A novel method for updating vector quantised codebooks

L. D. Scargall and S. S. Dlay

Department of Electrical & Electronic Engineering
University of Newcastle
Newcastle upon Tyne
UK

Abstract: We present a novel methodology for adaptive vector quantisation (AVQ), where the codebook is updated with new code-vectors. The new code-vectors replace less significant ones based on a scoring criterion that utilises a forgetting factor and a codebook half-life. The proposed method is shown to outperform conventional techniques of AVQ by approximately 1dB.

Key words: vector quantisation, very low bit-rate, image coding

1. **Introduction:** It is to some extent believed that future video codecs for very low bit-rates i.e. below 20 kbit/s, cannot be immune from vector quantisation (VQ) due to its overwhelming ability to reduce the spatial redundancy [1]. In this paper, we investigate the performance of an adaptive VQ image coding scheme for very low bit-rates. The codebook of the vector quantiser is updated with new code-vectors to take into account the time-variant nature of video sequences. In our adaptive approach new code-vectors replace less significant ones in the codebook based on a novel scoring criterion that utilises a forgetting factor derived from a codebook half-life, to calculate the local importance of each code-vector after every frame.

VQ [1-3] can be used as a powerful means of data compression by comparing an input vector to a finite set of reproduction code-vectors. The selected VQ blocks *x*, are then compared to a finite set of reproduction code-vectors *y*, in the codebook. The comparison measure is based upon minimizing the distortion penalty incurred by representing an input vector with one of the code-vectors. In theory, VQ can achieve a performance close to the rate distortion bound as the dimension of the code-vectors approach infinity.

2. **Proposed Methodology:** In AVQ [4, 5] the codebook is typically updated with new code-vectors that occur frequently but are not represented in the codebook. The increased flexibility of this method is unfortunately associated with an inflated bandwidth requirement, to transmit the new code-vectors to the receiver and an increase in computational complexity. In our adaptive approach, new code-vectors replace less significant ones in the codebook based upon a novel scoring system. Every time a code-vector is selected from the codebook, it inherits a unit score. Hence, a code-vector that is used more frequently is deemed more significant, will have a higher score. However, it is fair to assume that the non-stationary nature of video sequences may cause a code-vector to gain an artificially high score, if it is used excessively at one period and remains redundant for the rest of the video sequence. For this purpose it is vital to take into account the local scores of each code-vector. This is

accomplished by multiplying the scores of the codebook by a forgetting factor after each frame. The forgetting factor a, is a number less than one, and is determined by the pre-selected codebook half-life z. This is mathematically represented by Equation 1.

$$0.5 = a^z \tag{1}$$

The half-life determines the number of frames that must pass before the codebook scores will be reduced by exactly a half of their current value. This is exemplified in Figure 1 for a half-life of $z = 1$ frames. This implies that after every z frames the codebook scores are reduced by exactly a half. When a new code-vector is initiated into the codebook it must be given a chance to establish itself. For this reason, the unit score is inappropriate, so the new code-vector is assigned the value M. This is represented by Equation 2, where $P(X)$ is the probability that a vector will be selected from the codebook, multiplied by the discrete summation of the forgetting factor.

$$M = \sum_{i=0}^{\infty} a^i \cdot P(X) \tag{2}$$

As the modulus of the forgetting factor is less than 1, the discrete summation can be represented by a geometric progression. The Probability $P(X)$, that a code-vector will be randomly selected from the codebook is defined by Equation 3.

$$P(X) = \frac{Number\ of\ VQ\ blocks}{Codebook\ size} \tag{3}$$

In our adaptive approach, the codebook is configured to update with the code-vector that provides the highest encoding gain within the spatial frame. This is mathematically represented by Equation 4.

$$AVQ_{gain} = \sum_{i=1}^{k} \sum_{j=1}^{k} (y_n(i,j))^2 - \sum_{i=1}^{k} \sum_{j=1}^{k} (x(i,j) - y_n(i,j))^2 \tag{4}$$

In a typical head and shoulders video sequence the changes within the frame are relatively low. Hence, a code-vector within the codebook may have to remain significant for only several frames after it has been selected i.e. a code-vector may typically have to remain significant for up to 50 frames (or 5 seconds) for a slow video sequence. Conversely, a video sequence that contains large amounts of scene changes will require a code-vector to be significant for a much shorter period after it has been selected in order to refresh the codebook, i.e. a code-vector may now have to remain significant for only 5 frames (or ½a second). Based on this principle, the head and shoulders test sequences of "Miss America", "Carphone", "Silent" and "Claire" were simulated for half-lives of 5, 10 and 50 frames. The results are illustrated in Figures 2, 3, 4 and 5, and are also tabulated in Table 1. They represent a codebook of 64 entries with 2 new code-vectors per frame. It is clearly visible from these Figures that the proposed method of AVQ, which utilizes the forgetting factor, is superior to conventional based methods that rely on frequency scoring alone.

3. **Comparison of the ITU-T H.263 codec:** As a comparative basis, we configured the ITU-T standard H.263 codec, (version 2.0 of the Telenor simulator, using advanced prediction and PB-frames) to operate at a target bit-rate of 20 kbit/s and a frame rate of 10 fps. In our PSNR comparison, we compare the H.263 codec to our adaptive AVQ(1024-4) codec with 4 updates per frame and to our non-adaptive VQ(1024) codec that has a bit-rate of 11.4 kbit/s. The average PSNR values are tabulated in Table 2 for the four test video sequences. With reference to these results, the proposed adaptive scheme achieves a comparable PSNR performance to that of the H.263 codec.

A further comparison of the subjective visual quality by way of reconstructed frames from the output test sequences was also performed. Unbiased test persons confirmed that the proposed AVQ encoded images are sharper, contain more detail and are without blocking effects.

4. **Conclusion:** A new methodology has been presented for AVQ, which utilises a novel-scoring criterion based on a forgetting factor and codebook half-life. The new method gives rise to an additional performance enhancement of around 1dB over conventional techniques of AVQ that are based on frequency scoring. It has also been shown that the performance of the proposed adaptive scheme achieves a similar PSNR performance to the H.263 standard codec at a similar bit-rate.

5. References:

[1] A. Gersho, R.M. Gray: "*Vector Quantization and Signal Compression*", Kluwer Academic Publishers, 1992.

[2] L. Scargall and S. S. Dlay, "A Comparison of Vector Quantised Video Codecs", Proc. of SPIE-VVD'98, Boston, MA, America, pp. 502-508, Nov. 1-6, 1998.

[3] J. Streit, L. Hanzo: "Dual-Mode Vector Quantised Low Rate Cordless Videophone Systems for Indoors and Outdoors Applications", IEEE Trans. on V.T., Vol. 46, No. 2, pp 340-357, May 1997.

[4] J. E. Fowler: "Generalised Threshold Replenishment: An Adaptive Vector Quantizer Algorithm for the Coding of Nonstationary Sources", IEEE Trans. on Image Processing, Vol. 7, No. 10, pp. 1410-1424, Oct. 1998.

[5] R. Hamzaoui, D. Saupe, M. Wagner, "Rate-Distortion Based Video Coding with Adaptive Mean-Removed Vector Quantization", in Proc. of IEEE International conf. on image Processing, Chicago, IL, America, Oct. 1998.

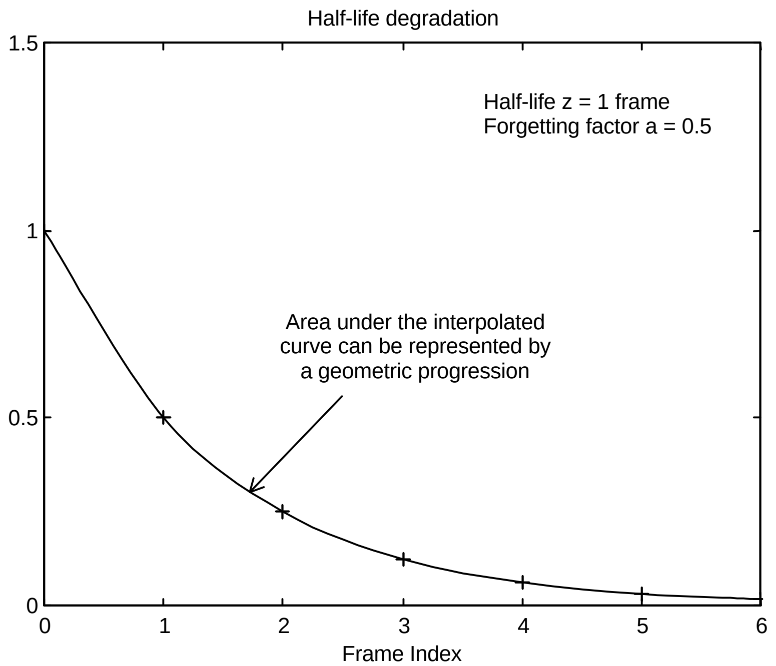

Figure 1: Half-life degradation

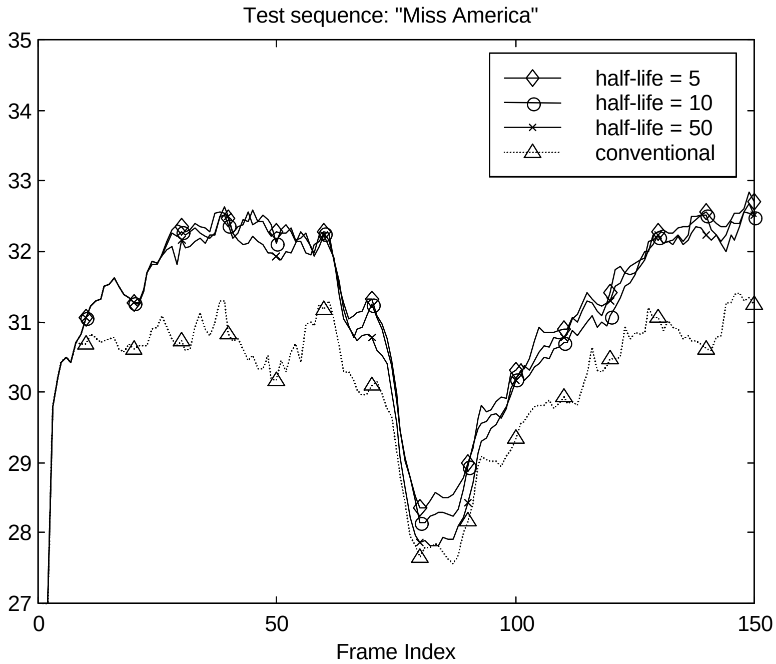

Figure 2: PSNR results of the AVQ(64-2) codec.

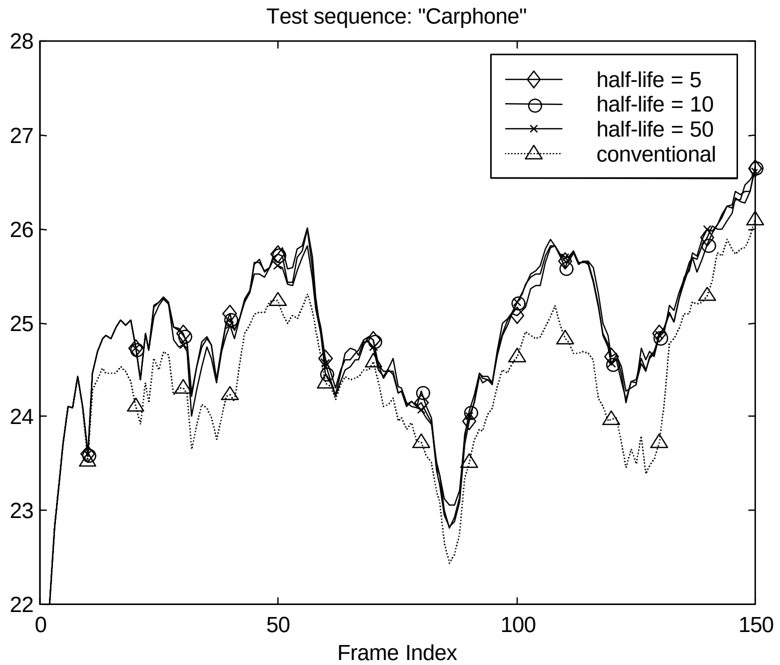

Figure 3: PSNR results of the AVQ(64-2) codec.

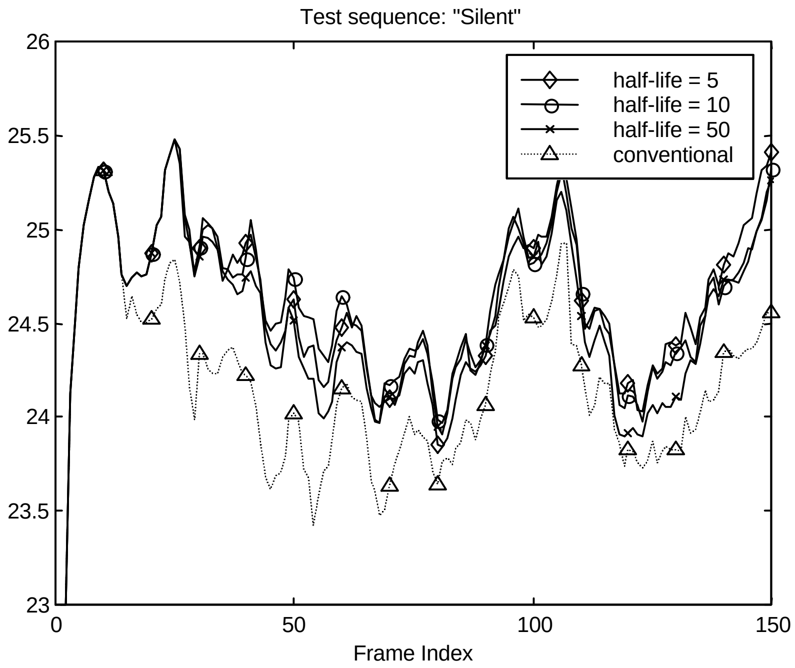

Figure 4: PSNR results of the AVQ(64-2) codec.

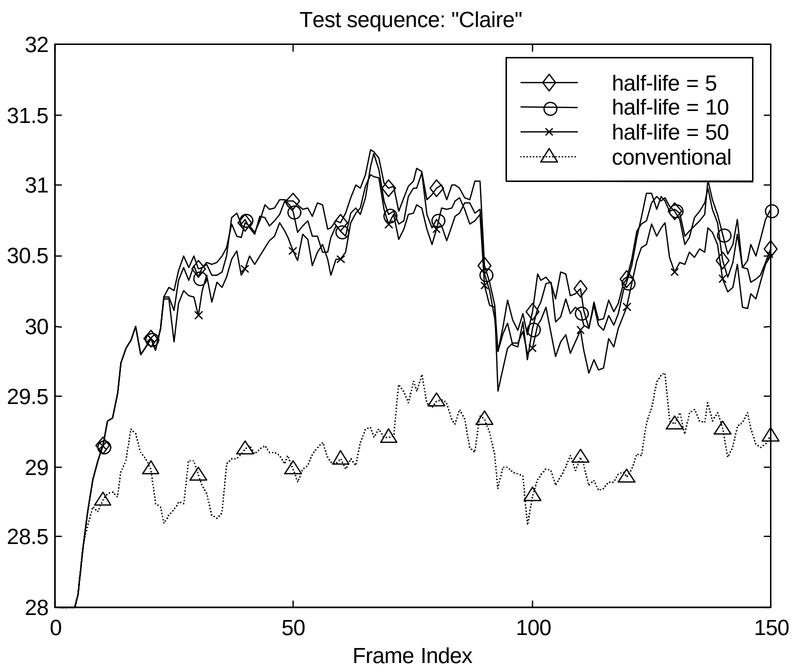

Figure 5: PSNR results of the AVQ(64-2) codec.

Half-life in frames AVQ(64-2)	Test video sequence			
	“Miss America”	“Carphone”	“Silent”	“Claire”
Conventional method	30.15dB	24.35dB	24.16dB	28.95dB
50	31.03dB	24.83dB	24.51dB	30.10dB
10	31.11dB	24.86dB	24.60dB	30.24dB
5	31.22dB	24.95dB	24.59dB	30.30dB

Table 1: Average PSNR performance of the conventional and proposed half-life codecs.

Codec Type	Test video sequence			
	“Miss America”	“Carphone”	“Silent”	“Claire”
H.263 at 20kbit/s	36.54dB	28.40dB	28.53dB	35.39dB
VQ(1024) at 11.4kbit/s	33.68dB	26.71dB	27.16dB	33.35dB
AVQ(1024-4) at 21.6kbit/s	34.93dB	28.00dB	27.91dB	35.12dB

Table 2: Average PSNR performance comparison of H.263 to the proposed codecs.