

PRACTICAL LOW BIT RATE PREDICTIVE IMAGE CODER USING MULTI-RATE PROCESSING AND ADAPTIVE ENTROPY CODING

Anna N. Kim and Tor A. Ramstad

Department of Telecommunications, NTNU
O.S. Bragstads plass 2B, N-7491 Trondheim, Norway

ABSTRACT

The modified Differential Pulse Coded Modulation (DPCM) codec with multi-rate processing has been shown to be able to code source with monotonically decreasing spectrum efficiently at low bit rates [1]. A practical image coder is designed based on this approach. Two dimensional DPCM is used along with decimation and interpolation to reduce the number of transmitted samples. The decimation rate depends on the signal spectrum and the bit rate. Further bit rate reduction is achieved through adaptive entropy coding. Wiener filter is appended in the decoder for minimizing distortion caused by quantization noise. The decimation filter can be implemented using simple IIR filters. The necessary side information is low. Simulation results show that the coder is able to give good compression performance at low bit rates which is superior to conventional DPCM codec and JPEG. Subjective quality can be as good as JPEG2000. While at very low bit rates the proposed codec is able to retain certain image characteristics better than JPEG2000.

1. INTRODUCTION

The main advantage of DPCM is its ability to perform efficient coding with low complexity. DPCM coders generally perform well at high rates (3-4 bits/pixel), when quantization steps are small and the optimal predictor is able to minimize the prediction error variance. However at low bit rate regions, the additive noise model is no longer applicable. An exact analysis of the system becomes very involved or even impossible. It is then interesting to modify the DPCM structure to comply with rate-distortion theory while at the same time using a rather applicable system model. The assumptions would have to be examined from experimental results.

In [1], we proposed a modified DPCM codec with multi-rate processing. It is designed for encoding of sources with monotonically decreasing spectra at low bit rates. Similar to the coder in [2], our codec was motivated by rate-distortion theory. It incorporates multi-rate processing and a Wiener receiver filter. Simulation with Gaussian AR(1) processes show that the decimation and interpolation process are able to provide further rate reduction so the overall rate-distortion performance is much closer to the rate-distortion bound at low bit-rate regions compared to [2].

In this paper we want to verify the potential of the proposed method in practice by designing an image coder based on this approach. The goal is to have a practical image coder with good low-bit rate performance, yet low-complexity implementation.

2. DPCM WITH MULTI-RATE PROCESSING

The rate-distortion theory states that the optimal lossy coding of a discrete time, continuous Gaussian amplitude source is done by coding only portions of the source spectrum that is above the noise floor. This is commonly described as the reverse "Water Filling" principle. Consider for example a 1st order Gauss-Markov process (AR(1)) which has monotonically decreasing spectrum. At low target bit rates (high distortion), the high frequency portion of the source spectrum that is below the additive noise should not contribute to the rate. This is achieved by applying a low-pass filter to shape the source spectrum prior to encoding with the standard DPCM.

However, the low-pass filtered source leads to a low-pass input to the quantizer, which is correlated and oversampled. To overcome this problem, we apply down-sampling to the source to lower the sampling rate before coding and corresponding up-sampling and interpolation is performed in the decoder. This modification leads to further overall rate reduction due to the fact that the number of samples for coding after down-sampling is reduced. The rate at the coder output is measured in bits per *coded* sample. However, for the overall rate-distortion performance, the rate is measured in bits per *source* sample. The actual bit rate is then the obtained rate divided by the down-sampling rate. A simple uniform quantizer is used for quantization of the prediction error followed by an entropy coder. We also append a Wiener filter to the decoder. Its role is to further suppress distortion due to quantization and it does not affect the rate.

Simulation results showed that the proposed codec was able to perform very close to the rate-distortion bound for highly correlated AR(1) process (correlation coefficient $\alpha = 0.9$) at low bit rates. When the codec is not operating at the optimal point, that is, when there is a mismatch between the low-pass filter cut-off frequency and the quantization noise level, the Wiener filter was able to provide significant compensation. More detailed description and discussion of the system can be found in [1].

3. LOW BIT-RATE PREDICTIVE IMAGE CODER WITH ADAPTIVE ENTROPY CODING

We now introduce the practical image coder based on the codec described in the previous section.

The image coder has essentially the same structure (Figure 1). The source image is first low-pass filtered and down-sampled two dimensionally. It is then fed through the DPCM encoder, where the difference between the down-sampled image and its prediction is quantized using a uniform quantizer and the quantization indices are entropy coded. On the decoder side, the received quantization

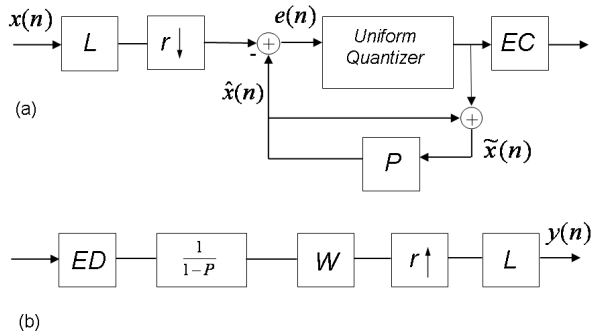


Fig. 1. Block Diagram of Image Coder: (a) Encoder (b) Decoder. *EC*: Entropy Encoder, *ED*: Entropy Decoder, *L* Low-pass Filter, *P*: Prediction Filter, *W*: Wiener Filter, *r*: Decimation/Interpolation Rate

indices first go through entropy decoder and DPCM decoder, followed by the Wiener filter. The resulting image is then up-sampled and interpolated using the same low-pass filter.

A two dimensional DPCM (2D-DPCM) is used in the proposed image coder. We model the down-sampled images as a separable 1st order Gauss-Markov model which has auto-correlation function:

$$R(i, j) = \sigma^2 \rho_1^{|i|} \rho_2^{|j|}, \quad |\rho_1| < 1, |\rho_2| < 1 \quad (1)$$

where ρ_1 and ρ_2 are the one step correlation coefficients in the row and column direction after decimation and σ^2 is the variance. The current pixel value can be calculated by its three surrounding pixels [3]:

$$\hat{x}(i, j) = \rho_1 \hat{x}(i-1, j) + \rho_2 \hat{x}(i, j-1) - \rho_1 \rho_2 \hat{x}(i-1, j-1) \quad (2)$$

The Wiener filter gives an optimal tradeoff in the mean squared sense between suppression of additive quantization noise and linear distortion due to filtering. Its frequency response is given by:

$$W(\omega_i, \omega_j) = \frac{S_{xx}(\omega_i, \omega_j)}{S_{xx}(\omega_i, \omega_j) + \sigma_q^2}, \quad (3)$$

where $S_{xx}(\omega_i, \omega_j)$ is the power spectral density of the down-sampled image and σ_q^2 is the quantization noise variance. The 2-D Wiener filter is generally not separable even if the point spread function and the variance covariance functions are [3]. The required auto-correlation function can be derived from the prediction coefficients, whereas the quantization noise variance that can be approximated by $\sigma_q^2 = \Delta^2/12$, where Δ is the quantization step.

We would like to stress that although the over-all image coder is designed for low-bit rate operation, due to down-sampling, the DPCM part of the system is in fact operating at fairly high rates (above 2 bits/pixel)¹. The additive noise model is then applicable and the assumption that the noise and the signal are uncorrelated is quite accurate when a uniform quantizer is applied. It is safe to consider that the quantization noise is uncorrelated with the quantizer input. The approximation of σ_q^2 is then also valid.

Further rate reduction of the system is obtained through adaptive entropy coding. For our codec in Section 2, the prediction

¹[4] defines high rate above 2 bits/pixel

error is quantized by the uniform quantizer then coded by one entropy coder. However, images have varying local statistics. When a fixed predictor is used in the DPCM encoder, the prediction error image will also have this property. It can be split into small blocks with size that makes their internal statistics constant. One should note that the block size must not be too small, in which case a large amount of side information is required. These blocks can then be classified into a small number of classes depending on their variances. After uniform quantization with the same quantizer for the entire image, an entropy coder can be designed for each class.

A method for optimal selection of class boundaries and entropy coders for each class for a given number of classes was designed in [5],[6]. It was shown that there is an optimal number of classes for which the gain in increasing the number of classes is counteracted by the increase of side information needed for informing the receiver about the block classification. The number of classes, of course, depends on the block size as well. [5] also showed it is adequate to keep the source splitting class number M between 4 and 6. For our proposed coder, $M = 4$ is chosen along with block size of 4x4 pixels. We then use the iteration algorithm in [5] to find the optimal variance decision levels to determine where to apply the four different entropy coders.

4. SIMULATION RESULTS AND DISCUSSION

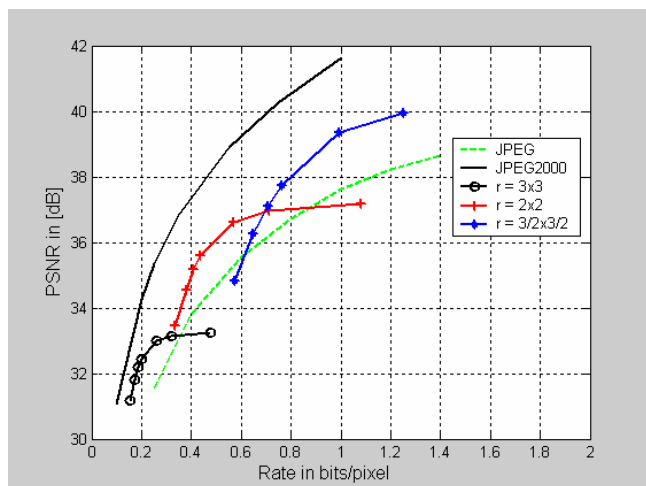


Fig. 2. Coding performance at decimation rate of 3/2x3/2, 2x2 and 3x3 comparing with JPEG. Image “Lena”.

In the first experiment we used the 512x512, monochrome “Lena” image to assess the source coding performance of our coder. Figure 2 shows the results for down-sampling rates $\frac{3}{2} \times \frac{3}{2}$, 2x2 and 3x3 compared with JPEG and JPEG2000². The conventional DPCM performance at such low bit rate region is unsurprisingly substantially worse, hence not included in the plot. The collection of curves of the proposed coder have the same shape as the ones show in [1]. The convex hull of the curves are the best operation point of the codec. For each down-sampling rate, there is an optimal quantization step Δ . When the quantization noise level is too low, the low-pass filtering before down-sampling removes too much of the source spectrum so the rate-distortion performance

²JPEG2000 standard used here was implemented in Java through The JJ2000 project. Source code v5.1 <http://jj2000.epfl.ch>

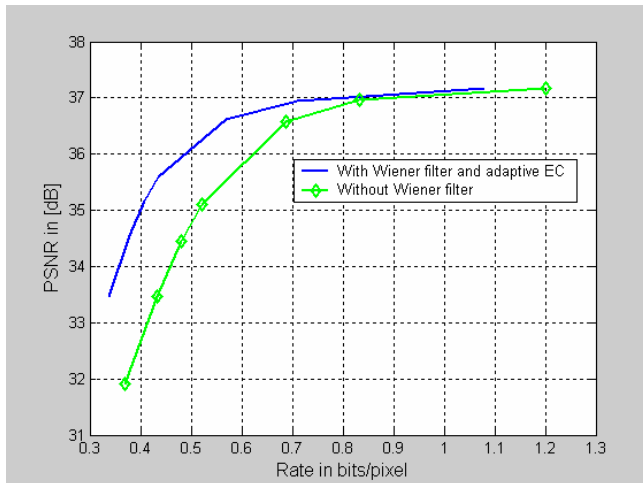


Fig. 3. Coding performance at decimation rate of 2x2. Image “Lena”.

is further away from the “elbow” of the curve. As the quantization noise increases, the performance improves. The “elbow” point is reached when the quantization noise level intercepts the source spectrum at the same frequency of the stop-band of low-pass filter. The performance is again worsened when the quantization noise becomes too high for the source shaping low-pass filter. In practice, for a chosen down-sampling rate, the quantization step Δ for the ideal operating point can be determined through simple iteration since the quantization noise level is approximated by $\Delta^2/12$.

Figure 3 shows the effect of adaptive entropy coding and Wiener filter. As our results in [1] indicated, improvement by Wiener filter is more pronounced beyond the convex hulls when the quantization noise level is higher. The adaptive entropy coding shows reduction rate around 0.15 bits/pixel. Figure 4 compares the original and coded version of “Lena”. The down-sampling rate is limited to $\frac{3}{2} \times \frac{3}{2}$, 2x2 and 3x3 since higher rates result more blurred images due to much removal of the high frequency components. It can be observed in Figure 2 however, that the three decimation factors are adequate to cover a large range of bit rates at almost optimal quality. The common blocking artifact seen in Transform Coding is not present. The ringing noise is very limited when decimation rate is low.

Necessary side information includes prediction coefficient (2 bytes), quantization step size (2 bits), decimation rate (2 bits) and entropy coder location table. The number of bits required for the first three quantities is negligible when the image size is large. Information needed for adaptive entropy coding can be coded as such: in the case of 4 entropy coders and 4x4 block size, the simplest way is to use 2 extra bits per block. which gives 0.125 bpp extra. With decimation rate r the number of bits spent per source pixel is then $0.125/r^2$. The alternative, which is more efficient, is to entropy code the entropy coder location table since the distribution of the coders is not uniform. The probability distribution needed for entropy coding can be modelled as generalized Gaussian distribution and requires 2 additional bytes for the necessary parameters. The resulting over-all side information is low.

FIR filters are generally used in image processing since they can be designed with linear phase. For lower complexity, IIR filters are good alternatives. Phase distortion can be avoided by filtering the image in the reverse direction during interpolation. As an ex-

ample, the low-pass filter for decimation and interpolation rate of 2x2 can be implemented in simple form using an allpass filter and a delay [7]:

$$H(z) = \frac{1}{2} \left[z^{-1} + \frac{\alpha + z^{-2}}{1 + \alpha z^{-2}} \right] \quad (4)$$

with $\alpha < 1$ for stability. The transfer function can be realized with only one multiplication and three additions.



Fig. 4. Comparison of original and coded versions of “Lena”. Top Left: Original. Top Right: Decimation Rate 3/2x3/2 Bit Rate 0.97 bits/pixel. Bottom Left: Decimation Rate 2x2 Bit Rate 0.55 bits/pixel. Bottom Right: Decimation Rate 3x3 Bit Rate 0.25 bits/pixel.

It is known that the latest image compression standard JPEG2000 is able to offer high quality performance at low bit rates. In terms of peak-to-peak signal-to-noise-ratio (PSNR), JPEG2000 has clearly an advantage (Figure 2). However in terms of subjective quality, the difference is barely noticeable (Figure 5). Here we compare the subjective quality by processing the 512x512, monochrome image “Peppers”. In Figure 6 one can see that at very low bit rates, JPEG2000 coder although has better edge sharpness, it “smears out” the object surface. For our proposed codec, the sharpness is compromised due to decimation but the retained texture of the objects in the image is clearly closer to the original.

Finally, we would like to comment on the complexity of the proposed codec. JPEG2000 requires at least 5 multiplications per pixel to achieve sufficient wavelet transform decomposition plus additional complexity from the context based arithmetic coder [8]. For our coder, 3 multiplications per pixel are needed for prediction and variance calculation for adaptive entropy coding. The Wiener filter can be implemented using FIR filters to lower computational

complexity [3]. Again since this is operated on decimated image, the actual multiplication per pixel is reduced depending on the decimation rate. As stated earlier decimation filters can be realized in very simple IIR filter forms so the overall complexity remains low.

5. CONCLUSIONS

In this paper, we proposed a practical image coder based on our modified DPCM codec with multi-rate processing. Simulation results show that the proposed coder provides great improvement over conventional DPCM coders and is able to outperform JPEG coders at low bit rates. In addition to the rate reduction by reducing the number of coded sample through down-sampling, we adopted the adaptive entropy coding scheme which provides further improvements of the source coding performance. The key feature of the image coder is its relatively low complexity. The low-pass decimation and interpolation filters can be implemented using very simple IIR filters. Very few down-sampling rates are needed to maintain the coder gain close to the ones provided by the optimal operating points. The inclusion of the Wiener filter further ensures less degradation when the coder is operating away from the ideal operation points. The adaptive entropy coding is a small and simple addition to the original codec but gives substantial rate gain. The necessary side information to be transmit to the receiver is little and can be efficiently coded. Our image coder is shown to be a good alternative for coding images with monotonically decreasing spectra at low bit rates.

6. REFERENCES

- [1] A.N. Kim and T.A. Ramstad, "Improving the rate distortion performance of dpcm," in *Proceedings of Int. Symposium on Signal Proc. and its Applications (ISSPA)*, 2003.
- [2] O. Guleryuz and M. T. Orchard, "On the dpcm compression of gaussian autoregressive sequences," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 945–956, 2001.
- [3] A. K. Jain, *Fundamentals of Digital Image Processing*, Information and System Science. Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A., 1st edition, 1989.
- [4] N.S. Jayant and P. Noll, *Digital Coding of Waveforms - Principles and Applications to Speech and Video*, Signal Processing. Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A., 1st edition, 1984.
- [5] J.M. Lervik and T.A. Ramstad, "Optimality of multiple entropy coder systems for nonstationary sources modelled by a mixture distribution," in *Proceedings of Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*, 1996.
- [6] J.M. Lervik, *Subband Image Communication Over Digital Transparent And Analog Waveform Channels*, PhD Dissertation, Department of Telecommunications, NTNU, Trondheim, Norway, 1996.
- [7] N. Viswanadham R.N. Madan and R.L. Kashyap, *Systems and Signal Processing*, Oxford and IBH, New Delhi, India, 1991.
- [8] D.S. Taubman and M.W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwar Academic Publishers, 2002.



Fig. 5. Subjective quality comparison of JPEG2000 and the proposed codec. Top: Original. Bottom left: Proposed codec. Decimation rate 2x2, Bit rate 0.42bpp; Bottom Right: JPEG2000 Bit rate 0.42bpp



Fig. 6. Comparison of JPEG2000 and the proposed codec. Left: Proposed codec. Decimation rate 2x2, Bit rate 0.27bpp; Right: JPEG2000 Bit rate 0.27bpp