EMBEDDED IMAGE CODING BASED ON LAPLACIAN PYRAMIDS WITH QUANTIZATION FEEDBACK

Bruno Aiazzi, Stefano Baronti, Franco Lotti
Nello Carrara Research Institute on Electromagnetic Waves IROE - CNR
Via Panciatichi, 64 - 50127 Florence, ITALY
e-mail: baronti@iroe.fi.cnr.it

Luciano Alparone
Department of Electronic Engineering, University of Florence
Via S. Marta, 3 - 50139 Florence, ITALY
e-mail: alparone@cosimo.ing.unifi.it

ABSTRACT

In this paper, a multi-layer SNR-scalable error-bounded image encoder is achieved in the framework of Laplacian pyramids with quantization noise feedback, by exploiting an entropy-minimizing optimum quantization strategy, a content-driven decision rule based on an L_∞ activity measure, and multistage quantizers to progressively upgrade quality at full scale. The resulting scheme yields intermediate versions with scale and SNR both increasing, and a further SNR scalability on the full resolution, with possibly lossless reconstruction, thereby expediting interactive browsing of remote data bases of images of any sizes and wordlength. The proposed encoder outperforms JPEG which does not possess all the above mentioned attractive characteristics.

1 INTRODUCTION

Progressive Image Transmission (PIT) [8] is more suitable than *block* image coding, e.g., JPEG [6], or *raster* image coding, e.g., Differential Pulse Coding Modulation (DPCM) [5], for image data retrieval and telebrowsing of image archives. Namely, low resolution coarse image versions are spatially expanded and upgraded to a higher resolution by adding refinements that constitute subtler and subtler details (spatial scalability) [10]. Alternatively, quantization residuals are recursively encoded by means of either multistage or embedded quantizers [8] (SNR scalability). In both cases, the error-free reconstruction may be possibly achieved. However, depending on the user's requirements, reception may stop at an intermediate decoding stage, thus yielding a lossy version, possibly at a lower spatial scale.

Image pyramids [3, 10, 2] have been found attractive for compression because of their characteristic progressiveness. The technique of quantization noise feedback in Laplacian pyramids (LP), besides being the key to lossless compression [10], is valuable whenever the L_∞ of the error has to be strictly upper-bounded (*near-lossless coding*) as, e.g., in contribution video, where post-production requires a high and homogeneous quality, and in medical imaging, where maximum errors have to be limited for legal reasons [9].

When prediction loops of quantization errors are enclosed in the pyramid scheme, errors coming from the top are interpolated onto underlying levels in which new errors are also introduced. Therefore, the quantizers at each level are no longer independent of each other [7]. The reconstruction error is determined by the quantization step at the base of the pyramid: hence, only one unit-sized step warrants lossless compression, that otherwise is *near-lossless*. The other step sizes might be chosen arbitrarily, without practically affecting the overall distortion; they are usually demanded by quality requirements of the scaled versions (*token images*).

The joint SNR-spatial scalability feature of feedback pyramids has been exploited to yield high-SNR, full-resolution image version from low-SNR, low-resolution versions [9]. However, only one SNR, although unconstrained to scale, is available at each spatial scale. In fact, LPs can not independently scale SNR and spatial resolution, i.e., yield image versions at a same spatial resolution with progressively increasing SNR, unless either embedded or multistage quantizers are employed at those layers in which SNR scalability is demanded.

In this work, quantization error feedback is employed in a modified version of Burt's LP, enhanced by means of a half- band interpolation filter (ELP) [2]. The *closed-loop* entropy-minimizing optimum quantizer used throughout was previously investigated for a multi-layer pyramid scheme [1]. As the differential pyramid still exhibits a local correlation, a *content-driven* decision strategy aimed at prioritizing the data to be coded is embodied in the feedback quantizer [4]. Errors stemming from both quantization and decision are simultaneously fedback. In order to allow SNR scalability, a two-stage quantizer is employed at the full-scale pyramid layer. The resulting multi-layer image encoder may be either lossless or near-lossless, although it benefits more from the content-driven feature in the lossy case.

2 A PYRAMID WITH NOISE FEEDBACK

Let $G_0 = \{G_0(m, n)\}$, be the $M \times N$ integer valued input image, with $M = p \times 2^K$, and $N = q \times 2^K$. The set of images $\{G_k, k = 0, 1, ..., K\}$ constitutes a Gaussian Pyramid (GP) [3], in which, for $k > 0$, G_k is a reduced version of G_{k-1}, where k identifies the pyramid level, and K the top level or *root*. *Reduction* corresponds to a separable linear lowpass filtering followed by *down-sampling* by two ($\downarrow$ 2). The *5-taps* Burt's parametric kernel $R(z) = a + \frac{1}{4}(z + z^{-1}) + (\frac{1}{4} - a/2)(z^2 + z^{-2})$ has been used for separable pyramid reduction. An *enhanced* Laplacian pyramid (ELP) [2, 4] $\{L_k, k = 0, 1, ..., K\}$, with $L_k = \{L_k(m, n)\}$ may be defined from GP: for $k = K$, $L_K \equiv G_K$, while for $k = 0, ..., K-1$ L_k is the difference between G_k and an expanded version of G_{k+1}, in which *expansion* signifies *up-sampling* by two ($\uparrow$ 2), i.e., interleaving with null samples, and linear lowpass filtering. A half-band kernel having five nonzero taps has been used for separable expansion: $E(z) = 1 + b(z + z^{-1}) + (\frac{1}{2} - b)(z^3 + z^{-3})$ corresponding to a parametric interpolation.

Laplacian pyramids achieve perfect reconstruction only in the case of infinite precision. When dealing with finite arithmetics, the output of $R(z)$ and $E(z)$ filters is rounded to integer, namely $[\cdot]$, in order to yield integer valued pyramids (denoted with a *). In this way, $G_0^* \equiv G_0$ may be exactly recovered from $L_k^* = G_k^* - [expand(G_{k+1}^*)], k = K-1, ..., 0$. Any step Δ_k can be used for quantizing L_k^*, provided that $\Delta_0 = 1$.

Let $Q_k(\cdot)$ indicate quantization with a step size Δ_k, namely, $Q_k(t) = [t/\Delta_k]$, and $Q^{-1}(\cdot)$ denote the inverse operation (i.e., $Q_k^{-1}(l) = l \cdot \Delta_k$). When non-unit step sizes are adopted, $\hat{G}_k^*$, the integer valued GP reconstructed at the receiving end for $k < K$, will be recursively given by the sum of the expanded $\hat{G}_{k+1}^*$ and of an approximate version of $\hat{L}_k^*$ due to quantization errors, in which $\hat{L}_k^* = G_k^* - [expand(\hat{G}_{k+1}^*)]$. Since the term $\hat{G}_{k+1}^*$ recursively accounts for previous quantization errors starting from the root level K down to level $k+1$ inclusive, setting $\Delta_0 = 1$ causes all errors previously introduced and delivered to the base of the pyramid to be compensated regardless of the other step sizes. Upper bounding of absolute errors is also guaranteed if $\Delta_0 > 1$.

An hybrid (pyramid + DPCM) encoder is outlined in Figure 1 for $K = 2$. Note that Q and Q^{-1} blocks are enclosed in the feedback loops. The K^{th} level of the GP (root) is DPCM-encoded for practical convenience.

The best choice of K in order to minimize entropy varies from one image to another. In fact, the pyramid exploits large-scale correlation, which becomes negligible for increasing K. Therefore, it is better to use the local correlation of the root by DPCM-encoding than to overly reduce the size of pyramid layers. Quantization errors of the root are interpolated as well and delivered to subsequent coding stages, as outlined in Fig. 1.

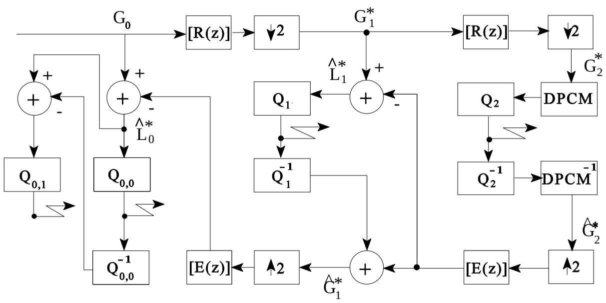

Figure 1: Block diagram of a three-layer hybrid (ELP + DPCM) encoder with quantization noise feedback and two-stage quantization at full scale.

A simple prediction scheme employs a linear regression of four neighboring pixels, three in the previous row and one in the current row. The optimal coefficients are found by Least Squares minimization of the residuals. For an Auto-Regressive (AR) model of the field, the inversion of a correlation matrix of a size equal to the prediction support is required [5].

The optimal choice of the set of quantization step sizes leading to the minimum equivalent entropy, $\{\hat{\Delta}_k, k = 1, ..., K\}$ has been derived in [1]: $\hat{\Delta}_k = 2\sigma_{k-1}^*/P_E$, in which P_E denotes the power gain of the expansion filter $E(z)$. The set $\{\sigma_k^*, k = 0, 1, ..., K-1\}$ represents the square roots of the variances of L_k^*, and can be measured by running the scheme of Fig. 1 with all unit-sized quantization steps. The step sizes $\hat{\Delta}_k$s may be approximated with odd integers, so as to minimize the maximum absolute quantization errors of the lower resolution image versions, without introducing significant performance penalty [1].

DPCM is employed on the root image to take advantage of its spatial redundancy. In this case, the optimum step size $\hat{\Delta}_K$ would no longer be optimal for minimizing entropy. Since the quantizer at level K is embedded in the DPCM loop (see Fig. 1), too coarse a quantization would reduce the prediction efficiency. The best value of $\Delta_k$is chosen empirically. A correct use of the error feedback strategy provides impressive benefits (12% to 15% in average) in the decorrelation performance [1].

The step size Δ_0 is established based on quality requirements, since the optimum quantizer is independent of its value. In the implementation of Fig. 1, the quantizer at level 0 is split into two cascaded stages to achieve SNR scalability at the full spatial scale. Errors stemming from the previous coarser quantizer $Q_{0.0}$ are quantized with a finer step size by the quantizer at the second stage, $Q_{0.1}$. In this way, a first full-scale image version is available at low cost, and the user can decide whether to upgrade it to a higher SNR, possibly to the lossless reconstruction, or to stop the retrieval process.

3 CONTENT-DRIVEN ENCODING

The scheme described in Sect. 2 has been further specialized in order to take advantage of the residual local correlation of the ELP and of image nonstationarity.

To this end, ELP levels are partitioned into adjacent 2×2 blocks. A quartet of nodes at level k is assumed to have one parent node at level $k+1$; also, each node from a quartet at level $k > 0$ is regarded as the parent node of a 2×2 underlying block at level $k-1$. A hierarchy is thus settled that enables the introduction of a split decision rule based on image content.

Content-driven transmission consists of a breadth-first tree-scan stage driven by a set of thresholds $\{T_k, k = 1, ..., K\}$ related to pyramid levels. An activity function of the parent node at level k, A_k, is computed on each underlying block at level $k-1$ of the ELP. If $A_k > T_k$, the four interpolation errors are encoded; otherwise, they are taken to be null. In this way, the information is prioritized: the most important pyramid coefficients are considered from the earliest stage on.

A critical point of the above outline is the choice of an efficient activity measure. In the present work, $A_k(m,n)$, the activity function of node (m,n) at level k, is taken as the L_∞ norm, or maximum absolute value of its four offspring at level $k-1$, i.e., four underlying interpolation errors. This choice represents a very simple, yet efficient, selection criterion.

The content-driven decision rule with uniform quantization of the retained quartets may be regarded as a data-dependent threshold quantization in which the quantization levels of the quartet of the coefficients to be discarded are all null, and quantization errors equal the values themselves. All the errors introduced by the decision rule can therefore be recovered at higher-resolution pyramid layers, thus extending quantization feedback also to content-driven pyramid schemes.

In practice, the set of thresholds can be related to the optimal steps of the uniform quantizer, e.g., by taking $\lfloor \hat{\Delta}_{k-1}/2 \rfloor < T_k \leq \hat{\Delta}_{k-1}$.

The introduction of a selective choice of the nodes to be split requires a synchronization overhead. Flag bits may be arranged to form a binary tree whose root corresponds to layer K and whose bottom corresponds to layer 1 of the ELP. Each one-bit marks the split of the related node into the quartet of its offspring. Conversely, each zero-bit indicates that the underlying quartet of interpolation errors has not been considered. A straightforward encoding of such a tree is not efficient because of the survived correlation. An easy way to exploit such a correlation is to run-length encode the sequences of zeroes and ones. Each level of the quad-tree is partitioned into square regions: each region is zig-zag scanned to yield runs of zeroes or ones, possibly continuing inside the next lower region. Experiments have shown that the average run-length is maximized when 8×8 regions are considered.

4 CODING RESULTS AND COMPARISONS

Tests were carried out to assess the compression capability of the ELP scheme, when the optimum quantizer was jointly employed with the content-driven strategy, in which errors caused by both quantization and activity-thresholding were simultaneously fed-back to be delivered throughout scales.

Figure 2: Multiscale encoding of Lena with SNR refinement at full scale through two-stage quantization. Coarser quantizer yields a 34.9 dB PSNR at cumulative 0.47 bpp; refinements yield a 38.8 dB at 1.0 bpp.

Coding results were compared with those of JPEG. Gross bit-rate in bits per pixel (bpp), with all overhead and side information including Huffman entropic codebooks, were considered throughout. Distortion was measured by means of the peak signal-to-noise ratio (PSNR).

Figure 2 shows the decoded pyramid of Lena (512×512, 8 bpp) with *four* spatial scales and *two* SNRs at full scale. Due to the coarse quantization of the lower resolution layers, provided by the optimum quantizer, and by the set of decision thresholds, the intermediate token images are rather *compact* in terms of rate, compared to their full-resolution versions.

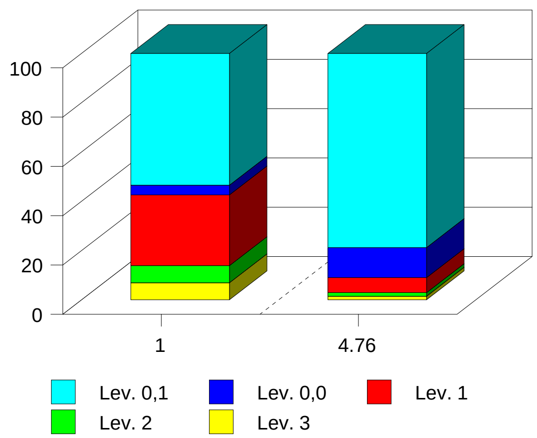

Figure 3: Percentages of progressive bit-rate associated to levels 3 (root), 2, 1, 0.0 (coarse full scale), and 0.1 (fine full- scale) of the multi-layer versions of *Lena* shown in Fig. 4, and of an analogous experiment yielding error-free reconstruction at level 0.1 with a rate of 4.76 bpp.

Figure 3 shows how the bit-rates produced in the test of Fig. 2 and in the lossless case were shared among the various layers. Note that the actual rates, as well as the PSNRs, of the lower scales are comparable in the two experiments: a property characteristic of the optimum quantizer developed in [1].

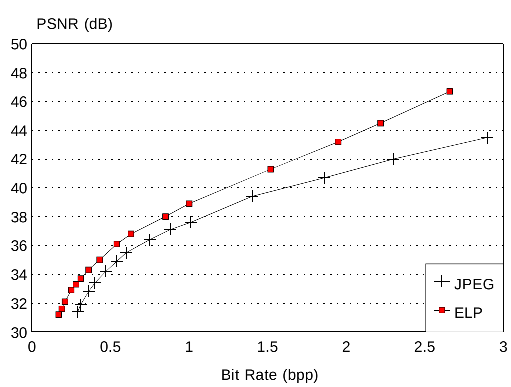

Figure 4: Performance plots of ELP and JPEG coding schemes for test image Lena: PSNR vs. Bit-Rate.

Performance curves are drawn in Figure 4 relating PSNR to Bit-Rate for *Lena* coded by means of the ELP scheme and JPEG. From the plots it appears that the former is more efficient than the latter, especially for very low rates, notwithstanding in the present implementation quantization is designed for minimizing the maximum absolute error. The PSNR gain over JPEG is about 2 dB for rates lower than 0.5 bpp, is superior by at least 1 dB in the middle region, and is steadily increasing for rates greater than 1.5 bpp.

References

[1] B. Aiazzi, L. Alparone, S. Baronti, and F. Lotti, "Lossless Image Compression by Quantizazione Feedback in a Content-Driven Enhanced Laplacian Pyramid", to appear on *IEEE Trans. Image Processing*.

[2] S. Baronti, A. Casini, F. Lotti, and L. Alparone, "Content-driven differential encoding of an enhanced image pyramid", *Signal Processing: Image Communication*, Vol. 6, No. 5, Oct. 1994, pp. 463–469.

[3] P. J. Burt, and E. H. Adelson, "The Laplacian Pyramid as a Compact Image Code", *IEEE Trans. Communications*, Vol. 31, No. 4, Apr. 1983, pp. 532–540.

[4] F. Lotti, B. Aiazzi, S. Baronti, A. Casini, and L. Alparone, "Content-Driven Pyramid Compression of Medical Images for Error-Free Archival and Progressive Transmission", *European Trans. Telecomm.*, Vol. 6, No. 3, May.-Jun. 1995.

[5] A. N. Netravali, and B. G. Haskell, *Digital Pictures*, Plenum Press, 1988.

[6] W. B. Pennebaker, and J. L. Mitchell, *JPEG Still Image Data Compression Standard*, Van Nostrand Reinhold, 1993.

[7] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit Allocation for Dependent Quantization with Applications to Multiresolution and MPEG Video Coders", *IEEE Trans. Image Processing*, Vol. 3, No. 5, pp. 533–545, Sep. 1994.

[8] K. H. Tzou, "Progressive image transmission: a review and comparison of techniques", *Optical Engineering*, Vol. 26, No. 7, Jul. 1987, pp. 581–589.

[9] K.M. Uz, M. Vetterli, and D.J. LeGall, "Interpolative Multiresolution Coding of Advanced Television with Compatible Subchannels", *IEEE Trans. Circuits Systems Video Technol.*, Vol. 1, No. 1, Mar. 1991, pp. 86–99.

[10] L. Wang, and M. Goldberg, "Comparative Performance of Pyramid Data Structures for Progressive Image Transmission", *IEEE Trans. Communications*, Vol. 39, No. 4, Apr. 1991, pp. 540–548.