Subband Image Coding with Three-tap Pyramids

Edward H. Adelson and Eero P. Simoncelli

MIT Media Laboratory
Cambridge, Massachusetts 02139

Subband coding is an effective means of data compression, but the quadrature mirror filters (QMF's) that are generally used have many taps, and often require many floating-point multiplications [1, 2]. As we have previously noted [3], it is possible to perform subband coding with extremely simple decoding filters if one is willing to encode the image with larger filters that are tailored for the task.

Consider the band-splitting filter pair [1 2 1] and [-1 2 -1]. These simple filters may be implemented with arithmetic shifts and additions, and thus are ideal for implementation on ordinary personal computers. In two dimensions they can be applied separably. The only problem is that the filters violate the standard QMF criteria, and therefore a different set of filters must be designed for the encoding process.

The image vector **e** may be written as a weighted sum of basis vectors corresponding to shifted versions of the filters, which appears as columns in the matrix **F**. If the weighting coefficients form a vector **p** we have:

$$\mathbf{e} = \mathbf{F}\mathbf{p}$$

$$\mathbf{F} = \begin{bmatrix} 1 & & & & \\ 2 & -1 & & & \\ 1 & 2 & 1 & & \\ & -1 & 2 & -1 & \\ & & 1 & 2 & \\ & & & -1 & \ddots \end{bmatrix}$$

For encoding, we seek a matrix **G** that will deliver the coefficients when applied to the image **e**. That is:

$$\mathbf{p} = \mathbf{G}^t\mathbf{e}$$

and thus

$$\mathbf{G} = (\mathbf{F}^{-1})^t.$$

The task then is simply to invert **F**; the columns of **G** will correspond to the encoding filters. These filters are generally non-zero over the entire length of the image, but they fall off rapidly and can be approximated by filters of finite length.

We have computed a set of optimal finite-length inverse filters with a simplex method, using an error criterion of maximum absolute value error on the reconstruction of a step

---

This research was supported by grants from NSF (IRI 871-939-4), Ford (SC-922465-ES), and DARPA (Rome Airforce F30602-89-C-0022).

edge. The optimized tap values are shown in table 1, for filter lengths of 15, 17, and 21. The accuracy improves as the filter length increases.

Figure 1 shows an original 256x256 8-bit image, and figure 2 shows the same image encoded at 0.8 bits per pixel, using a 4-level subband pyramid and a modified Huffman code. The performance is nearly as good as that obtained with QMF's.

We compared the decoding time of the 3-tap pyramid to a 9-tap QMF pyramid (with floating point coefficients) on a Sun 3/60 with a 20 Mhz 68020 CPU and 68881 floating-point coprocessor. The 3-tap pyramid was decoded in just under 1 second, and the 9-tap pyramid was slower by a factor of more than 20. Since the 3-tap pyramid is almost as effective as the larger QMF pyramids from the standpoint of compression, its simplicity and speed make it very attractive for many applications.

n	15	17	21
0	0.8648855700	0.8662753700	0.8660005000
1	0.3589060300	0.3588442800	0.3586960400
2	-0.1476441600	-0.1488108800	-0.1486006000
3	-0.0618851260	-0.0616580880	-0.0615359620
4	0.0244434030	0.0257062400	0.0255328510
5	0.0106931890	0.0102884290	0.0105768030
6	-0.0030558493	-0.0044906090	-0.0043832410
7	-0.0015278960	-0.0012884160	-0.0017810371
8		0.0006442405	0.0007449251
9			0.0002303323
10			-0.0001151661

Table 1: Filter impulse response values for 15, 17, and 21-tap inverses. Half of the impulse response sample values are shown for each of the normalized lowpass filters (All filters are symmetric about $n = 0$). The appropriate highpass filters are obtained by multiplying with the sequence $(-1)^n$ and shifting by one pixel.

References

[1] D. Esteban and C. Galand. Application of quadrature mirror filters to split band voice coding schemes. In *Proceedings ICASSP*, pages 191–195, 1977.

[2] J. D. Johnston. A filter family designed for use in quadrature mirror filter banks. In *Proceedings ICASSP*, volume 1, pages 291–294, 1980.

[3] Edward H. Adelson, Eero Simoncelli, and Rajesh Hingorani. Orthogonal pyramid transforms for image coding. In *Proceedings of SPIE*, volume 845, pages 50–58, Cambridge, MA, October 1987.

[4] Eero P. Simoncelli. Orthogonal sub-band image transforms. Master's thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, May 1988.

Figure 1: The original "Lena" image at 256 × 256 pixels, 8 bits per pixel.

Figure 2: Compressed image, using the 15-tap filter given in table 1. The compressed image occupied 6462 bytes or approximately 0.8 bits per pixel.