DOWNSCALED INVERSES FOR M-CHANNEL LAPPED TRANSFORMS

Ricardo L. de Queiroz

Xerox Corporation
800 Phillips Rd. M/S 128-27E
Webster, NY 14580
queiroz@wrc.xerox.com

Reiner Eschbach

Xerox Corporation
800 Phillips Rd. M/S 128-27E
Webster, NY 14580
Reiner_Eschbach@wb.xerox.com

ABSTRACT

Compressed images may be decompressed for devices using different resolutions. Full decompression and rescaling in space domain is a very expensive method. We studied downscaled inverses where the image is decompressed partially and a reduced inverse transform is used to recover the image. We studied the design of fast inverses, for a given forward transform. General solutions are presented for M-channel FIR filter banks of which block and lapped transforms are a subset.

1. INTRODUCTION

Suppose an image is scanned and compressed using an M-channel block or lapped transforms. It is then stored and printed or displayed on one of several devices, each supporting a distinct resolution. Thus, one must be able to decompress the image and resize it to fit the desired resolution. It is reasonable to assume that the image will be stored at a higher resolution and downsized to fit the lower resolution devices. It is also desirable to avoid buffering the full resolution image or to process more pixels than necessary. The alternative is to decompress the image directly into its lower resolution.

Transform coding is very popular for image compression. Downsampling by factors which are powers of 2 is trivial if we use the discrete wavelet transform (DWT) or other similar subband approaches [1]. However, we focus our attention on other attractive transforms which are called "parallel" or "M-channel" [2]. In these, the input image is directly decomposed into several subbands at once. Examples of these transforms are block transforms such as the discrete cosine and sine transforms (DCT,DST) [3], along with lapped transforms such as the lapped orthogonal transform (LOT) [4, 5], the generalized LOT (GenLOT) [6] and extended lapped transforms (ELT) [4, 7], as well as other M-channel filter banks [2].

2. BACKGROUND

2.1. Filter banks and polyphase matrices

We use FIR uniform filter banks, of which block and lapped transforms are special cases. There are M analysis filters $f_k(n)$ and M synthesis filters $g_k(n)$ $(0 \leq k \leq M-1)$. The signal $x(n)$ is decomposed into M subband signals $y_k(m)$ $(0 \leq k \leq M-1)$. After processing or quantization, the subband signals $\hat{y}_k(m)$ are used to construct the signal $\hat{x}(n)$. For simplicity, let $f_k(n)$ and $g_k(n)$ have $L = NM$ taps each, padding zeros if necessary.

It is more convenient to work with the polyphase transfer matrix (PTM) of the system which is a linear multi-input multi-output (MIMO) system of FIR filters relating M polyphase components of the input signal ($x_i(n) = x(nM+i)$) to the M subbands [2]. Conversion to and from polyphase components are called blocking and unblocking operations. A PTM is an LTI system iff it is pseudo circulant [2], otherwise (which is often the case) it represents a linear periodically time-varying (LPTV) filter.

The signal is blocked and passed through the analysis PTM $\mathbf{F}(z)$. It is reconstructed from the subbands using the PTM $\mathbf{G}^T(z)$ followed by an unblocking device. (The rows of $\mathbf{G}(z)$ correspond to the filters $g_k(n)$.) As $L = NM$, the PTMs have polynomial entries of order $N-1$. In practice, neither PTM represents an LTI system. However, for perfect reconstruction (PR) analysis-synthesis systems [2, 4] we have

$$\mathbf{T}(z) = \mathbf{G}^T(z)\mathbf{F}(z) = z^{-N+1}\mathbf{I}_M. \qquad (1)$$

and the overall system is an LTI filter (a pure delay).

2.2. Resampling as post-processing

The straightforward method to rescale the compressed image is to perform an inverse transform and, then, scale the image, as in Fig. 1.

A simplification can be achieved if we allow non-uniform resampling and assume $K < M$. For this, we employ the symbol $\uparrow K/M \downarrow$ in Fig. 1, which means: retain K out of M samples. More generically, one can retain nK out of nM samples. In Fig. 1 the filter $H(z)$ is a low-pass with cutoff on $K\pi/M$. The filtered signal $v(n)$ is resampled to the final sampling rate yielding $u(n)$. $H(z)$ can be moved past the unblocking device as a MIMO system as shown in Fig. 1. Let $\mathbf{x}(z)$ have entries $X_i(z) = \mathcal{Z}\{x_i(n)\}$. Hence,

$$\mathbf{v}(z) = \mathbf{H}(z)\mathbf{G}^T(z)\mathbf{F}(z)\mathbf{x}(z) = \mathbf{H}(z)\mathbf{T}(z)\mathbf{x}(z) \qquad (2)$$

and, for PR filter banks, $V(z) = z^{-L+1}H(z)X(z)$ and the overall system is LTI followed by a resampler. If M/K is an integer, the system becomes a trivial uniform downsampler, where $P = 1$ and $Q = M/K$, and the filter has cutoff at π/Q. The synthesis system and filters are effectively

0-8186-7919-0/97 $10.00 © 1997 IEEE

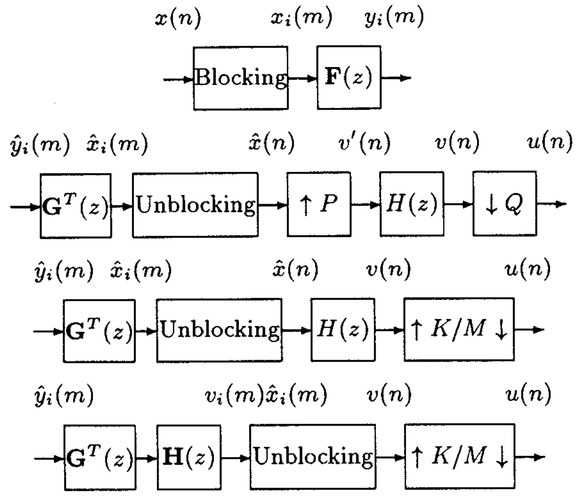

Figure 1: The analysis and synthesis systems with resampling as post-processing. If non-uniform downsampling is used, the filter can be moved past the unblocking device.

$$\mathbf{S}(z) = \mathbf{H}(z)\mathbf{G}^T(z). \tag{3}$$

$$s_k(n) = h(n) * g_k(n). \tag{4}$$

However, any fast algorithm for $\mathbf{G}(z)$ may be lost.

2.3. Discarding subbands

A computationally efficient way to implement the filter may borrow the filtering properties (stopband attenuation, etc.) of the filters $f_k(n)$ and $g_k(n)$. If the filters have good stopband attenuation, we can simply discard the highest frequency $M-K$ subbands and keep the lower K subbands. Then, we can either set $y_i(n) = 0$ for $K \le i \le M-1$, or set $g_i(n) = 0$ for $K \le i \le M-1$ and $0 \le n \le L-1$. Let

$$\mathbf{I}'_{K,M} = diag\{\underbrace{1, 1, \ldots, 1}_{K\ 1's}, \underbrace{0, 0, \ldots, 0}_{M-K\ 0's}\} \tag{5}$$

The equivalent synthesis system is given by

$$\mathbf{S}(z) = \mathbf{G}^T(z)\mathbf{I}'_{K,M}. \tag{6}$$

$$\mathbf{T}(z) = \mathbf{G}^T(z)\mathbf{I}'_{K,M}\mathbf{F}(z) \tag{7}$$

Unless severe restrictions are imposed to the filter bank design, $\mathbf{T}(z)$ will not represent an LTI system. However, if the aliasing terms are sufficiently small the overall transfer filter is approximately LTI and given by

$$T(z) \approx \frac{1}{M}\sum_{k=0}^{K-1} G_k(z)F_k(z). \tag{8}$$

So, "good" filters may approximate a reasonably selective filter after discarding subbands. We cannot compare LPTV and LTI filters but the above approximation is useful to give a reference point. Fig. 2 shows a comparison of approximated LPTV filters and LTI ones.

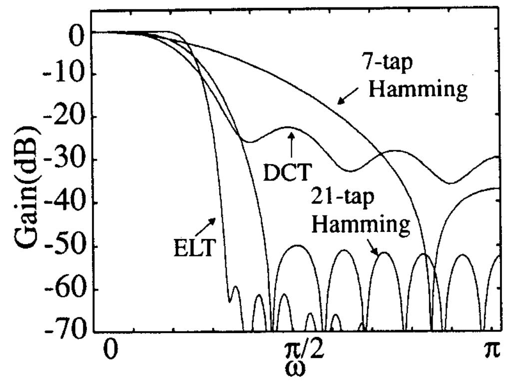

Figure 2: Frequency response plots for some useful filters for downsampling a signal at a 4:1 ratio.

3. BLOCK RESAMPLING

The goal in this paper is to directly design a synthesis system which would output K samples of $u(n)$ for every block of M subband samples ($y_i(m)$ for $0 \le i \le M-1$). In this case, we can design a synthesis PTM such that

$$\mathbf{S}(z) = \begin{bmatrix} S_{00}(z) & \cdots & S_{0,M-1}(z) \\ \vdots & \ddots & \vdots \\ S_{K-1,0}(z) & \cdots & S_{K-1,M-1}(z) \end{bmatrix}, \tag{9}$$

and $\mathbf{T}(z)$ becomes a $K \times M$ matrix given by

$$\mathbf{T}(z) = \mathbf{S}(z)\mathbf{F}(z). \tag{10}$$

$\mathbf{S}(z)$ implies filtering and resampling. We can design $\mathbf{S}(z)$ by block resampling of the output signal of a regular analysis-synthesis system and implement filtering by processing the subbands, i.e.

$$\mathbf{S}(z) = \mathbf{\Phi}(z)\mathbf{G}^T(z)\mathbf{C}(z), \tag{11}$$

where $\mathbf{\Phi}(z)$ is a $K \times M$ resampling matrix and $\mathbf{C}(z)$ is the filtering operator. For example, if $\mathbf{C}(z)$ is a diagonal matrix with zero order entries, it will perform filtering by weighting the subbands. Let

$$\mathbf{v}(z) = \mathbf{G}^T(z)\mathbf{C}(z)\mathbf{F}(z)\mathbf{x}(z) = \mathbf{A}(z)\mathbf{x}(z), \tag{12}$$

so that

$$\mathbf{u}(z) = \mathbf{\Phi}(z)\mathbf{v}(z). \tag{13}$$

The signal $v(n)$ has M samples per block, while the final signal $u(n)$ has only K samples per block. Thus $\mathbf{u}(z)$ has the K polyphases $U_i(z)$ and $\mathbf{v}(z)$ has M polyphases $V_i(z)$. We have chosen to construct a continuous curve from which $u(n)$ and $v(n)$ can be found by uniform sampling. This approach only works if the samples of $v(n)$ generate smooth curves, i.e. if the LPTV filter $\mathbf{A}(z)$ does a good job of removing high-frequency components. Linear interpolation (fitting a straight line in between every two samples of $u(n)$) is generally visually pleasing. Splines and higher order polynomials may be used as well. The sampling grid we

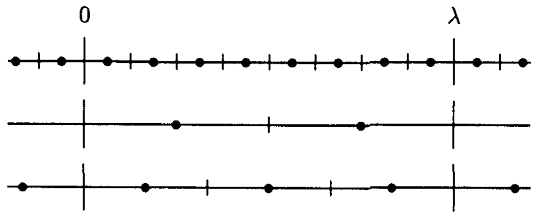

Figure 3: The interval of a block originally with 8 samples, is resampled at 2 and 3 samples per block.

have chosen is illustrated in Fig. 3, where sampling occurs at the center of K uniform cells that fill the block interval. It is straightforward to design $\mathbf{\Phi}(z)$ to allow the proposed resampling. For example, with a piecewise linear interpolation between samples, $M = 8$ and $K = 2$, we have

$$\mathbf{\Phi}(z) = \begin{bmatrix} 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 \end{bmatrix}$$

For the filtering operator, we can further simplify the problem by using the approach in Sec. 2.3 ($\mathbf{C}(z) = \mathbf{I}'_{K,M}$). Then, filtering is accomplished by retaining only the K lower frequency subbands. Let $\bar{\mathbf{y}}(z)$ be the vector with the K lowest frequency subbands. We want to design a PTM, which we denote by $\mathbf{R}(z)$, directly relating these K subbands to the K output polyphase components as:

$$\mathbf{u}(z) = \mathbf{R}^T(z)\bar{\mathbf{y}}(z). \tag{14}$$

Let $\mathbf{\Lambda} = [\mathbf{I}_K, \mathbf{0}_{K\times M-K}]$ so that $\bar{\mathbf{G}}(z) = \mathbf{\Lambda}\mathbf{G}(z)$. Hence,

$$\mathbf{u}(z) = \mathbf{R}^T(z)\mathbf{\Lambda}\hat{\mathbf{y}}(z) = \mathbf{R}^T(z)\mathbf{\Lambda}\mathbf{F}(z)\mathbf{x}(z) \tag{15}$$

$$\mathbf{S}(z) = \mathbf{R}^T(z)\mathbf{\Lambda} \tag{16}$$

As $\mathbf{C}(z) = \mathbf{I}'_{K,M} = \mathbf{\Lambda}^T\mathbf{\Lambda}$, we have that

$$\mathbf{S}(z) = \mathbf{\Phi}(z)\mathbf{G}^T(z)\mathbf{\Lambda}^T\mathbf{\Lambda} \tag{17}$$

$$\mathbf{R}^T(z) = \mathbf{\Phi}(z)\mathbf{G}^T(z)\mathbf{\Lambda}^T = \mathbf{\Phi}(z)\bar{\mathbf{G}}^T(z). \tag{18}$$

This result implies that the K synthesis filters should be resampled versions of the K filters corresponding to the K lowest-frequency subbands of the original synthesis filter bank $\mathbf{G}(z)$. Thus, the actual synthesis filters $r_k(n)$ have length NK and are found as

$$\mathbf{r}_k(z) = \mathbf{\Phi}(z)\mathbf{g}_k(z). \tag{19}$$

We can define a continuous function for the filters $\mu_k(t)$, for $0 \le k \le K-1$, which are uniformly sampled. Thus,

$$r_k(n) = \mu_k\left(\frac{2n+1}{2K}\lambda\right) \tag{20}$$

for λ representing the normalized support of M samples of the impulse response of $g_k(n)$. See [8] for more in the subject including upsampling and more general factors.

4. FAST TRANSFORMS

For fast transforms, we design $\mathbf{R}(z)$ possessing a fast algorithm to approximate given $\mathbf{G}(z)$ and $\mathbf{\Phi}(z)$. Table 1 shows the filters response for DCT and DST of types II (the common one) and IV [3]. It also gives a desirable continuous function that will interpolate the original filter samples, from which the new filters $r_k(n)$ are found. In this case, if the listed block transforms have their $k \times k$ transform matrix denoted as $\mathbf{D}_k$, the forward transform employs $\mathbf{D}_M$ while the synthesis filters are found as

$$\mathbf{R}^T(z) = \sqrt{\frac{K}{M}}\,\mathbf{D}_K^T. \tag{21}$$

Resizing in the DCT domain has been studied before [9]. Similar resampling methods for the DCT have also been recently reported [10, 11].The table also includes results for a cosine modulated filter bank known as the ELT [4]. If $\mathbf{E}_k(z)$ is the analysis PTM for the k-channels ELT, the downscaled synthesis is found as:

$$\mathbf{R}^T(z) = \sqrt{\frac{K}{M}}\,\mathbf{E}_K^T(z). \tag{22}$$

Note that for the ELT, one may have to find an appropriate continuous modulation prototype (window) $w(t)$. We recommend interpolating $w^{(M)}(n)$ (M-channels case) into $w(t)$ with a smooth function such that $w^{(K)}(n)$ (K-channels case) can be found by $\mathbf{w}_k^{(K)}(z) = \mathbf{\Phi}(z)\mathbf{w}_k^{(M)}(z)$.

The same concepts also apply to other filter banks and lapped transforms. The LOT and GenLOTs are filter banks (lapped transforms) whose filters have linear-phase (symmetric bases) [4, 5, 6]. The LOT can be viewed as a special case of a GenLOT [6]. In this case, if $\mathbf{L}(z)$ is the K-channel synthesis GenLOT (with proper reoptimization of its parameters for given M-channel analysis GenLOT [8]) we have

$$\mathbf{R}^T(z) = \sqrt{\frac{K}{M}}\,\mathbf{L}^T(z). \tag{23}$$

We also carried tests, by compressing an image at 1 bit/pel using JPEG (DCT) and decompressing it at a quarter resolution ($M = 8, K = 2$). The proposed method takes 12 operations (multiplies plus adds) to construct a block of 2×2 pixels from the block with the 8×8 DCT coefficients. We compared it to space-domain subsampling techniques including: no postfilter (672 ops); 4×4 averaging filter (736 ops); 21×21 Hamming filter (4196 ops). Portions of the decompressed images are shown in Fig. 4. We also show in Fig. 5 the reconstructed image using the ELT in place of the DCT in JPEG, with and without mismatch in the window design, in order to highlight its importance.

5. CONCLUSIONS

We studied methods to decompress a compressed image to a lower resolution. A general framework was presented allowing the use of an arbitrary FIR uniform paraunitary filter bank along with specific algorithms aimed at popular block and lapped transforms. As results have shown, the solutions we propose yield higher quality of the decompressed

Table 1: Continuous functions for popular block and lapped transforms

Transform	filter elements $g_k(n)$	continuous function $\mu_k(t)$
DCT	$\sqrt{\frac{2}{M}}\alpha_k \cos((2n+1)k\pi/2M)$	$\alpha_k\sqrt{\frac{2}{M}}\cos(tk\pi)$
DST	$\sqrt{\frac{2}{M}}\alpha_k \sin((2n+1)k\pi/2M)$	$\alpha_k\sqrt{\frac{2}{M}}\sin(tk\pi)$
DST-IV	$\sqrt{\frac{2}{M}}\sin((2n+1)(2n+1)\pi/4M)$	$\sqrt{\frac{2}{M}}\sin(t(2k+1)\pi/2M)$
DCT-IV	$\sqrt{\frac{2}{M}}\cos((2n+1)(2n+1)\pi/4M)$	$\sqrt{\frac{2}{M}}\cos(t(2k+1)\pi/2M)$
ELT	$w(n)\sqrt{\frac{2}{M}}\cos\left[\left(k+\frac{1}{2}\right)\frac{\pi}{M}\left(n+\frac{M+1}{2}\right)\right]$	$w(t)\sqrt{\frac{2}{M}}\cos\left[\left(k+\frac{1}{2}\right)\left(t+\frac{\pi}{2}\right)\right]$

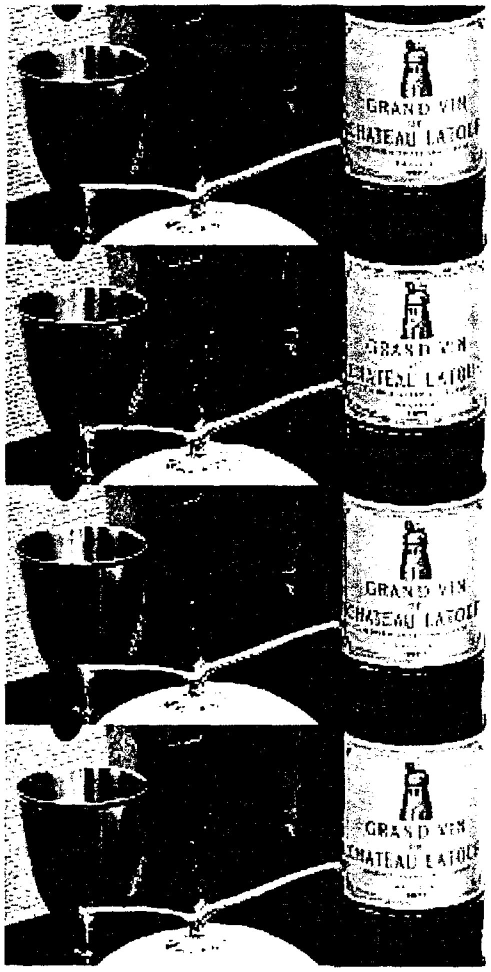

Figure 4: Reconstructed JPEG images at a quarter resolution. From top to bottom: proposed method; no filter; 4×4 averaging; and 21×21 Hamming filter.

image compared to other simple reconstruction methods. Furthermore, it also leads to much faster implementation. The basic idea is to resample the synthesis filters instead of resampling the image.

6. REFERENCES

[1] M. Vetterli and J. Kovacevic. *Wavelets and Subband Coding*. Prentice-Hall. 1995.

[2] P.P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice-Hall. 1993.

[3] K. Rao and P. Yip *Discrete Cosine Transform : Algorithms, Advantages, Applications*. Academic Press. 1990.

[4] H. S. Malvar. *Signal Processing with Lapped Transforms*. Artech House. 1992.

[5] H. S. Malvar and D. H. Staelin. "The LOT: transform coding without blocking effects." *IEEE Trans. ASSP*. ASSP-37. pp. 553–559. Apr. 1989.

[6] R. L. de Queiroz. T. Q. Nguyen. and K. R. Rao. "GenLOT: generalized linear-phase lapped orthogonal transforms." *IEEE Trans. Signal Processing*. Vol. 44. pp. 497–507. Apr. 1996.

[7] H. S. Malvar. "Extended lapped transforms: properties. applications and fast algorithms." *IEEE Trans. Signal Processing*. vol. 40. pp. 2703–2714. Nov. 1992.

[8] R. de Queiroz and R. Eschbach. "Fast downscaled inverses for images compressed with M-channel lapped transforms." *IEEE Trans. Image Proc.*. 1997.

[9] K. N. Ngan. "Experiments on 2D decimation in time and orthogonal transform domains." *Signal Processing*. Vol. 11. pp. 249–263. Oct. 1986.

[10] S. Martucci. "Image resizing in the DCT domain." *Proc. ICIP*. Vol. II. pp. 244-247. 1995.

[11] B. Natarajan and B. Vasudev. "A fast approximate algorithm for scaling down digital images in the DCT domain." *Proc. ICIP*. Vol. II. pp. 241-243. 1995.

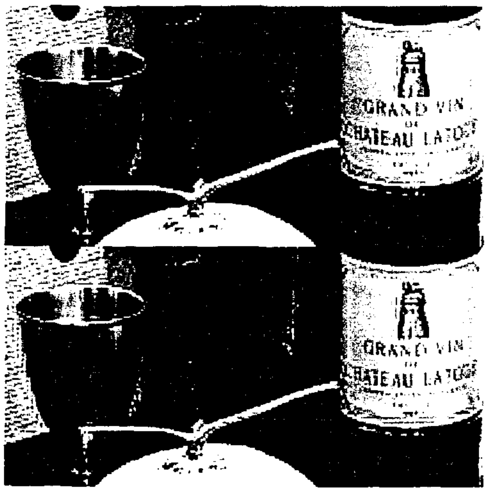

Figure 5: Same experiment for the ELT (top) and for an ELT designed with window mismatch (bottom).