NOISE PROCESSING FOR SIMPLE LAPLACIAN PYRAMID SYNTHESIS BASED ON DUAL FRAME RECONSTRUCTION

Aditya Mavlankar, David Chen, Sameh Zakhary, Markus Flierl, and Bernd Girod

Information Systems Laboratory
Department of Electrical Engineering, Stanford University
Stanford, CA, 94305, U.S.A.
{maditya, dmchen, zakhary, mflierl, bgirod}@stanford.edu

ABSTRACT

The Laplacian pyramid (LP) provides a frame expansion. Thus, there exist infinitely many synthesis operators which achieve perfect reconstruction in the absence of quantization. However, if the subbands are quantized in the open-loop mode then the dual frame synthesis operator, which is the pseudo-inverse of the analysis operator, minimizes the mean squared error (MSE) in the reconstruction. Note that this requires modification of the conventional simple synthesis scheme. For the open-loop mode, we propose novel quantization noise processing at the encoder that allows us to achieve the same performance as dual frame reconstruction and yet retain the simple synthesis scheme at the decoder. This has the advantage that the decoder can be simple in structure as well as be agnostic of whether the encoder was open-loop or closed-loop and achieves minimum MSE reconstruction for both cases. Experimental results show a gain of around 1 dB with the dual frame reconstruction compared to the simple synthesis operator. Furthermore, experiments confirm that this gain can also be obtained by retaining the simple synthesis operator and performing the proposed quantization noise processing at the encoder.

Index Terms— Laplacian pyramid, dual frame, framed pyramid.

1. INTRODUCTION AND MOTIVATION

In 1983, Burt and Adelson [1] proposed the Laplacian pyramid (LP) as a multiresolution image representation. In the first stage of the decomposition, the original image is transformed into a coarse signal and a detail signal. The coarse signal has fewer samples than the original image but the detail signal has the same number of samples as the original image. The coarse signal is a filtered and downsampled version of the original image. It is then upsampled and filtered to predict the original image. The prediction residual constitutes the detail signal. The coarse signal can be decomposed further and this process can be repeated a few times iteratively. This yields a pyramid consisting of the coarsest version thus obtained and the detail signals at various scales. Thus, the LP is an overcomplete representation. For the purpose of image compression, discrete wavelet transforms and other critically sampled filterbanks are usually preferred since they lead to more efficient rate-distortion performance. However, the LP has certain advantages. The overcomplete representation provides freedom in the design of filters, allows the LP to support multiple reconstruction methods, and reduces the aliasing in the highpass subbands. Aside from compression, the LP has been used in several coarse-to-fine algorithms in image processing, for example [2]. Some scalable video coding schemes have also used modified LP representations for each successive frame of the video as a starting point for interframe coding, for example [3]. However, most applications in the literature employ the simple synthesis operator to reconstruct the image from the LP. In 2003, based on their observation that this longstanding multiresolution pyramid is in fact a frame expansion, Do and Vetterli [4] pointed out that this simple synthesis operator is not optimal in terms of minimizing the distortion propagation from the subbands of the LP to the reconstructed image in the open-loop mode. In [4], they also propose a practical structure which implements the dual frame reconstruction that is optimal. Furthermore, they expound conditions on the filters such that the proposed practical structure is in fact the pseudo-inverse, a.k.a. the dual frame, of the LP analysis frame. In this paper, we propose novel quantization noise processing at the encoder which allows us to use a simple synthesis operator and yet obtain similar gains as offered by the dual frame reconstruction.

2. DUAL FRAME RECONSTRUCTION

A single level of decomposition along with the conventional simple synthesis operator implementation is shown in Fig. 1. In order to simplify the mathematical representation and without loss of generality, we consider 1-D signals and combine the filtering operation denoted by F_H and the downsampling by 2 into one matrix operator H. Similarly, we combine the upsampling by 2 and the filtering operation denoted by F_G into one matrix operator G. If F_H and F_G are associated with filter kernels $h[\bullet]$ and $g[\bullet]$ respectively, then the matrices H and G have the following structure:

$$H = \begin{bmatrix} \ddots & & & & & & & \\ \cdots & h[2] & h[1] & h[0] & \cdot & \cdot & & \\ & \cdot & \cdot & h[2] & h[1] & h[0] & \cdots & \\ & & & & & & \ddots & \end{bmatrix}$$

$$G = \begin{bmatrix} \ddots & \vdots & & \\ & g[0] & \cdot & \\ & g[1] & \cdot & \\ & g[2] & g[0] & \\ & \cdot & g[1] & \\ & \cdot & g[2] & \\ & & \vdots & \ddots \end{bmatrix}$$

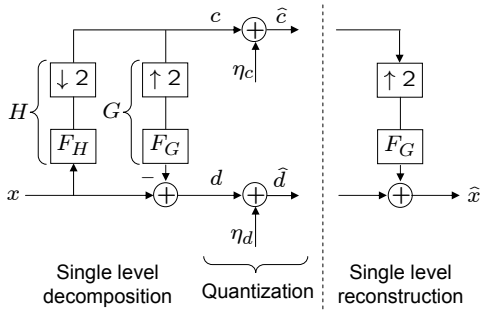

Fig. 1. Single level decomposition yields coarse signal c and difference signal d. These are then quantized, which is represented by addition of quantization noise η_c and η_d. Also shown is the conventional reconstruction.

Now let the input image be represented by a vector x. Then the analysis operation is given by:

$$y = \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} H \\ I - GH \end{bmatrix} x = Ax$$

Since the analysis operator can be represented by left multiplication with a tall matrix A, there are infinitely many synthesis operators, which are left-inverses of A and achieve perfect reconstruction in the absence of quantization. The synthesis frame which is the dual frame, given by $S_{opt} = A^{\dagger} = (A^T A)^{-1} A^T$, achieves the minimum distortion propagation from the subbands of the LP to the reconstructed image. Do and Vetterli [4] show that if the filters satisfy certain conditions then the synthesis structure shown in Fig. 2 (b) implements the pseudo-inverse. The conventional synthesis is shown in Fig. 2 (a) for comparison. The conditions on the filters are repeated here for clarity:

- The structure implements a left-inverse *if and only if* H and G are biorthogonal with respect to the sampling factor, i.e., $HG = I$.
- Furthermore, the structure implements a pseudo-inverse *if and only if* GH is an orthogonal projector, i.e., in addition to $HG = I$ as above, $H^T G^T = GH$ also holds.

As a special case, if the filters are orthogonal with respect to the sampling factor then $H = G^T$ and $G^T G = I$ and the structure in Fig. 2 (b) implements the dual frame since both the conditions above are met. It was shown in [4] that even in the case of biorthogonal filters, for example the 9-7 biorthogonal filters, which satisfy the first condition and are only approximately orthogonal, the structure in Fig. 2 (b) yields a gain of about 1 dB over the simple synthesis scheme.

3. QUANTIZATION NOISE PROCESSING AT ENCODER

3.1. Single Decomposition Level

Compared to the simple synthesis scheme, the implementation for the dual frame has one additional branch. Notice that if $HG = I$, then $Hd = H(x - GHx) = 0$. This means that the additional branch in Fig. 2 (b) is subtracting $H\eta_d$ from $\hat{c}$, the received coarse signal. To imitate this behavior at the encoder, we propose to subtract $H\eta_d$ from the coarse signal prior to its quantization at the encoder; this way we can retain the simple synthesis scheme and still achieve

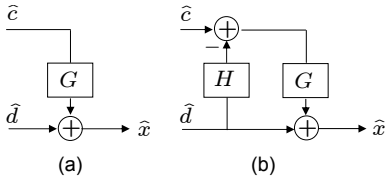

Fig. 2. (a) Conventional reconstruction with simple synthesis operator; it is suboptimal. (b) Pseudo-inverse reconstruction with modified synthesis operator; it is optimal.

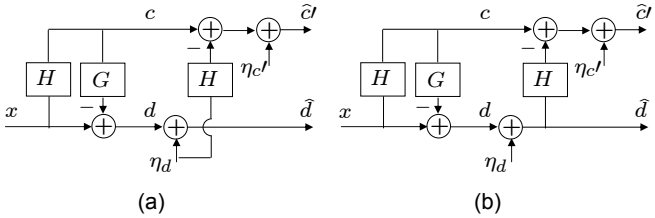

Fig. 3. (a) Proposed quantization noise processing at encoder. (b) Simplification due to $Hd = 0$. The simple synthesis scheme shown in Fig. 2 (a) can be used after this.

dual frame reconstruction. This is shown in Fig. 3 (a). Note that since $Hd = 0$, this simplifies to the structure in Fig. 3 (b).

The are two differences between the proposed method and the dual frame reconstruction proposed in [4]. The first difference is that if the same quantizer is used on the coarse signal as in Fig. 1, then the quantization noise added is slightly different; this is denoted by $\eta_c\prime$ instead of η_c. But this does not affect performance in any noticeable way, as demonstrated in Section 5. The second difference is that if the detail signal is modified *after* the quantization noise processing, i.e., the decoder receives a copy which has some noise in addition to η_d, then the simple synthesis scheme cannot guarantee the same gains as the dual frame reconstruction proposed in [4].

Notice that $Hd = 0$ implies that the detail signal is orthogonal to every row of H. Since the rows of H are linearly independent, they span a subspace with dimensionality equal to the number of rows of H. This means that the detail signal d is restricted to a lower dimensional subspace. This has been recently exploited to compress d further, for example in [5]. Note that the approach in [5] aims to produce a critical representation of the LP and is different from the dual frame reconstruction approach.

3.2. Multiple Decomposition Levels

The generic block diagram of a conventional LP encoder with $N > 1$ decomposition levels is shown in Fig. 4. The extension of this conventional encoder with $N > 1$ levels to include the proposed quantization noise processing is not unique. A straightforward extension is shown in Fig. 5 (a). Another possible extension is shown in Fig. 5 (b), which attempts to keep the coarse signal cleaner since the successive levels of decomposition do not decompose the quantization noise fed back in the levels below. The input signal for the first level is $x_1\prime = x_1 = x$. Also for the scheme in Fig. 5 (b), $d_1 = d_1\prime\prime$. Notice that in Fig. 5 (b), we subtract $H\eta_{d_i}\prime\prime$ from the next subband since $Hd_i\prime\prime$ might not be zero.

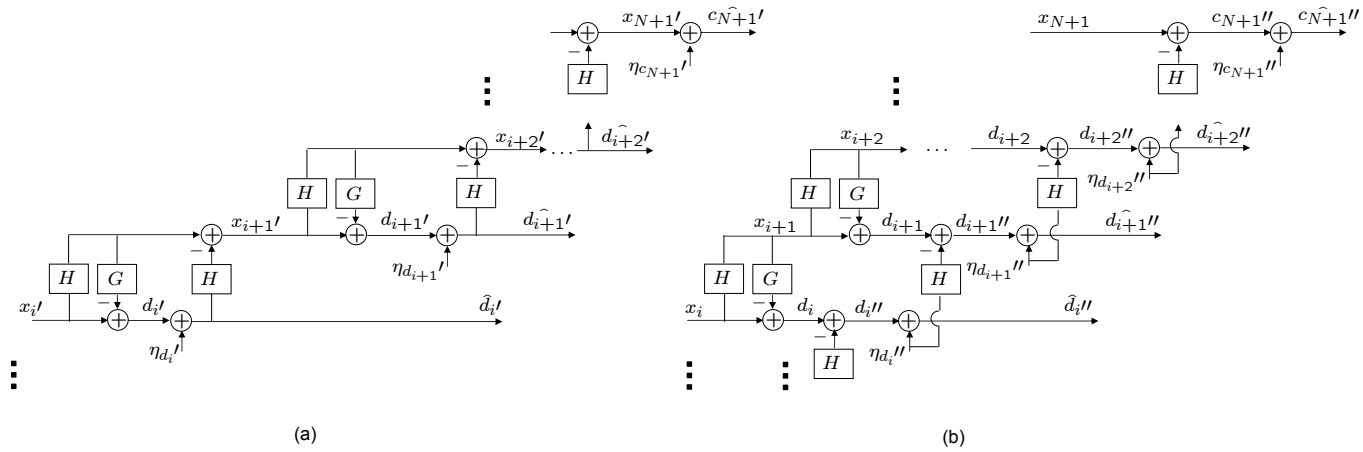

Fig. 5. Extension of encoder-side quantization noise processing to multiple levels. (a) Straightforward extension. (b) Possible extension which attempts to keep the coarse signal cleaner during the decomposition process compared to (a).

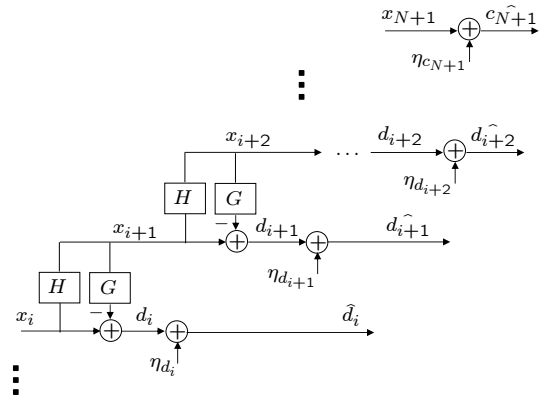

Fig. 4. Generic block diagram of conventional LP encoder consisting of $N > 1$ decomposition levels. This yields N detail subbands and 1 coarse subband; for example, $N = 2$ yields subbands $\hat{d}_1$, $\hat{d}_2$ and $\hat{c}_2$. The original signal x is denoted by x_1 here.

4. RATE ALLOCATION

The rate allocation algorithm ensures a good rate-distortion performance by spending the available bit-budget judiciously for coding the different subbands. If there are N subbands and we spend R_b bits per coefficient for coding the b^{th} subband then the distortion in the reconstructed image can be modeled as

$$D(R) = \sum_{b=1}^{N} \alpha_b G_b D_b(R_b), \qquad R = \sum_{b=1}^{N} \alpha_b R_b,$$

where R is the resulting bit per pixel for coding the image, α_b is the fraction of all coefficients in the b^{th} subband, G_b is the synthesis gain for the b^{th} subband and $D_b(\cdot)$ is the empirical distortion-rate function for coding the b^{th} subband individually. There is a closed-form solution at high rates, as given in [6], to assign R_b bits for coding the b^{th} subband given by

$$R_b = R + \frac{1}{2}\log_2\left(\frac{G_b\varepsilon_b^2\sigma_b^2}{\prod_{n=1}^{N}(G_n\varepsilon_n^2\sigma_n^2)^{\alpha_n}}\right)$$

where σ_b^2 is the variance in the b^{th} subband and ε_n^2 is 1 or 0.87 if the subband distribution is Gaussian or Laplacian, respectively. At high rates, once the rate R_b is known, a uniform scalar quantizer can be chosen with step-size

$$\Delta_b = \sqrt{12\varepsilon_b^2\sigma_b^2} \times 2^{-R_b}.$$

In the low-rate regime, we use the above formulas to get an initial estimate of the optimal subband rates. A local search of subband rates is then performed to yield better rate-distortion trade-off.

5. EXPERIMENTAL RESULTS

We used the luminance component of the images, *Lena*, *Einstein*, *Barbara* and *Cameraman* with 512x512 pixels each for our experiments. Fig. 6 shows the rate-distortion performance for the *Lena* image with one and four levels of decomposition. Fig. 7 shows the results for the *Einstein* image. The 9-7 filters have been used for the decomposition in the LP. The schemes depicted in Fig. 5(a) and (b) are denoted as noise processing A and B respectively in the rate-distortion plots and their performance is very similar. Notice that for a single level of decomposition, the two methods are identical.

The subbands obtained are quantized to yield $c_{\hat{N}+1}\prime$, $\hat{d}_i\prime$ with $i = 1 \ldots N$ for scheme A and $c_{\hat{N}+1}\prime\prime$, $\hat{d}_i\prime\prime$ with $i = 1 \ldots N$ for scheme B. These quantized subbands are then run-length coded in raster scan using individual Huffman tables which are conveyed in the bit-stream. The Huffman tables are computed from the empirical statistics of the respective quantized subband. The rate allocation described in Section 4 yields an initial set of quantizers for the subbands. If the quantizer step-size is greater than the dynamic range for any subband then the quantizer step-size is set equal to or less than the dynamic range. For the quantizer set, the Huffman tables are computed as well as the total rate. If the total rate exceeds the target rate then the quantizer set is made coarser by multiplying all step-sizes with a factor of 1.1. Similarly, if the total rate is below the

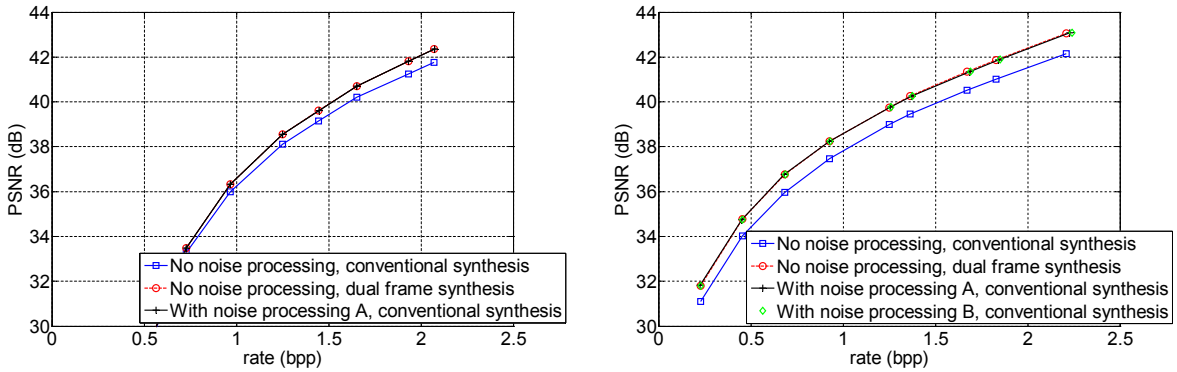

Fig. 6. Rate-distortion performance for the *Lena* image. One level of LP decomposition (*left*) and four levels of LP decomposition (*right*).

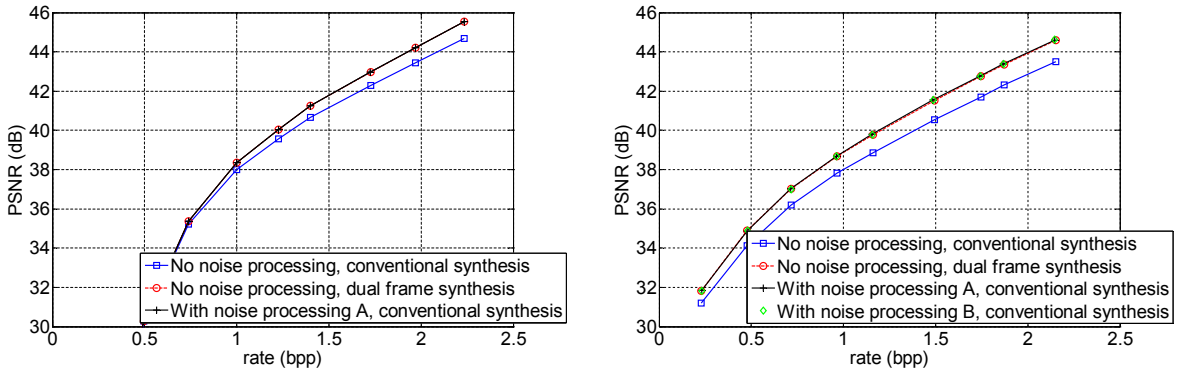

Fig. 7. Rate-distortion performance for the *Einstein* image. One level of LP decomposition (*left*) and four levels of LP decomposition (*right*).

rate constraint then the quantizer set is made finer by multiplying all step-sizes with a factor of 0.95. If the total rate is within 5% of the target rate, then the local search is terminated. This local search is less computationally expensive than an exhaustive search that also varies the ratios of the step-sizes.

It can be seen that the dual frame reconstruction offers a gain of about 1 dB over the simple synthesis scheme. This gain can also be achieved using quantization noise processing at the encoder and retaining the simple synthesis at the decoder. Fewer decomposition levels have the disadvantage of coding more lowpass coefficients but the advantage that the LP representation is less overcomplete. It can be seen that at high rates it pays to have fewer decomposition levels and at low rates it pays to have more decomposition levels. Similar rate-distortion performance was observed for the other images as well. The total complexity of the encoder and the decoder combined is the same as in the dual frame reconstruction approach. The proposed schemes shift roughly half of the complexity from the decoder side to the encoder side.

6. CONCLUSIONS

For the Laplacian pyramid, we consider the case when the dual frame reconstruction is possible using the decoder structure as proposed in [4]. For encoding in the open-loop mode, we propose novel quantization noise processing at the encoder that allows us to achieve the same performance as dual frame reconstruction and yet retain the simple synthesis scheme at the decoder. We assume that the detail subband does not undergo further distortion after being used for the noise processing at the encoder. The proposed method simplifies the decoder. Furthermore, the decoder is identical for open-loop and closed-loop encoding and achieves minimum MSE reconstruction for both cases. As argued in [4], the simple synthesis scheme generally gives the minimum MSE for the closed-loop encoder.

7. REFERENCES

[1] P.J. Burt and E.H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. on Communications*, vol. 31, no. 4, pp. 532–540, April 1983.

[2] P.J. Burt and E.H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Trans. on Graphics*, vol. 2, no. 4, pp. 217–236, 1983.

[3] M. Flierl and P. Vandergheynst, "Inter-resolution transform for spatially scalable video coding," *Proc. of Intl. Picture Coding Symposium (PCS '04), San Francisco, CA, USA*, December 2004.

[4] M.N. Do and M. Vetterli, "Framing pyramids," *IEEE Trans. on Signal Processing*, vol. 51, no. 9, pp. 2329–2342, September 2003.

[5] G. Rath and C. Guillemot, "Compressing the Laplacian pyramid," *Proc. of IEEE Workshop on Multimedia Signal Processing (MMSP '06), Victoria, Canada*, pp. 75–79, October 2006.

[6] D.S. Taubman and M.W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer, Norwell, MA, 2002.