda}\left[\left(\frac{\Delta+2\alpha}{2\lambda}\right)^2 + \left(\frac{\Delta+2\alpha}{\lambda}\right) + 2\right]\right\}. \quad \text{(B.2)}$$

We may add Eqs. (B.1) and (B.2) and simplify to obtain

$$\varepsilon^2 = 2\lambda^2 - \frac{2\lambda\Delta e^{-\alpha/\lambda}e^{-\Delta/2\lambda}}{(1-e^{-\Delta/\lambda})}\left[\frac{\alpha}{\lambda} + 1\right].$$

References

[1] J. Caviedes, J. Jung, No-reference metric for a video quality control loop, SCI 2001, Orlando, July 2001.
[2] Y. Chen, K. Challapali, M. Balakrishnan, Extracting quantization parameters from previously compressed MPEG-2 video, Proceedings of the International Conference on Image Processing, Santa Barbara, October 1998.
[3] S. Karunasekera, N. Kingsbury, A distortion measure for blocking artifacts in images based on human visual sensitivity, IEEE Trans. Image Proc. 4 (6) (1995) 713–724.
[4] M. Knee, The picture appraisal rating (PAR)—a single-ended picture quality measure for MPEG-2, Snell& Wilcox website http://www.snellwilcox.com.
[5] C. Lambrecht, O. Verscheure, Perceptual quality measure using a spatio-temporal model of the human visual system, Proceedings of SPIE, Vol. 2688, January 1996, pp. 450–61.
[6] S. LoPresto, K. Ramchandran, M. Orchard, Image coding based on mixture modeling of wavelet coefficients and a

fast estimation-quantization framework, DCC '97, Snowbird, March 1997.

[7] J. Lubin, D. Fibush, Objective video quality measurement using a JND-based full reference technique, ATIS Technical Report T1.TR.75-2001, October 2001.

[8] Methodology for subjective assessment of the quality of television pictures, Recommendation ITU-R BT.500-10, 2000.

[9] M. Miyahara, K. Kotani, V. Algazi, Objective picture quality scale for image coding, IEEE Trans. Commun. 46 (9) (September 1998) 1215–1226.

[10] S. Smoot, L. Rowe, Study of DCT coefficient distributions, Proceedings of SPIE, January 1996, pp. 403–411.

[11] S. Wolf, M. Pinson, Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system, SPIE International Symposium on Voice, Video, and Data Communications, Boston, MA, September 11–22, 1999.

[12] M. Yuen, H. Wu, A survey of hybrid MC/DPCM/DCT video coding distortions, Signal Processing 70 (1998) 247–278.