LAPPED TRANSFORMS

Ricardo L. de Queiroz
Advanced Color Imaging
Xerox Corporation
queiroz@wrc.xerox.com

0.1 Introduction

The idea of a lapped transform (LT, for short) maintaining orthogonality and non-expansion of the samples was developed in the early 80's at MIT by a group of researchers unhappy with the blocking artifacts so common in traditional block transform coding of images. The idea was to extend the basis function beyond the block boundaries, creating an overlap, in order to eliminate the blocking effect. This idea was not new, but the new ingredient to overlapping blocks would be the fact that the number of transform coefficients would be the same as if there was no overlap, and that the transform would mantain orthogonality. Cassereau [1] introduced the lapped orthogonal transform (LOT), and Malvar [5],[6],[7] gave the LOT its design strategy and a fast algorithm. It was later pointed by Malvar [9] the equivalence between an LOT and a multirate filter bank. Based on cosine modulated filter banks [15], modulated lapped transforms were designed [8],[25]. Modulated transforms were generalized for an arbitrary overlap later, creating the class of extended lapped transforms (ELT) [10]–[13]. Recently a new class of LTs with symmetric bases were developed yielding the class of generalized LOTs (GenLOT) [17],[19],[20]. As we mentioned, filter banks and LTs are the same, although studied independently in the past. We, however, refer to LTs for paraunitary uniform FIR filter banks with fast implementation algorihtms based on special factorizations od the basis functions.

We assume a one-dimensional input sequence $x(n)$ which is transformed into several coefficients $y_i(n)$, where $y_i(n)$ would belong to the i-th subband. We also will use the discrete cosine transform [23] and another cosine transform variation, which we abbreviate as DCT and DCT-IV (DCT type 4), respectively [23].

0.2 Orthogonal block transforms

In traditional block-transform processing, such as in image and audio coding, the signal is divided into blocks of M samples, and each block is processed independently [2], [3], [12], [14], [22], [23], [24]. Let the samples in the m-th block be denoted as

$$\mathbf{x}_m^T = [x_0(m), x_1(m), \ldots, x_{M-1}(m)], \tag{1}$$

for $x_k(m) = x(mM + k)$ and let the corresponding transform vector be

$$\mathbf{y}_m^T = [y_0(m), y_1(m), \ldots, y_{M-1}(m)]. \tag{2}$$

For a real unitary transform $\mathbf{A}$, $\mathbf{A}^T = \mathbf{A}^{-1}$. The forward and inverse transforms for the m-th block are

$$\mathbf{y}_m = \mathbf{A}\mathbf{x}_m, \tag{3}$$

and

$$\mathbf{x}_m = \mathbf{A}^T \mathbf{y}_m. \tag{4}$$

The rows of $\mathbf{A}$, denoted $\mathbf{a}_n^T$ $(0 \leq n \leq M-1)$, are called the basis vectors because they form an orthogonal basis for the M-tuples over the real field [24]. The transform vector coefficients $[y_0(m), y_1(m), \ldots, y_{M-1}(m)]$ represent the corresponding weights of vector $\mathbf{x}_m$ with respect to this basis.

If the input signal is represented by vector $\mathbf{x}$ while the subbands are grouped into blocks in vector $\mathbf{y}$, we can represent the transform $\mathbf{T}$ which operates over the entire signal as a block diagonal matrix:

$$\mathbf{T} = diag\{\ldots, \mathbf{A}, \mathbf{A}, \mathbf{A}, \ldots\}, \tag{5}$$

where, of course, $\mathbf{T}$ is an orthogonal matrix.

0.2.1 Orthogonal lapped transforms

For lapped transforms [12], the basis vectors can have length L, such that $L > M$, extending across traditional block boundaries. Thus, the transform matrix is no longer square and most of the equations valid for block transforms do not apply to an LT. We will concentrate our efforts on *orthogonal* LTs[12] and consider $L = NM$, where N is the overlap factor. Note that N, M, and hence L are all integers. As in the case of block transforms, we define the transform matrix as containing the orthonormal basis vectors as its rows. A lapped transform matrix $\mathbf{P}$ of dimensions $M \times L$ can be divided into square $M \times M$ submatrices $\mathbf{P}_i$ $(i = 0, 1, \ldots, N-1)$ as

$$\mathbf{P} = [\mathbf{P}_0 \; \mathbf{P}_1 \; \cdots \; \mathbf{P}_{N-1}]. \tag{6}$$

The orthogonality property does not hold because $\mathbf{P}$ is no longer a square matrix and it is replaced by other properties which we will discuss later.

If we divide the signal into blocks, each of size M, we would have vectors $\mathbf{x}_m$ and $\mathbf{y}_m$ such as in (1) and (2). These blocks are not used by LTs in a straightforward manner. The actual vector which is transformed by the matrix $\mathbf{P}$ has to have L samples and, at block number m, it is composed of the samples of $\mathbf{x}_m$ plus $L - M$ samples. These samples are chosen by picking $(L-M)/2$ samples at each side of the block $\mathbf{x}_m$, as shown in Fig. 1, for $N = 2$. However, the number of transform coefficients at each step is M, and, in this respect, there is no change in the way we represent the transform-domain blocks $\mathbf{y}_m$.

The input vector of length L is denoted as $\mathbf{v}_m$, which is centered around the block $\mathbf{x}_m$, and is defined as

$$\mathbf{v}_m^T = \left[x\left(mM - (N-1)\frac{M}{2}\right) \cdots x\left(mM + (N+1)\frac{M}{2} - 1\right)\right]. \tag{7}$$

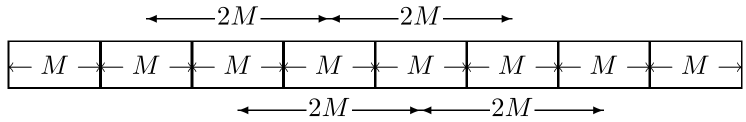

Figure 1: The signal samples are divided into blocks of M samples. The lapped transform uses neighboring blocks samples, as in this example for $N = 2$, i.e. $L = 2M$, yielding an overlap of $(L - M)/2 = M/2$ samples on either side of a block.

Then, we have

$$\mathbf{y}_m = \mathbf{P}\mathbf{v}_m. \tag{8}$$

The inverse transform is not direct as in the case of block transforms, i.e., with the knowledge of $\mathbf{y}_m$ we do not know the samples in the support region of $\mathbf{v}_m$, and neither in the support region of $\mathbf{x}_m$. We can reconstruct a vector $\hat{\mathbf{v}}_m$ from $\mathbf{y}_m$, as

$$\hat{\mathbf{v}}_m = \mathbf{P}^T\mathbf{y}_m. \tag{9}$$

where $\hat{\mathbf{v}}_m \neq \mathbf{v}_m$. To reconstruct the original sequence, it is necessary to accumulate the results of the vectors $\hat{\mathbf{v}}_m$, in a sense that a particular sample $x(n)$ will be reconstructed from the sum of the contributions it receives from all $\hat{\mathbf{v}}_m$, such that $x(n)$ was included in the region of support of the corresponding $\mathbf{v}_m$. This additional complication comes from the fact that $\mathbf{P}$ is not a square matrix [12]. However, the whole analysis-synthesis system (applied to the entire input vector) is orthogonal, assuring the PR property using (9).

We can also describe the process using a sliding rectangular window applied over the samples of $x(n)$. As an M-sample block $\mathbf{y}_m$ is computed using $\mathbf{v}_m$, $\mathbf{y}_{m+1}$ is computed from $\mathbf{v}_{m+1}$ which is obtained by shifting the window to the right by M samples, as shown in Fig. 2.

As the reader may have noticed, the region of support of all vectors $\mathbf{v}_m$ is greater than the region of support of the input vector. Hence, a special treatment has to be given to the transform at the borders. We will discuss this fact later and assume infinite-length signals until then, or assume the length is very large and the borders of the signal are far enough from the region to which we are focusing our attention.

If we denote by $\mathbf{x}$ the input vector and by $\mathbf{y}$ the transform-domain vector, we can be consistent with our notation of transform matrices by defining a matrix $\mathbf{T}$ such that $\mathbf{y} = \mathbf{T}\mathbf{x}$ and $\hat{\mathbf{x}} = \mathbf{T}^T\mathbf{y}$. In this case, we have

$$\mathbf{T} = \begin{bmatrix} \ddots & & & & \\ & \mathbf{P} & & & \\ & & \mathbf{P} & & \\ & & & \mathbf{P} & \\ & & & & \ddots \end{bmatrix}. \tag{10}$$

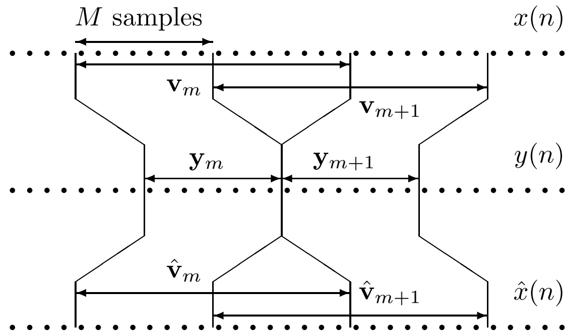

Figure 2: Illustration of a lapped transform with $N = 2$ applied to signal $x(n)$, yielding transform domain signal $y(n)$. The input L-tuple as vector $\mathbf{v}_m$ is obtained by a sliding window advancing M samples, generating $\mathbf{y}_m$. This sliding is also valid for the synthesis side.

---

where the displacement of the matrices $\mathbf{P}$ obeys the following

$$\mathbf{T} = \begin{bmatrix} \ddots & \ddots & & \ddots & & & \\ & \mathbf{P}_0 & \mathbf{P}_1 & \cdots & \mathbf{P}_{N-1} & & \\ & & \mathbf{P}_0 & \mathbf{P}_1 & \cdots & \mathbf{P}_{N-1} & \\ & & & \ddots & \ddots & & \ddots \end{bmatrix}. \tag{11}$$

$\mathbf{T}$ has as many block-rows as transform operations over each vector $\mathbf{v}_m$.

Let the rows of $\mathbf{P}$ be denoted by $1 \times L$ vectors $\mathbf{p}_i^T$ $(0 \le i \le M-1)$, so that $\mathbf{P}^T = [\mathbf{p}_0, \cdots, \mathbf{p}_{M-1}]$. In an analogy to the block transform case, we have

$$y_i(m) = \mathbf{p}_i^T \mathbf{v}_m. \tag{12}$$

The vectors $\mathbf{p}_i$ are the basis vectors of the lapped transform. They form an orthogonal basis for an M-dimensional subspace (there are only M vectors) of the L-tuples over the real field.

Assuming that the entire input and output signals are represented by the vectors $\mathbf{x}$ and $\mathbf{y}$, respectively, and that the signals have infinite length, then, from (10), we have

$$\mathbf{y} = \mathbf{T}\mathbf{x} \tag{13}$$

and, if $\mathbf{T}$ is orthogonal,

$$\mathbf{x} = \mathbf{T}^T \mathbf{y}. \tag{14}$$

The conditions for orthogonality of the LT is expressed as the orthogonality of $\mathbf{T}$. Therefore, the following equations are equivalent in a sense that they state the PR property along with the orthogonality of the LT.

$$\sum_{i=0}^{N-1-l} \mathbf{P}_i \mathbf{P}_{i+l}^T = \sum_{i=0}^{N-1-l} \mathbf{P}_i^T \mathbf{P}_{i+l} = \delta(l) \mathbf{I}_M. \tag{15}$$

$$\mathbf{T}\mathbf{T}^T = \mathbf{T}^T\mathbf{T} = \mathbf{I}_\infty \tag{16}$$

It is worthwhile to reaffirm that orthogonal LTs are uniform maximally-decimated FIR filter bank. Assume the filters in such filter bank have L-tap impulse responses $f_i(n)$ and $g_i(n)$ ($0 \leq i \leq M-1, 0 \leq n \leq L-1$), for the analysis and synthesis filters, respectively. If the filters have originally a length smaller than L one can pad the impulse response with 0's until $L = NM$. In other words, we force the basis vectors to have a common length which is an integer multiple of the block size. Assume the entries of $\mathbf{P}$ are denoted by $\{p_{ij}\}$. One can translate the notation from LTs to filter banks by using

$$p_{kn} = f_k(L-1-n) = g_k(n) \tag{17}$$

0.3 Useful transforms

0.3.1 Extended lapped transform (ELT)

Cosine modulated filter banks are filter banks based on a low-pass prototype filter modulating a cosine sequence. By a proper choice of the phase of the cosine sequence, Malvar developed the modulated lapped transform (MLT) [8], which led to the so-called extended lapped transforms (ELT) [10], [11], [12], [13]. The ELT allows several overlapping factors N, generating a family of LTs with good filters' frequency response and fast implementation algorithm.

In the ELTs, the filters' length L is basically an even multiple of the block size M, as $L = NM = 2kM$. The MLT-ELT class is defined by

$$p_{k,n} = h(n)\cos\left[\left(k+\frac{1}{2}\right)\left(\left(n-\frac{L-1}{2}\right)\frac{\pi}{M} + (N+1)\frac{\pi}{2}\right)\right] \tag{18}$$

for $k = 0, 1\ldots, M-1$ and $n = 0, 1, \ldots, L-1$. $h(n)$ is a symmetric window modulating the cosine sequence and the impulse response of a low-pass prototype (with cutoff frequency at $\pi/2M$) which is translated in frequency to M different frequency slots in order to construct the uniform filter bank. The ELTs have as their major plus a fast implementation algorithm, which is depicted in Fig. 3 in an example for $M = 8$. The free parameters in the design of an ELT are the coefficients of the prototype filter. Such degrees of freedom are translated in the fast algorithm as rotation angles.

For the case $N = 4$ there is a useful parameterized design [13][12][11]. In this design, we have:

$$\theta_{k0} = -\frac{\pi}{2} + \mu_{M/2+k} \tag{19}$$

$$\theta_{k1} = -\frac{\pi}{2} + \mu_{M/2-1-k} \tag{20}$$

where

$$\mu_i = \left[\left(\frac{1-\gamma}{2M}\right)(2k+1) + \gamma\right] \tag{21}$$

and γ is a control parameter, for $0 \leq k \leq (M/2)-1$. γ controls the trade-off between of the atenuation and transition region of the prototype filter. For $N = 4$, the relation between angles and $h(n)$ is:

$$h(k) = \cos(\theta_{k0})\cos(\theta_{k1}) \tag{22}$$

$$h(M-1-k) = \cos(\theta_{k0})\sin(\theta_{k1}) \tag{23}$$

$$h(M+k) = \sin(\theta_{k0})\cos(\theta_{k1}) \quad (24)$$
$$h(2M-1-k) = -\sin(\theta_{k0})\sin(\theta_{k1}) \quad (25)$$

for $k = 0, 1, \ldots, M/2 - 1$. See [12] for optimized angles for ELTs. Further details on ELTs can be found in [10], [11], [12], [13], [17].

0.3.2 Generalized linear-phase lapped orthogonal transform (GenLOT)

The generalized linear-phase lapped orthogonal transform (GenLOT) is also a useful family of LTs possessing symmetric bases (linear-phase filters). The use of linear-phase filters is a popular requirement in image processing applications. Let

$$\mathbf{W} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_{M/2} & \mathbf{I}_{M/2} \\ \mathbf{I}_{M/2} & -\mathbf{I}_{M/2} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Psi}_i = \begin{bmatrix} \mathbf{U}_i & \mathbf{0}_{M/2} \\ \mathbf{0}_{M/2} & \mathbf{V}_i \end{bmatrix}, \quad (26)$$

where $\mathbf{U}_i$ and $\mathbf{V}_i$ can be any $M/2 \times M/2$ orthogonal matrices. Let the transform matrix $\mathbf{P}$ for the GenLOT be constructed iteractively. Let $\mathbf{P}^{(i)}$ be the partial reconstruction of $\mathbf{P}$ after including up to the i-th stage. We start by setting $\mathbf{P}^{(0)} = \mathbf{E}_0$ where $\mathbf{E}_0$ is an orthogonal matrix with symmetric rows. The recursion is given by:

$$\mathbf{P}^{(i)} = \boldsymbol{\Psi}_i \mathbf{W} \mathbf{Z} \begin{bmatrix} \mathbf{W}\mathbf{P}^{(i-1)} & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{W}\mathbf{P}^{(i-1)} \end{bmatrix} \quad (27)$$

where

$$\mathbf{Z} = \begin{bmatrix} \mathbf{0}_{M/2} & \mathbf{0}_{M/2} & \mathbf{I}_{M/2} & \mathbf{0}_{M/2} \\ \mathbf{0}_{M/2} & \mathbf{I}_{M/2} & \mathbf{0}_{M/2} & \mathbf{0}_{M/2} \end{bmatrix}. \quad (28)$$

At the final stage we set $\mathbf{P} = \mathbf{P}^{(N-1)}$. $\mathbf{E}_0$ is usually the DCT while the other factors ($\mathbf{U}_i$ and $\mathbf{V}_i$) are found through optimization routines. More details on GenLOTs and its design can be found in [17], [19], [20]. The implementation flow-graph of a GenLOT with $M = 8$ is shown in Fig. 4.

0.4 Remarks

We hope this introductory work could be helpful to understand the basic concepts of lapped transforms. Filter banks are covered in other parts of this book. An excellent book by Vaidyanathan [28] has a thorough coverage of such subject. The interrelations of filter banks and LTs are well covered by Malvar [12] and Queiroz [17]. For image processing and coding, it is necessary to process finite-length signals. As we discussed, such issue is not so straightforward in a general case. Algorithms to implement LTs over finite-length signals are discussed in [7],[12],[16],[17],[18], [21]. These algorithms can be general or specific. The specific algorithms are generally targeted to a particular LT invariantly seeking a very fast implementation. In general, Malvar's book [12] is an excellent reference for lapped transforms and its related topics.

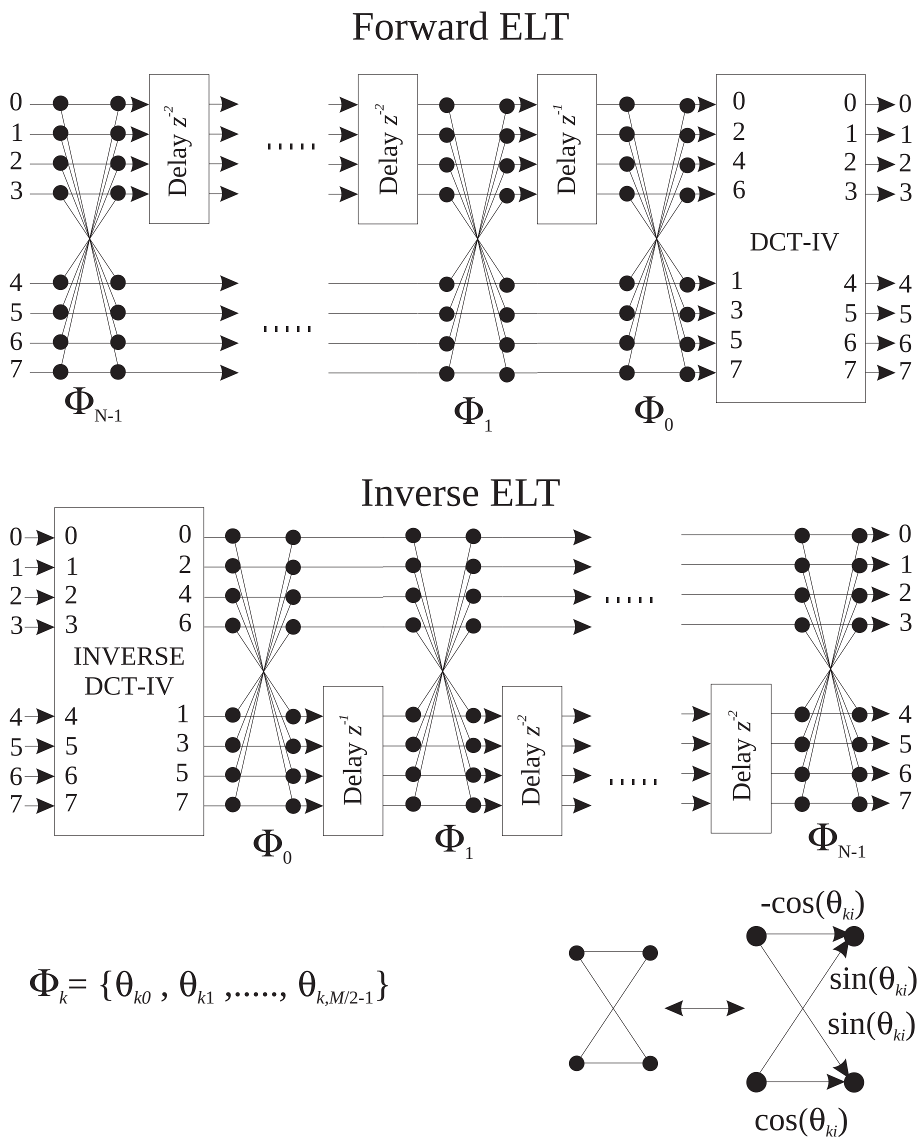

Figure 3: Implementation flow-graph for the ELT with $M = 8$.

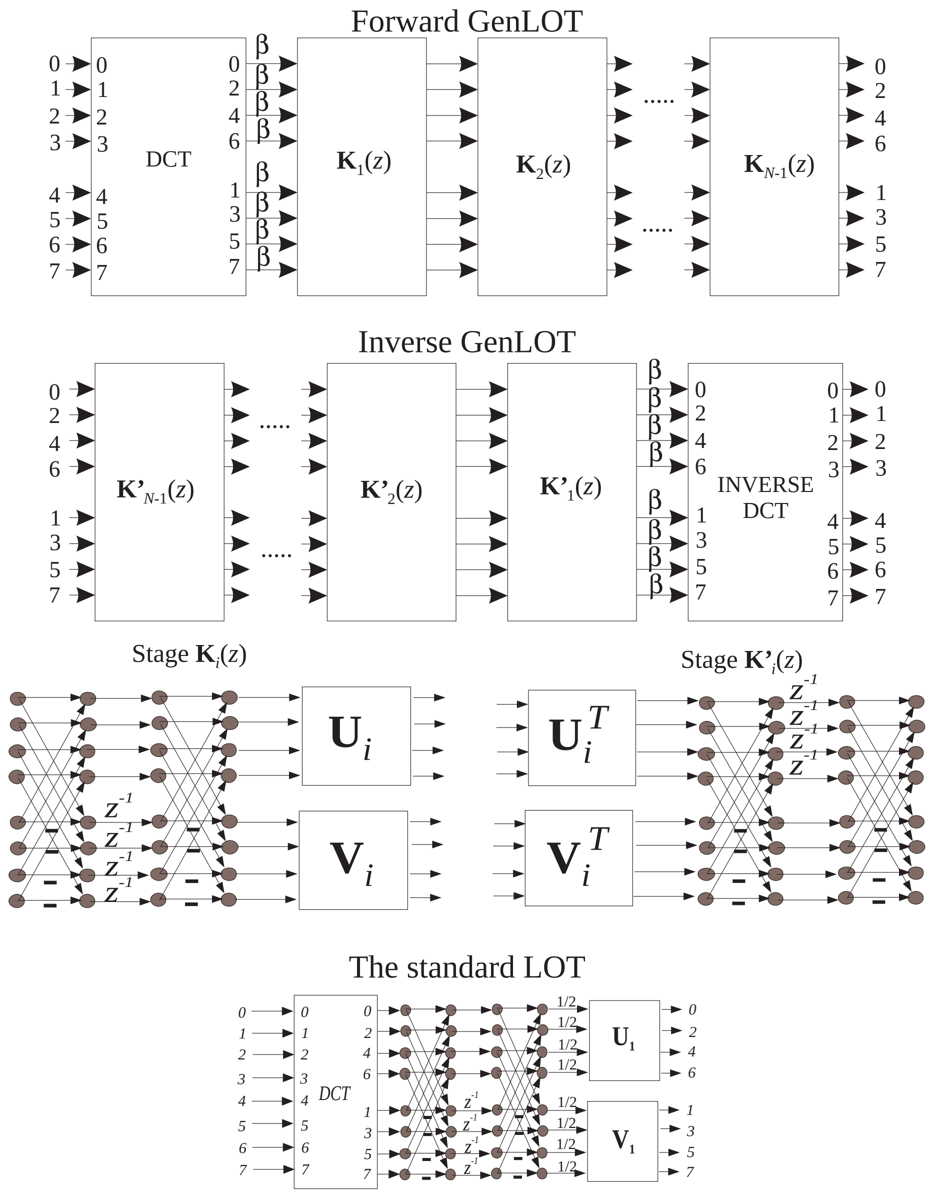

Figure 4: Implementation flow-graph for the GenLOT with $M = 8$, where $\beta = 2^{N-1}$.

Bibliography

[1] P. Cassereau, *A New Class of Optimal Unitary Transforms for Image Processing*, Master's Thesis, Mass. Inst. Tech., Cambridge, MA, May 1985.

[2] R. J. Clarke, *Transform Coding of Images,* Orlando, FL: Academic Press, 1985.

[3] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[4] H. Jozawa and H. Watanabe, "Intrafield/Interfield adaptive lapped transform for compatible HDTV coding," *4th International Workshop on HDTV and Beyond,* Torino, Italy, Sept. 4-6, 1991.

[5] H. S. Malvar, *Optimal pre- and post-filtering in noisy sampled-data systems,* Ph.D. Dissertation, Mass. Inst. Tech., Cambridge, MA, Aug. 1986.

[6] H. S. Malvar, "Reduction of blocking effects in image coding with a lapped orthogonal transform," *Proc. of Intl. Conf. on Acoust., Speech, Signal Processing*, Glasgow, Scotland, pp. 781-784, Apr. 1988.

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

[8] H. S. Malvar, "Lapped transforms for efficient transform/subband coding," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-38, pp. 969–978, June 1990.

[9] H. S. Malvar, "The LOT: a link between block transform coding and multirate filter banks," *Proc. Intl. Symp. Circuits and Systems,* Espoo, Finland, pp. 835–838, June 1988.

[10] H. S. Malvar, "Modulated QMF filter banks with perfect reconstruction," *Elect. Letters,* vol.26, pp. 906-907, June 1990.

[11] H. S. Malvar, "Extended lapped transform: fast algorithms and applications," *Proc. of Intl. Conf. on Acoust., Speech, Signal Processing*, Toronto, Canada, pp. 1797–1800, 1991.

[12] H. S. Malvar, *Signal Processing with Lapped Transforms.* Norwood, MA: Artech House, 1992.

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

[14] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Compression Standard,"* New York, NY: Van Nostrand Reinhold, 1993.

[15] J. P. Princen and A. B. Bradley, "Analysis/synthesis filter bank design based on time domain aliasing cancellation," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-34, pp. 1153–1161, Oct. 1986.

[16] R.L. de Queiroz and K. R. Rao, "Time-varying lapped transforms and wavelet packets," *IEEE Trans. on Signal Processing,* vol. 41, pp. 3293–3305, Dec. 1993.

[17] R. L. de Queiroz, *On Lapped Transforms,* Ph.D Dissertation, The University of Texas at Arlington, August 1996.

[18] R. L. de Queiroz and K. R. Rao, "The extended lapped transform for image coding," *IEEE Trans. on Image Processing,* Vol. 4, pp. 828–832, June, 1995.

[19] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, "The class of GENLOTs: generalized linear-phase lapped orthogonal transforms," *Proc. IEEE Intl. Symp. Circuits and Systems,* London, UK, 1994.

[20] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, "The generalized lapped orthogonal transforms," *Electronics Letters,* vol. 30, pp. 107–107, Jan. 1994.

[21] R. L. de Queiroz, and K. R. Rao, "On orthogonal transforms of images using paraunitary filter banks," *Journal on Visual Communications and Image Representation,* Vol. 6, No. 2, pp. 142–153, June 1995.

[22] M. Rabbani and P.W. Jones, *Digital Image Compression Techniques,* SPIE Optical Engineering Press, Bellingham, WA, 1991.

[23] K. R. Rao and P. Yip *Discrete Cosine Transform : Algorithms, Advantages, Applications,* San Diego, CA : Academic Press, 1990.

[24] K. R. Rao (ed.), *Discrete transforms and their applications,* New York, NY: Van Nostrand Reinhold, 1985.

[25] H. Schiller, "Overlapping block transform for image coding preserving equal number of samples and coefficients," *Proc. SPIE, Visual Communications and Image Processing,* vol. 1001, pp. 834–839, 1988.

[26] A. K. Soman, P. P. Vaidyanathan, and T. Q. Nguyen, "Linear-phase paraunitary filter banks: theory, factorizations and applications," *IEEE Trans. on Signal Processing,* vol. 41, pp. 3480–3496, Dec. 1993.

[27] M. Temerinac and B. Edler, "A unified approach to lapped orthogonal transforms," *IEEE Trans. Image Processing*, vol. 1, pp. 111–116, Jan. 1992.

[28] P.P. Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[29] R. W. Young and N. G. Kingsbury, "Frequency domain estimation using a complex lapped transform," *IEEE Trans. Image Processing,* vol. 2, pp. 2–17, Jan. 1993.