To be published in the Proceedings of the International Conference on Image Processing
26-29 October 1997, Santa Barbara, USA.

Bit allocation methods for closed-loop coding of oversampled pyramid decompositions

Uwe Horn, Thomas Wiegand, and Bernd Girod

Telecommunications Institute
University of Erlangen-Nuremberg
Cauerstrasse 7/NT, D-91058 Erlangen, Germany
email: [uhorn|wiegand|girod]@nt.e-technik.uni-erlangen.de

Abstract

Oversampled pyramid decompositions have been successfully applied to scalable video coding. Quantization noise feedback at the encoder ("closed-loop coding") has several advantages compared to "open-loop coding" where noise feedback is not included. Finding optimal bit allocations for closed-loop coders is a complex task since the rate-distortion ($R-D$) performance of one quantizer depends on the selected operational points of all previous ones. We present simplified models for closed-loop and open-loop coding. From those, analytical solutions to the optimal bit allocation problem can be derived. In the optimal case, closed-loop outperforms open-loop coding although distortion in lower resolution layers of the closed-loop coder may become rather high which is undesirable in scalable coding applications. In this case it turns out to be a good solution to use an optimal open-loop bit allocation for closed-loop coding. Theoretical results are verified by coding experiments.

1 Introduction

Oversampled pyramid decompositions for scalable video coding have been proposed e.g. in [1]. They exhibit several advantages compared to critically subsampled subband coding schemes mainly because motion compensation can be easily incorporated [2, 3]. Such an oversampled representation can be encoded with a pyramid coder as shown in Fig. 1. The coder works either in closed-loop (CL) or in open-loop (OL) mode. The decoder is equal in both cases. As can be seen, the input signal is successively filtered and downsampled. In OL-mode, for each layer a bandpass signal is formed by subtracting the interpolated lower resolution signal from the spatially subsampled original signal within that layer. The bandpass signals are quantized and transmitted to the decoder. Note that we assume vector quantizers (VQ) to be used for quantization. From the received layers the decoder can reconstruct the different spatial resolutions as shown.

In contrast to OL-coding, CL-coding uses the already reconstructed signals for interpolation. Therefore, the layer bandpass signals contain additional quantization noise introduced by the quantizer of the lower resolution layer. Although the overall quantization noise can be controlled by the last quantizer, its $R-D$ performance strongly depends on the number of bits allocated to the previous quantizers. This is the reason why finding optimal bit allocations for CL-coders is quite a complex task [4].

In this paper, we first present simple Gaussian $R-D$ models for OL- and CL-coding where we assume high bit-rate, fine quantization coding. From these models, we derive general solutions to the bit allocation problem for OL- and CL-coding.[1] This allows us to compare the performance of both coding approaches. In the optimal case, CL- outperforms OL-coding although distortion in lower resolution layers of the CL-codec may become rather high which is undesirable in scalable coding applications. In this case a good solution is to use an optimal OL-coder bit allocation for CL-coding.

2 Open-loop coding

Fig. 2 shows a model where we have replaced the quantizers of an OL-coder by additive white noise. We neglect the quantization of the lowpass layer which is generally very fine and has little influence on the $R-D$ functions of the higher resolution layers. The overall mean squared error distortion in the full resolution layer 0 is computed as sum of the noise variances contributed by the quantizers within the different layers.

[1] The calculations which lead to the solutions given in Eqs. (2) and (4) can be found at: http://www-nt.e-technik.uni-erlangen.de/~wiegand/icip97-proof.ps.gz

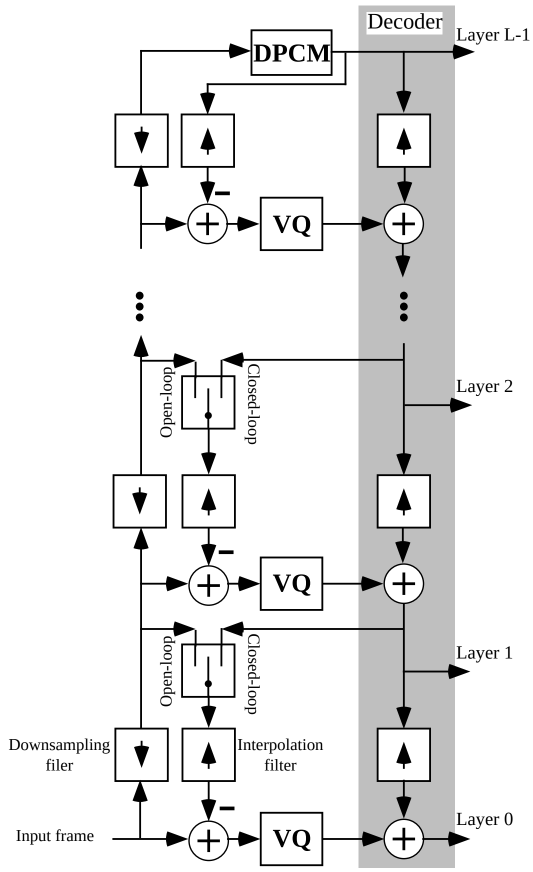

Figure 1: N-layer pyramid codec

The noise contributions are modeled by weighting the quantization noise $d_l(r_l)$ in layer l by an appropriate factor α_l which depends on l and the chosen interpolation filter. Note that α_1 corresponds to the power transfer factor (PTF, [5]) of the interpolation filter and that $\alpha_0 = 1$. If we assume a Gaussian source, we obtain for an L-layer decomposition the following overall distortion and rate equations:

$$D_{OL}(R) = \sum_{l=0}^{L-1} \alpha_l \cdot g_l \cdot \sigma_l^2 \cdot 2^{-2r_l} \qquad (1)$$

$$R = \sum_{l=0}^{L-1} n_l \cdot r_l .$$

Here g_l takes into account the spectral flatness [5] and the σ_l^2 are the variances of the various interpolation error signals. $n_l = N_l/N_0$ is defined as the ratio between N_l the number of samples in layer l and N_0 the number of samples in the full resolution layer 0. The redundancy of the oversampled decomposition is therefore expressed by $M = \sum_{l=0}^{L-1} n_l$.

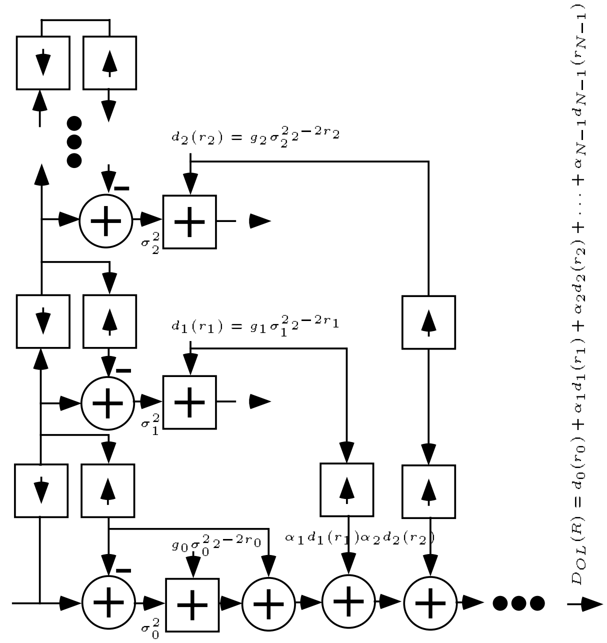

Figure 2: Open-loop model

We can solve the bit allocation problem like for critically sampled subband coding schemes [6] and obtain as solution

$$r_l = \frac{R}{M} + \frac{1}{2}\log_2 \frac{\alpha_l \cdot g_l \cdot \sigma_l^2 / n_l}{\prod_{k=0}^{L-1} (\alpha_k \cdot g_k \cdot \sigma_k^2 / n_k)^{\frac{n_k}{M}}} \qquad (2)$$

which is quite similar to the well-known solution for critically sampled subband coding schemes. Our solution is more general because it can also deal with oversampled decompositions.

3 Closed-loop coding

Fig. 3 shows the corresponding model for CL-coding. Due to noise-feedback, filtered noise introduced by the previous quantizer is added to the bandpass signal *before* quantization. If we neglect that g_l depends on $r_{l+1}, \ldots, r_{L-1}$, the overall distortion in the closed-loop case is:

$$D_{CL}(R) = \sum_{l=0}^{L-1} \alpha^l \cdot g_l \cdot \sigma_l^2 \cdot 2^{-2\sum_{k=0}^{l} r_k} . \qquad (3)$$

R is computed as in Eq. 1. We can find an analytical solution to the corresponding bit allocation problem as:

$$r_l = \begin{cases} \frac{1}{2}\log_2\left(\frac{\alpha \cdot g_l \cdot \sigma_l^2}{g_{l-1} \cdot \sigma_{l-1}^2} \cdot w_l\right), & l > 0 \\ R - \sum_{l=1}^{L-1} n_l \cdot r_l, & l = 0 \end{cases} \qquad (4)$$

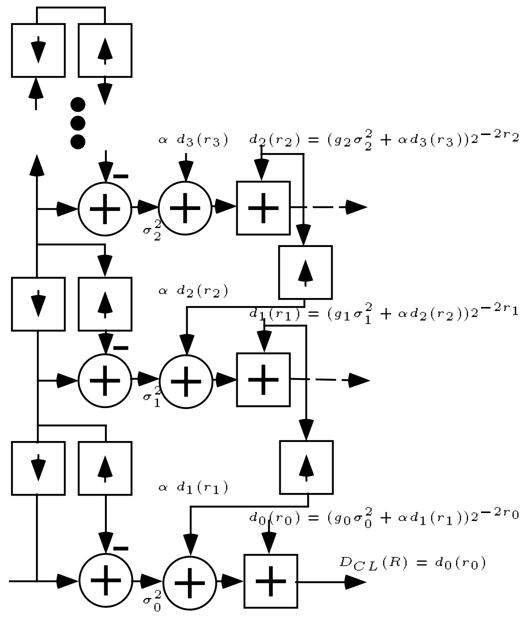

Figure 3: Closed-loop model

$$w_l = \begin{cases} \frac{n_{l-1}-n_l}{n_l-n_{l+1}}, & \frac{n_{l-1}-n_l}{n_l-n_{l+1}} > 0 \\ 1, & n_l = n_{l-1} = n_{l+1} \end{cases}, n_L = 0$$

Recently an equivalent solution for the special case of a regular decomposition with $n_l = \frac{1}{4^l}$ has been published independently from our work in [7]. In contrast to OL-coding, r_l for $l > 0$ is independent of the overall bitrate R and determined only by the statistics of the bandpass signals! In layer 0 the remaining bits after quantization of all other layers are spent. The drawback of the optimal solution is that the distortion *within* lower resolution layers can become quite high compared to the full resolution distortion which is undesirable for scalable coding applications. As shown in the next section, a good solution to this problem is to use an optimal *open-loop* bit allocation for CL-coding.

4 Comparison

In Fig. 4, we compare the behavior of different bit allocation and coding methods for a three layer decomposition according to the models given above. As mentioned before, we neglect quantization of the lowpass signal in the lowest resolution layer. For the remaining layers 1 and 0 we plot layer distortion expressed as SNR in dB versus overall rate R. Three different cases are considered: OL-coding with optimal bit allocation according to Eq. 2, CL-coding with

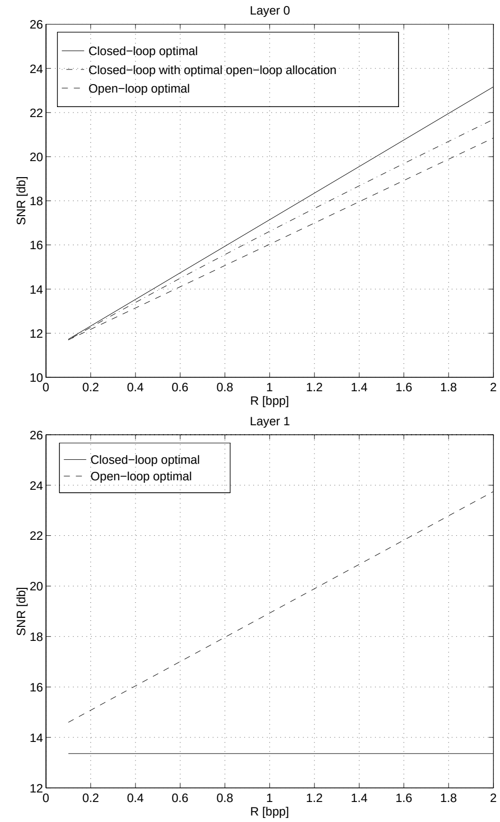

Figure 4: Different bit allocation and coding methods

optimal bit allocation according to Eq. 4 and optimal OL-coder bit allocation used for CL-coding. Note that the first and the third case lead to identical results for layer 1. The following parameters are used: $\sigma_0^2 = 0.054, \sigma_1^2 = 0.061, g_0 = g_1 = 1, \alpha = 0.39$. The variances are obtained from a decomposition of the first frame of the CIF video test sequence 'Students', normalized to the energy contained in the full resolution input signal. The value of α corresponds to the PTF of an interpolation filter which is used in an existing pyramid codec implementation [3] and given later in this section.

As can been seen from Fig. 4, CL-coding with optimal bit allocation gives the best performance in layer 0 but the worst in layer 1. The gain obtained from CL-coding compared to OL-coding increases with increasing overall rate.

The disadvantage of optimal CL-coder bit allocations is, that lower resolution layers are coded with rather high distortion. As can be seen, using optimal OL-coder bit allocation for CL-coding is a good compromise. Distortion in layer 1 decreases dramatically, while the overall distortion in layer 0 increases by only

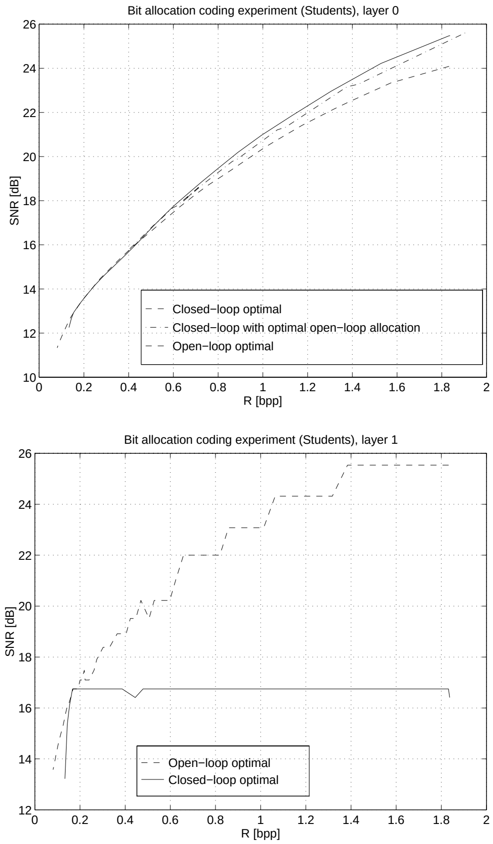

Figure 5: Comparison by a coding experiment

a small amount.

Fig. 5 shows results obtained by a coding experiment where compression is achieved by an E_8-lattice vector quantizer followed by entropy-coding. We use $[\frac{1}{2}\ \frac{1}{2}]$ as downsampling and $[\frac{1}{4}\frac{3}{4}\frac{3}{4}\frac{1}{4}]$ as interpolation filter both applied separately in horizontal and vertical direction [3]. The optimal solutions are found by tracing the convex hull of $D(R)$ for layer 0 for more than 1000 considered bit allocations. The $R-D$ behavior of the different bit allocation and coding methods is quite similar to that predicted from our theoretical models. Since we have assumed $g_l = 1$ for the plots of Fig. 4 and since a lattice vector quantizer is used for the coding experiments, absolute distortion values for a given rate differ from the model predictions. Again it can be seen that at higher bitrates CL-coding outperforms OL-coding. By using optimal OL-coder allocations for CL-coding the overall performance becomes no more than 0.2 dB worse compared to optimal CL-coding.

5 Conclusion

By using simplified models for pyramid coders based on a Gaussian $R-D$ model, analytical solutions for the optimal bit allocation problem in open-loop as well as in closed-loop pyramid coders can be obtained. A comparison between optimal closed-loop and open-loop coding shows that closed-loop outperforms open-loop coding in the full resolution layer. The disadvantage of closed-loop optimal bit allocations is that lower resolution layers are encoded with only a few bits independent of the overall bitrate. This is undesirable for scalable coding. A good heuristic is to use optimal open-loop bit allocations for closed-loop coding. The overall performance decreases only by a small amount. Real coding experiments show the usefulness of the presented models and the derived bit allocation heuristic for closed-loop coding.

References

[1] M.K. Uz, M. Vetterli, and D.J. LeGall. Interpolative multiresolution coding of advanced television with compatible subchannels. *IEEE Trans. on Circuits and Systems for Video Technology*, 1(1):86–99, Mar. 1991.

[2] B. Girod, U. Horn, and Ben Belzer. Scalable video coding with multiscale motion compensation and unequal error protection. In Y. Wang, S. Panwar, S.-P. Kim, and H. L. Bertoni, editors, *Multimedia Communications and Video Coding*, pages 475–482. Plenum Press, New York, Oct. 1996.

[3] U. Horn and B. Girod. Performance analysis of multiscale motion compensation techniques in pyramid coders. In *Proc. ICIP'96*, volume III, pages 255–258, 1996.

[4] K. Ramchandran, A. Ortega, and M. Vetterli. Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders. *IEEE Trans. on Signal Processing*, 3(5):533–545, Sep. 1994.

[5] N.S. Jayant and P. Noll. *Digital Coding of Waveforms*. Prentice-Hall, Englewood Cliffs, New Jersey, 1984.

[6] J.W. Woods (ed.). *Subband Image Coding*. Kluwer Academic Publishers, Boston, 1991.

[7] L. Alparone. Quantization noise feedback in Laplacian pyramid-based image coding: A rate-distortion approach. In *Proc. IEEE DSP 97*, pages 849–852, Santorini, Jul. 1997.