A NON-EXPANSIVE PYRAMIDAL MORPHOLOGICAL IMAGE CODER

Dinei A. F. Florêncio and Ronald W. Schafer

Digital Signal Processing Laboratory
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332
floren@eedsp.gatech.edu rws@eedsp.gatech.edu

ABSTRACT

Pyramid Image Coding is a natural coding scheme for applications where progressive transmission is desired. In this kind of coder, versions of the original image at several resolution levels are formed by successive filtering and subsampling. Then, beginning from the coarsest image, the image is used to produce an estimate for the next (higher resolution) level and the error is coded and transmitted. While expansive pyramids (e.g., Burt's Laplacian Pyramid) are usually less efficient, non-expansive pyramids tend to produce ringing (e.g., Subband/Wavelet) and/or blocking (e.g., DCT). In this paper we introduce a non-expansive pyramid that does not produce ringing or blocking effects. Instead, the main artifact is texture removal. The simulations have produced images with entropies in the range of .2 to 1.5 bpp, with SNR figures similar to or better than JPEG at equivalent rates. The proposed coder has several atractive features, including 8 bit integer operations only, a perfect reconstruction mode, progressive transmission and an interesting progressive computation property.

1. INTRODUCTION

When transmitting images over slow channels, progressive image transmission may be of interest. Several algorithms allow for "progressive transmission modes" [1]. Among these, the pyramid coders are specially interesting, since the coding of the image at several different resolutions is the essence of the coder, with no extra processing being necessary to compute the intermediate images.

In a pyramid coder, versions of the original image at several resolution levels are formed and (progressively) transmitted. Figure 1 shows the basic structure of a 3-stage pyramid coder. In Burt's "Laplacian Pyramid" [2], the lower resolution images are produced from the original image by lowpass gaussian filtering (to reduce aliasing) and subsampling, and the reconstruction process uses a similar filter for interpolation. These linear filters have the undesired effect of blurring the edges of the images, and several authors have applied mathematical morphology to this pre-filter in order to improve edge preservation [3, 4, 5]. They have used morphological operations (openings, closing, or dilations) to avoid aliasing without blurring the edges, but their success has been limited by the use of non-critical sampling strategies, implying a higher entropy in the difference pyramids, and therefore reducing the efficiency of the coder. The efficiency in the lower resolution levels of the pyramid was so low that a scheme that does not involve subsampling has been proposed [6].

With the recently introduced *Critical Morphological Sampling Theorem* [7, 8], 100% better efficiency can be obtained in these lower resolution levels. In [8] we used critical morphological sampling to produce a pyramid with lower redundancy. Nevertheless, that pyramid has the disadvantage of expanding the number of points to be coded by approximately 33% (as in the Laplacian and many other pyramid structures). In this paper we propose a *non-expansive* version of that pyramidal structure. This has been acomplished by completely removing the anti-aliasing filter (and therefore guaranteeing that one out every four samples in each difference image will be zero). The final structure is therefore similar to a hierarchical DPCM structure[9].

When compared to other coding techniques, the proposed coder has several advantages, including 8 bit arithmetic (i.e., it involves only fixed range, integer operations), multiplier-free filters (a division by 2 may be required), a perfect reconstruction mode, progressive transmission, and progressive computation on both decoder and encoder. A SNR comparison with JPEG shows a slight advantage to the proposed coder. A subjective comparison of the coded images shows a clear advantage of the proposed coder on flat regions and sharp edges, with no ringing or blocking (both common in JPEG images). On the other hand, JPEG performs

This work was supported in part by the CNPq (Brazil), under contract 200.246/90-9 and the Joint Services Electronics Program under contract DAAH-04-93-G-0027.

0-8186-6950-0/94 $4.00 © 1994 IEEE

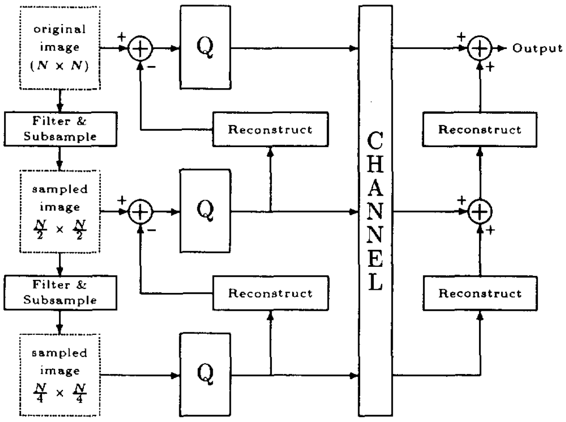

Figure 1: Basic structure of a (3-level) pyramid coder

better on regions characterized by textures.

Section 2 reviews the general structure of a pyramid coder. Section 3 discusses the reconstruction algorithm. Section 4 presents results of some simulations and Section 5 summarizes the conclusions.

2. CODER STRUCTURE

The structure of a generic pyramid coder/decoder is shown in Figure 1. It contains several similar stages (three stages in this figure). In each of these stages the signal is filtered, and then subsampled. The subsampled image is then used as the image for the next stage, where the same process is repeated. At each stage, an estimate for the higher resolution level is reconstructed, and the difference between this estimate and the true image at that level is quantized and transmitted. The same reconstruction process is used in the receiver in order to recover the image. The differences among pyramid coders are in the filtering, reconstruction and quantization processes. Burt's Laplacian Pyramid Coder [2] uses Gaussian-like filters for both anti-aliasing filtering and reconstruction. Several other schemes have been proposed based on Burt's scheme. While linear and morphological filters are both common as the anti-aliasing filters, the reconstruction process has been most often a linear process. In the quantization step both scalar and vectorial quantization are common. In [8] we introduced a pyramid coder where we used the recent Critical Morphological Sampling Theorem to help design effective anti-aliasing and reconstruction filters. In this paper we introduce a version of that coder where the anti-aliasing filters have been removed. Therefore the boxes labeled *"Filter & Subsample"* in Figure 1 consist only of a decimation process. The associated reconstruction process is designed taking into account the fact the aliasing may have occured.

3. THE RECONSTRUCTION PROCESS

The reconstruction process can be described as a single, fixed, WOS filter (Weighted Order Statistics filter), but a description as a space varying filter is more informative. Suppose we have an $N \times N$ image X, and want to estimate the $2N \times 2N$ version Y. The reconstruction process can then be described in the following way:

1. For samples where both coordinates are even make $Y(2i, 2j) = X(i, j)$.
2. For samples with first coordinate even, and second coordinate odd, i.e., $Y(2i, 2j + 1)$ compute a weighted median of $\{X(i-1, j), X(i-1, j+1), X(i, j), X(i, j+1), X(i+1, j), X(i+1, j+1)\}$ with weights $\{1, 1, 3, 3, 1, 1\}$.
3. For samples of the form $Y(2i + 1, 2j)$, compute the symmetric version of the previous step.
4. For samples of the form $Y(2i + 1, 2j + 1)$, compute the median of $\{X(i, j), X(i+1, j), X(i, j+1), X(i+1, j+1)\}$.

Note in particular that items 2 and 3, although using a mask that includes samples reasonably far from the sample being reconstructed, has been designed in such a way the the resulting value will always be between the values of the two closest pixels. An immediate consequence of this is that edge ringing is completely eliminated.

This reconstruction algorithm has been shown to be capable of producing images with very good subjective quality, and with the additional advantage of having a smaller reconstruction error than the linear reconstruction. The good performance of the filter can be attributed to the fact that it is capable of partially recovering the directional information contained in the subsampled signal, and the interpolated values indirectly take this into account. Note that this produces a kind of directional reconstruction without the need of estimating any directional information, as in [10].

Notice also that the reconstruction process does not involve any multiplications[1], and that all operations are 8-bit operations. In fact it can be shown that even the difference operation can be restricted to 8-bits by using arithmetic module-256. If no further quantization is used, this will atomatically produce an exact reconstruction, since no round-off errors are involved.

[1]a sum followed by division by two is used to solve the ties in the WM filter. A subtraction is necessary to compute the residual. All other operations are comparisons.

Figure 2: original 256x256 image

4. PERFORMANCE

To evaluate the performance of the coder, we implemented a 5-stage coder and compared with the equivalent Laplacian pyramid coder and with a JPEG coder, using a 256x256 version of the image Lenna (Figure 2). We have used the (first-order) entropy of the quantized signal as an estimate for the bit-rate of the coded image. Of course, to attain that rate, arithmetic coding should be used. When comparing rates with JPEG, we allow for a 5% margin to account for non-ideal coding.

The comparison with the Laplacian Pyramid coder can be done in a very direct way, because the basic structure of the two coders is similar. For comparison, we used the same quantization steps of 16, 8, 4, 2, and 1, in the difference images at levels 0, 1, 2, 3, and 4 respectively in both images, producing the results in Table 1. In that table, each line corresponds to one level of the pyramid, beginning from the coarser image (16x16). Note that the proposed scheme produces a result that is 1.12 dB better than the Laplacian Pyramid, and yet, the entropy of the error pyramid is smaller (1.43 bpp vs. 1.68). This result is 0.14 bpp better than the one obtained with the pyramid presented in [8]. As one would expect, the artifacts introduced by both coders are similar, and the subjective quality of the images match the 1 dB objective quality improvement.

The reconstruction process produces images with good subjective quality, even if no information about the difference image is transmitted at the highest resolution level. Based on this, we designed a more appropriate (scalar) quantizer for the proposed pyramid, and we were able to obtain very good results with even smaller (estimated) bit rates. The modified quantizer is essentially a uniform quantizer with a larger center bin, as the optimal quantizer for a Laplacian distribution [11]. Table 2 shows the bit rate and SNR rates for the proposed coder and JPEG. The bit rates correspond to true bit rates for the JPEG coder. For the proposed coder, we estimated the equivalent bit rate, based on the entropy of the difference signal plus a 5% margin to compensate for side information and the non-optimality of the entropy coders. Observe that the performance of the two coders on bit rates around .7 bpp is practically identical in terms of SNR. But the absence of blocking and ringing artifacts in the images produced by the proposed coder results in an image with a superior subjective quality. In fact, the JPEG coder introduces many artifacts in the image (blocking, ringing near the edges, false textures, etc), while the proposed coder introduces very few such artifacts, the main distortion being the removal of the texture in some regions. For higher rates, distortion is essentially negligible at 1.75 bpp for both coders. The proposed coder seems to have a much higher dynamic

Level (# pels)	Quant. step	Entropy (bpp)	Equiv. bpp	Total bpp
Proposed Pyramid (PSNR = 38.5 dB)				
4(192)	1	6.769	0.026	0.026
3(768)	2	5.725	0.067	0.093
2(3,072)	4	4.102	0.192	0.285
1(12,288)	8	2.417	0.453	0.738
0(49,152)	16	0.927	0.695	1.433
Laplacian Pyramid (PSNR = 37.4 dB)				
4(256)	1	6.682	0.026	0.026
3(1,024)	2	5.164	0.081	0.107
2(4,096)	4	3.572	0.223	0.330
1(16,384)	8	2.044	0.511	0.841
0(65,536)	16	0.845	0.845	1.686

Table 1: Entropy and SNR for a direct comparison between Laplacian and the proposed Morphological Pyramid

Proposed Method		**JPEG**	
Entropy + 5%	SNR	Bit Rate	SNR
1.75	40.38	1.75	38.72
1.00	35.69	1.00	34.72
0.85	34.29	0.85	33.74
0.70	33.19	0.70	32.64
0.55	31.53	0.55	31.46
0.20	26.95	0.20	26.42

Table 2: SNR comparison between JPEG and the proposed coder

range; it can be pushed up to perfect reconstruction (at entropy = 4.63 bpp) or down to rates around 0.15 bpp, where JPEG starts to break down and become very blocky. Corresponding images for JPEG and the proposed coder are shown in Figures 3 4 and 5, for bit rates of 1.75 bpp, 0.70 bpp and 0.15 bpp respectively. While the edge ringing and blocking artifacts of JPEG have proved difficult to overcome, it may be relatively easy to combine the proposed coder with some texture coding algorithm to improve the texture preservation of the proposed coder.

5. CONCLUSIONS

We have presented a new Non-Expansive Pyramid coder. The non-expansive characteristic has been obtained by removing the anti-aliasing filter, and using a non-linear reconstruction process that limits the effects of aliasing, while producing sharp reconstructed images. We showed that the coder compares favorably to a Laplacian Pyramid coder in relation to both SNR and bit rate. When compared to a JPEG coder, we observed that the proposed coder (using a similar bit rate) pro-

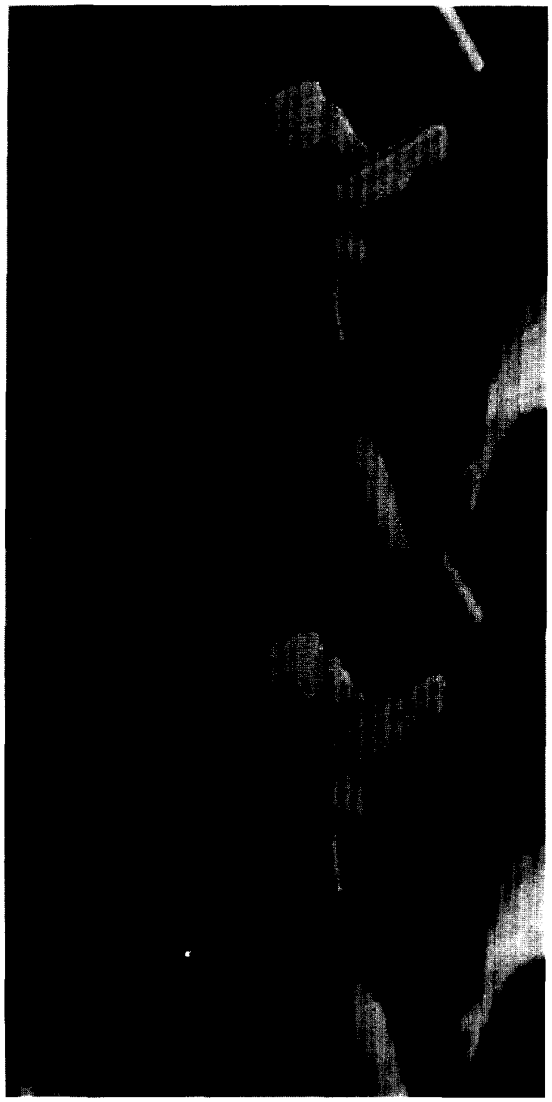

Figure 3: JPEG (top) and proposed coder (bottom) at 1.75 bpp.

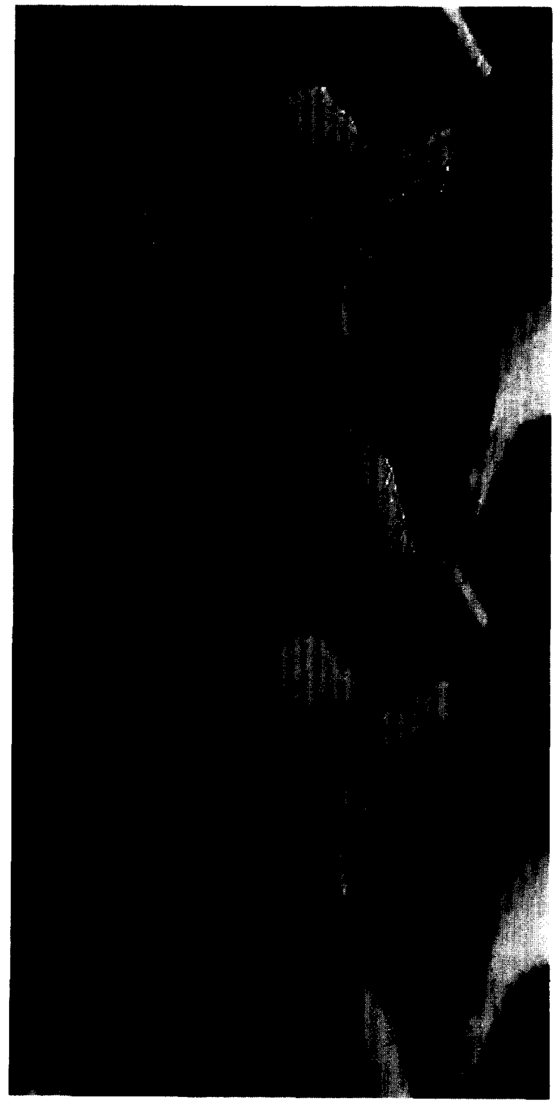

Figure 4: JPEG (top) and proposed coder (bottom) at 0.70 bpp

duces an image with very few artifacts, with a higher subjective quality, even though it has a similar SNR.

The coder has many attractive features for real-time implementations, including the fact that all operations can be performed with 8-bit precision (for a 8-bit image), low complexity (in particular, multiplier-free filters are used), high locality (the filters are within a 3x3 window), and progressive transmission. Additionally, it is interesting to note that the processing is done from bottom to top (i.e., from the lower resolution level to the higher resolution level). This — combined with progressive transmission — may be important in some applications where the available computing power is not known a *priori*, but a limited time is available for transmission (e.g., video coding on a CPU running several other process in parallel).

Future work will explore the redundancy remaining in the difference pyramid. This will include different subsampling geometries, vector quantization of the residuals, and improvement of the reconstruction process.

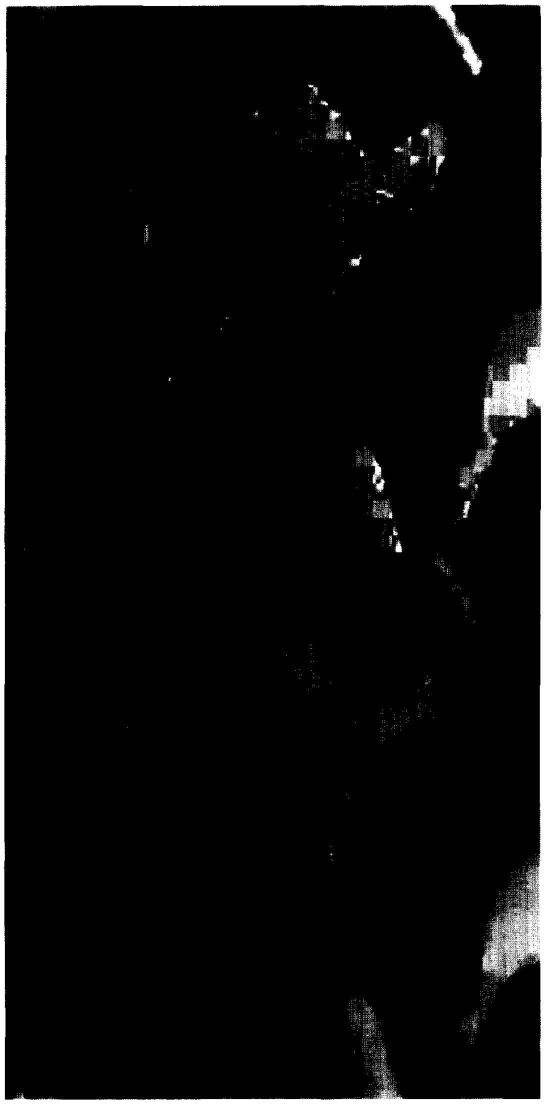

Figure 5: JPEG (top) and proposed coder (bottom) at 0.15 bpp

6. REFERENCES

[1] K.-H. Tzou, "Progressive image transmission: A review and comparison of techniques," *Optical Engineering*, vol. 26, pp. 581–589, July 1987.

[2] P. J. Burt, "The laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, pp. 532–540, Apr. 1983.

[3] L. A. Overturf, M. L. Comer, and E. J. Delp, "Color image coding using morphological pyramid decomposition," in *Human Vision, Visual Processing, and Digital Display III*, pp. 265–275, SPIE Vol. 1666, 1992.

[4] F.-K. Sun and P. Maragos, "Experiments on image compression using morphological pyramids," in *Visual Communication and Image Processing IV*, pp. 1303–1312, SPIE Vol. 1199, 1989.

[5] A. Toet, "A morphological pyramidal image decomposition," *Pattern Recog. Lett.*, vol. 9, pp. 255–261, May 1989.

[6] Z. Zhou and A. N. Venetsanopoulos, "Morphological methods in image coding," in *IEEE Intl. Conf. Acoust., Speech, Signal Processing*, pp. 481–484, Vol. III, 1992.

[7] D. A. F. Florêncio and R. W. Schafer, "Critical morphological sampling," to be submitted to IEEE Trans. on Signal Processing.

[8] D. A. F. Florêncio and R. W. Schafer, "Critical morphological sampling and applications to image coding," in *1994 ISMM Workshop on Mathematical Morphology and its Applications to Signal Processing*, Sept. 1994.

[9] R. L. de Queiroz and J. B. T. Yabu-uti, "Hierarchical image coding with pyramid DPCM," in *SBT/IEEE Intl. Telecomm. Symposium - ITS*, pp. 23.4.1–23.4.5, 1990.

[10] Y.-S. Ho and A. Gersho, "A pyramidal image coder with contour-based interpolative vector quantization," in *Visual Communication and Image Processing IV*, pp. 733–740, SPIE Vol. 1199, 1989.

[11] G. J. Sullivan, "Optimum entropy constrained scalar quatization for exponential and Laplacian random variables," in *IEEE Intl. Conf. Acoust., Speech, Signal Processing*, 1994.