Proc. of IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing, Salt Lake City, 2001

FRAME RECONSTRUCTION OF THE LAPLACIAN PYRAMID

Minh N. Do[†] *and Martin Vetterli*[†§]

[†]Laboratory for Audio-Visual Communications
Swiss Federal Institute of Technology Lausanne (EPFL), 1015 Lausanne, Switzerland

[§]Department of EECS, University of California at Berkeley, Berkeley CA 94720, USA

Email: {Minh.Do, Martin.Vetterli}@epfl.ch; Web: lcavwww.epfl.ch

ABSTRACT

We study the Laplacian pyramid (LP) as a frame operator, and this reveals that the usual reconstruction is suboptimal. With orthogonal filters, the LP is shown to be a tight frame, thus the optimal linear reconstruction using the dual frame operator has a simple structure as symmetrical with the forward transform. For more general cases, we propose an efficient filter bank for reconstruction in the LP that is shown to perform better than the usual method. Numerical results indicate that gains of more than 1 dB are actually achieved.

1. INTRODUCTION

Multiscale data representation is a powerful idea. In particular, it captures data in hierarchical structures where each level corresponds to a reduced-resolution approximation. One of the early examples of such a scheme is the Laplacian pyramid (LP), proposed by Burt and Adelson [1] for image coding. The basic idea of the LP is the following. First, derive a coarse approximation of the original signal by lowpass filtering and downsampling. Based on this coarse version, predict the original (by upsampling and filtering) and then calculate the difference as the prediction error. The process can be iterated on the coarse version. A diagram for the LP is shown in Fig. 1.

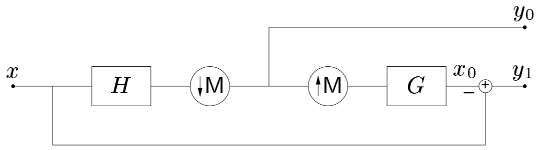

Fig. 1. Laplacian pyramid scheme.

A drawback of the LP is the implicit oversampling. Therefore, in compression applications it is normally replaced by subband coding or wavelet transform which is a critically sampled scheme and often an orthogonal decomposition. However, the LP has the advantage over the critically sampled wavelet scheme in the sense that each pyramid level generates only one bandpass image (even for multidimensional cases) which does not have "scrambled" frequencies. This frequency scrambling happens in the wavelet filter bank when a highpass channel after downsampling is folded back into the low frequency, and thus its spectrum is reflected. In the LP, this effect is avoided by only downsampling the lowpass channel. Therefore the LP permits further subband decomposition to be applied on its bandpass images. A possible scheme is a pyramidal decomposition where bandpass images of the LP are fed into directional filter banks. The end result is a set of directional subband images at multiple scales [2].

For many applications like compression and denoising, the coefficients in the transform domain are processed further and this can introduce errors due to quantization or thresholding. The processed coefficients are then used to reconstruct the original data. For the LP, the usual reconstruction – adding the prediction from the coarse version with the difference, produces perfect reconstructed image in the absence of noise but turns out to be suboptimal otherwise.

Our key observation is that the LP is a *frame* operator and that one should use the dual frame operator for reconstruction. While this seems a somewhat trivial observation, it has not been used in practice, probably because the usual reconstruction, while suboptimal, is so simple. Yet, we will show that gains of more than 1 dB are actually possible.

This paper is organized as follows. Section 2 introduces the notations and sets up the LP in the frame analysis where the dual frame operator or the pseudo-inverse is defined. Section 3 considers an important case where the LP is in fact a *tight frame*, and thus the dual frame operator has a simple structure as symmetrical with the forward transform. In Section 4, inspired by the structure of the new reconstruction algorithm for the tight frame case, we study a more general LP where it has better inverse operator with an efficient filter bank. Section 5 compares the two reconstruction algorithms for multilevel LP's. The performance improvement of the new reconstruction algorithm is illustrated with experimental results in Section 6.

2. PRELIMINARIES

2.1. Burt and Adelson's Laplacian Pyramid

We consider the general d-dimensional case where a discrete-time signal is a sequence of real-valued numbers defined on the integer lattice $\mathbb{Z}^d$, e.g. $x[n],\ n \in \mathbb{Z}^d$. The sampling operation is represented by a $d \times d$ nonsingular integer matrix M [3]. Signals with finite energy belong to a Hilbert space $l_2(\mathbb{Z}^d)$ with the inner product is defined as

$$\langle x, y \rangle = \sum_{n \in \mathbb{Z}^d} x[n] y[n],$$

This work was supported by a Department of Communication Systems - EPFL PhD Fellowship and the Swiss National Science Foundation grant number 21-52439.97.

and thus the l_2-norm is $\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum_{n\in\mathbb{Z}^d} x[n]^2}$.

With the notation as shown in Fig. 1, the filtering and downsampling operation for the LP yields

$$y_0[n] = \sum_{k\in\mathbb{Z}^d} x[k]h[\mathsf{M}n - k] = \langle x[\bullet], \tilde{h}[\bullet - \mathsf{M}n]\rangle, \qquad (1)$$

where we denote $\tilde{h}[n] = h[-n]$.

Similarly, the upsampling and filtering operation results in

$$x_0[n] = \sum_{k\in\mathbb{Z}^d} y_0[k]g[n - \mathsf{M}k]. \qquad (2)$$

Writing signals as column vectors, for example $\boldsymbol{x} = (x[n] : n \in \mathbb{Z}^d)^T$, we can express those operations as left matrix multiplications

$$\boldsymbol{y}_0 = \mathsf{H}\boldsymbol{x} \qquad \text{and} \qquad \boldsymbol{x}_0 = \mathsf{G}\boldsymbol{y}_0,$$

where H and G correspond to $(\downarrow \mathsf{M})H$ and $G(\uparrow \mathsf{M})$, respectively. In particular, H has $\left\{\tilde{h}[n - \mathsf{M}k]\right\}_{n\in\mathbb{Z}^d}$ as its rows and G has $\{g[n - \mathsf{M}k]\}_{n\in\mathbb{Z}^d}$ as its columns. In the sequel, we denote I as the identity matrices with appropriate sizes depending on the context.

With this matrix notation, the output of the LP in the highpass channels is

$$\boldsymbol{y}_1 = \boldsymbol{x} - \boldsymbol{y}_0 = \boldsymbol{x} - \mathsf{GH}\boldsymbol{x} = (\mathsf{I} - \mathsf{GH})\boldsymbol{x}.$$

By combining the previous relations, we can write the analysis operator of the LP as

$$\underbrace{\begin{pmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{pmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{pmatrix} \mathsf{H} \\ \mathsf{I} - \mathsf{GH} \end{pmatrix}}_{\mathsf{T}_a} \boldsymbol{x}. \qquad (3)$$

The usual inverse transform of the LP compute: $\hat{\boldsymbol{x}} = \mathsf{G}\boldsymbol{y}_0 + \boldsymbol{y}_1$, or

$$\hat{\boldsymbol{x}} = \underbrace{\begin{pmatrix} \mathsf{G} & \mathsf{I} \end{pmatrix}}_{\mathsf{T}_s} \begin{pmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{pmatrix}. \qquad (4)$$

It is easy to check that $\mathsf{T}_s\mathsf{T}_a = \mathsf{I}$ for any H and G. This agrees with the well-known fact that the LP can be perfectly reconstructed with any pair of filters H and G.

2.2. Frame Analysis

The frame analysis is a powerful concept in analyzing redundant representations [4]. For a detailed introduction to frames, readers are referred to [5] (Chapter 2). It can be shown that a linear operator is a frame if and only if it is invertible on its range with a bounded inverse. For the Laplacian pyramid, there always exists a bounded reconstruction algorithm (i.e. the usual algorithm), hence the LP is guaranteed to be a frame. In this case, the frame operator is represented by the left matrix multiplication with T_a.

Since the LP is a redundant transform, its frame operator admits an infinite number of left inverses. Among those, the most important one is called the *dual frame* operator, which is represented by the *pseudo inverse* of T_a [6]

$$\mathsf{T}_a^\dagger = (\mathsf{T}_a^T\mathsf{T}_a)^{-1}\mathsf{T}_a^T. \qquad (5)$$

When there is additive "noise" in the frame coefficients, the pseudo inverse eliminates the influence of errors that are orthogonal to the range of the frame operator. Therefore, if instead of having access to $\boldsymbol{y} = \mathsf{T}_a\boldsymbol{x}$ we have $\hat{\boldsymbol{y}} = \boldsymbol{y} + \boldsymbol{e}$, then the pseudo inverse provides the solution $\hat{\boldsymbol{x}} = \mathsf{T}_a^\dagger\hat{\boldsymbol{y}}$ that minimizes the residual $\|\mathsf{T}_a\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}\|$. This is called the least-squares solution.

Let T_a^{-1} be an arbitrary left inverse of T_a. With the noise model setup as above, we can write the reconstruction error when using T_a^{-1} as

$$\begin{aligned} \hat{\boldsymbol{x}} - \boldsymbol{x} &= \mathsf{T}_a^{-1}\hat{\boldsymbol{y}} - \boldsymbol{x} = \mathsf{T}_a^{-1}(\boldsymbol{y} + \boldsymbol{e}) - \boldsymbol{x} \\ &= \mathsf{T}_a^{-1}(\mathsf{T}_a\boldsymbol{x} + \boldsymbol{e}) - \boldsymbol{x} = \mathsf{T}_a^{-1}\boldsymbol{e}. \qquad (6) \end{aligned}$$

3. A TIGHT FRAME CASE

First, let us consider the special case where filters in the LP are orthogonal filters with respect to the sampling matrix M. More precisely, the orthogonal condition requires that

$$\langle g[\bullet], g[\bullet - \mathsf{M}n]\rangle = \delta[n], \qquad \text{and} \qquad (7)$$

$$h[n] = g[-n], \qquad \text{or} \qquad \mathsf{H} = \mathsf{G}^T. \qquad (8)$$

Such filters can be designed using well-known methods, including separable and non-separable ones [3, 7]. The LP with orthogonal filters possess an important property as stated in the following proposition.

Proposition 1 *The Laplacian pyramid with orthogonal filters, i.e. satisfying (7) and (8), is a* tight frame.

Proof: Under the orthogonality condition, (2) becomes

$$x_0[n] = \sum_{k\in\mathbb{Z}^d} \langle x[\bullet], g[\bullet - \mathsf{M}k]\rangle g[n - \mathsf{M}k].$$

Therefore if we denote V as the space spanned by the set of orthogonal vectors $\{g[\bullet - \mathsf{M}k]\}_{k\in\mathbb{Z}^d}$ then $\boldsymbol{x}_0$ is the orthogonal projection of $\boldsymbol{x}$ onto V. Together with the fact that $\boldsymbol{y}_1 = \boldsymbol{x} - \boldsymbol{x}_0$, using the *Pythagorean theorem* leads to

$$\|\boldsymbol{x}\|^2 = \|\boldsymbol{x}_0\|^2 + \|\boldsymbol{y}_1\|^2 = \|\boldsymbol{y}_0\|^2 + \|\boldsymbol{y}_1\|^2. \qquad (9)$$

where the equality $\|\boldsymbol{x}_0\| = \|\boldsymbol{y}_0\|$ comes from the fact that $\boldsymbol{y}_0$ are the coefficients in the orthonormal expansion (2) of $\boldsymbol{x}_0$ in V. □

Since T_a is a tight frame with $\mathsf{T}_a^T\mathsf{T}_a = \mathsf{I}$, its pseudo inverse is simply the transposed matrix. Thus we have

$$\mathsf{T}_a^\dagger = \begin{pmatrix} \mathsf{G}^T \\ \mathsf{I} - \mathsf{GG}^T \end{pmatrix}^T = \begin{pmatrix} \mathsf{G} & \mathsf{I} - \mathsf{GG}^T \end{pmatrix}. \qquad (10)$$

So the optimal (least-square) reconstruction using the pseudo inverse is

$$\hat{\boldsymbol{x}} = \mathsf{T}_a^\dagger\boldsymbol{y} = \mathsf{G}\boldsymbol{y}_0 + (\mathsf{I} - \mathsf{GG}^T)\boldsymbol{y}_1 = \mathsf{G}(\boldsymbol{y}_0 - \mathsf{H}\boldsymbol{y}_1) + \boldsymbol{y}_1.$$

The last expression above is derived in order to reduce the computational complexity of the pseudo inverse $\mathsf{T}_a^\dagger$. It leads to an efficient filter bank structure for reconstruction of the LP that is shown in Fig. 2.

It is worth to stress two important facts here. First, the usual inverse is different from the pseudo inverse, and thus it is suboptimal. Secondly, the pseudo inverse in this case has a symmetrical structure with the forward transform hence it has the same order of complexity.

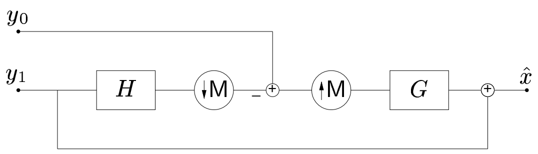

Fig. 2. The proposed reconstruction scheme for the Laplacian pyramid in Fig. 1. It is the pseudo inverse when the filters are orthogonal.

4. GENERAL CASES

In this section we consider the more general cases where H and G are arbitrary filters. Even though any frame operator has a pseudo inverse, for complexity reason we would like to only consider the cases where the inverse operator has a fast structured transform. Inspired by the tight frame case, we restrict our attention to the reconstruction algorithm that has a structure as shown in Fig. 2. We then turn the problem around by asking for which filters such an algorithm is indeed an inverse or pseudo inverse. This has the same flavor as the filter design problem for perfect reconstruction filter banks. The following proposition provides an answer to our question.

Proposition 2 ([8])
1. *The reconstruction shown in Figure 2 is an inverse transform of the LP* if and only if *two filters H and G are biorthogonal with respect to the sampling lattice* M, *which means the upper branch of the LP (*GH*) is a projector, or* $\mathsf{HG} = \mathsf{I}$.
2. *Furthermore, that reconstruction is the pseudo inverse* if and only if *the upper branch of the LP (*GH*) is an orthogonal projector.*

Recall that given a Hilbert space H, a linear operator P mapping H onto itself is called a *projector* if $P^2 = P$. Furthermore, if P is self-adjoint or $P = P^T$ then P is called an *orthogonal projector*.

Remark: It is interesting to note that the two conditions for the LP in the above proposition, i.e. projection and orthogonal projection, are exactly the same with the conditions for the improved LP's that was studied in [9]. Those conditions lead to LP with interpolation and least squares LP, respectively. The motivation for those modifications in [9] is to minimize the prediction error $\boldsymbol{y}_1$ of the LP, whereas our motivation in this work is to have a better reconstruction algorithm for the LP.

So the minimum requirement for the filter bank shown in Fig. 2 to be a bona fide inverse of the LP is the biorthogonality condition of the filters H and G. In this case, the LP has an interesting geometrical interpretation. Let us define two subspaces V and $\tilde{V}$ that are spanned by $\{g[\bullet - \mathsf{M}k]\}_{k\in\mathbb{Z}^d}$ and $\{\tilde{h}[\bullet - \mathsf{M}k]\}_{k\in\mathbb{Z}^d}$, respectively. These are also the column and row spaces of G and H. For all $\boldsymbol{x}$ in $l_2(\mathbb{Z}^d)$, the result of the approximation channel of the LP: $\boldsymbol{x}_0 = \mathsf{GH}\boldsymbol{x}$, is a projection of $\boldsymbol{x}$ onto V. Furthermore, since $\mathsf{HG} = \mathsf{I}$ the approximation error $\boldsymbol{y}_1 = \boldsymbol{x} - \boldsymbol{x}_0$ has property that

$$\mathsf{H}\boldsymbol{y}_1 = \mathsf{H}(\boldsymbol{x} - \mathsf{GH}\boldsymbol{x}) = \mathsf{H}\boldsymbol{x} - \mathsf{H}\boldsymbol{x} = \mathbf{0}.$$

In other words, the approximation error is perpendicular to $\tilde{V}$. This fact is illustrated in Fig. 3. The upper channel in the LP (GH) can be called an *oblique projector* [10] and denoted by $\tilde{P}_V$.

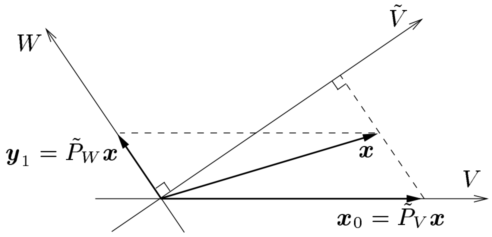

Fig. 3. Graphical representation of a oblique projector. $\boldsymbol{x}_0 = \tilde{P}_V \boldsymbol{x}$ is the projection of x onto V such that the approximation error $\boldsymbol{x} - \tilde{P}_V \boldsymbol{x}$ is perpendicular to $\tilde{V}$.

Let us define W the orthogonal complementary subspace of V, then it is easy to verify that $\boldsymbol{y}_1 = \boldsymbol{x} - \tilde{P}_V \boldsymbol{x}$ is a projection of $\boldsymbol{x}$ onto W such that the error is parallel to V (refer to Fig. 3). Denote this projection as $\tilde{P}_W$, $\tilde{P}_W = \mathsf{I} - \mathsf{GH}$.

Now let us compare the reconstruction using the usual method, denoted REC-1:

$$\boldsymbol{x}_1 = \mathsf{T}_s \boldsymbol{y} = \mathsf{G}\boldsymbol{y}_0 + \boldsymbol{y}_1, \tag{11}$$

and the one using the filter bank in Fig. 2, denoted REC-2:

$$\boldsymbol{x}_2 = \mathsf{T}_r \boldsymbol{y} = \mathsf{G}\boldsymbol{y}_0 + (\mathsf{I} - \mathsf{GH})\boldsymbol{y}_1. \tag{12}$$

Those two reconstruction algorithms are different in the way of handling the prediction error $\boldsymbol{y}_1$. More specifically, the REC-1 method adds directly $\boldsymbol{y}_1$ while the REC-2 method adds the $\tilde{P}_W$ projection of $\boldsymbol{y}_1$ to the reconstructed signal. So when there is noise in the LP coefficients, the REC-2 method eliminates the influence of errors in $\hat{\boldsymbol{y}}_1$ that are parallel to V.

In order to have more quantitative measurements in the performance of the two methods, suppose we wish to approximate $\boldsymbol{x}$ given $\hat{\boldsymbol{y}} = \mathsf{T}_a \boldsymbol{x} + \boldsymbol{e}$. With no further information about the error in the LP coefficients $\boldsymbol{e}$, it makes sense to chose $\hat{\boldsymbol{x}}$ to minimize the residual $\|\mathsf{T}_a \hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}\|$. As mentioned before, the optimal linear solution to this problem is the pseudo inverse of T_a. Using this as the measurement for the performance in reconstruction, the following result states that REC-2 always performs better than REC-1.

Proposition 3 ([8]) *Assume that H and G are biorthogonal filters. Let $\hat{\boldsymbol{x}}_1$ and $\hat{\boldsymbol{x}}_2$ be the results of reconstruction from noisy LP coefficients $\hat{\boldsymbol{y}}$ using REC-1 and REC-2, respectively. Then we have*

$$\|\mathsf{T}_a \hat{\boldsymbol{x}}_1 - \hat{\boldsymbol{y}}\| \geq \|\mathsf{T}_a \hat{\boldsymbol{x}}_2 - \hat{\boldsymbol{y}}\|, \tag{13}$$

where equality holds if and only if $\hat{\boldsymbol{y}}_1 = \mathbf{0}$.

5. MULTILEVEL LAPLACIAN PYRAMIDS

We now consider the situation where the LP scheme is iterated on the coarse version. We concentrate, for the sake of simplicity, on the one-dimensional case where $\mathsf{M} = 2$. By separating the sequences y_1 into two subsequences y_{10} and y_{11} of even and odd numbered samples, respectively, we can re-formulate the LP as a three channels filter banks (see Fig. 4).

Then, using polyphase-domain analysis, we can write the new analysis filters as [8]

$$\begin{aligned} K_0(z) &= 1 - H(z)G_0(z^2), \\ K_1(z) &= z - H(z)G_1(z^2), \end{aligned} \tag{14}$$

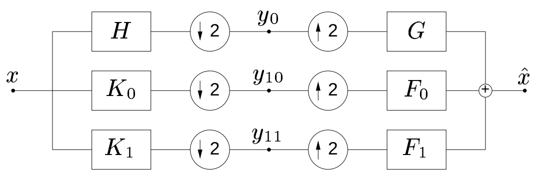

Fig. 4. Laplacian pyramid as an oversampled filter bank.

where $G_i(z)$ are the polyphase components of $G(z)$, or

$$G(z) = G_0(z^2) + z^{-1}G_1(z^2).$$

Since $H(z)$ and $G(z)$ are both lowpass filters, $K_0(z)$ and $K_1(z)$ are high pass filter.

On the synthesis side, for the REC-1 reconstruction we simply have

$$\dot{F}_0(z) = 1, \quad \text{and} \quad \dot{F}_1(z) = z. \tag{15}$$

Whereas, for the REC-2 reconstruction, the synthesis filters are

$$\ddot{F}_0(z) = K_0(z), \quad \text{and} \quad \ddot{F}_1(z) = K_1(z). \tag{16}$$

When the filter bank in Fig. 4 is iterated on the lowpass channel y_0, then by using the multirate identity which says that $G(z)$ followed by upsampling by 2 is equivalent to upsampling by 2 followed by $G(z^2)$ [7], we have equivalent synthesis filters at the n-level of a multilevel LP as

$$F_i^{(n)}(z) = F_i(z^{2^{n-1}}) \prod_{k=0}^{n-2} G(z^{2^k}), \quad i = 0, 1. \tag{17}$$

preceded by upsampling by 2^n. The impulse responses of $F_i^{(n)}(z)$ are the synthesis functions for the LP.

Now, consider what happen when (15) and (16) are substituted into (17). In the REC-1 method, we see that the synthesis functions for the LP are all low frequency signals. Thus, the errors from highpass subbands of a multilevel LP do not remain in these subbands but appear as broadband noise in the reconstructed signal. In [11], this effect was noted as the most serious disadvantages of the LP for coding applications.

On the other hand, in the REC-2 method, the synthesis functions have the similar frequency characteristics as the analysis functions. Therefore, reconstruction using REC-2 method remedies the previous mentioned problem of the LP.

6. EXPERIMENTAL RESULTS

In order to test the actual performance of the new reconstruction method with "real" signals we set up the experiments as follows. An input signal is transformed with a Laplacian pyramid. The LP coefficients are uniformly quantized, and this introduces errors depending to the quantization step. The quantized LP coefficients are then used to reconstruct the original signal using the usual and the new inverses. This is the typical scenario in lossy compression. We use the signal-to-noise ratios (SNR) between the original and the reconstructed signals as the measures of quality.

Experiments with different test signals and filters (using both orthogonal and biorthogonal ones) indicate that the new reconstruction provides significant improvements over the usual method. Fig. 5 plots a typical result. In this case the input signal is the 512×512 "Lena" image and the LP is computed with 5 levels using the Daubechies' orthogonal maximally flat 8-taps filters [5]. As can be seen, gains of more than 1 dB are actually obtained at each quantization step.

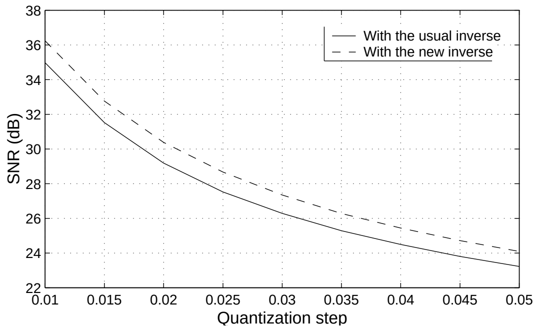

Fig. 5. Comparison of reconstruction from quantized Laplacian pyramid coefficients using two different inverses.

Acknowledgments: The authors would like to thank Prof. Michael Unser, Dr. Vivek Goyal and Irena Maravic for their helpful comments and suggestions.

7. REFERENCES

[1] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, April 1983.

[2] M. N. Do and M. Vetterli, "Pyramidal directional filter banks and curvelets," in *Proc. IEEE Int. Conf. on Image Proc.*, 2001, submitted.

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

[4] V. K. Goyal and J. Kovačević, "Quantized frame expansions with erasures," *Applied and Computational Harmonic Analysis*, 2001, to appear.

[5] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[7] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[8] M. N. Do and M. Vetterli, "Framing pyramids," *IEEE Trans. Signal Proc.*, 2001, submitted.

[9] M. Unser, "An improved least squares Laplacian pyramid for image compression," *Signal Proc.*, vol. 27, pp. 187–203, May 1992.

[10] M. Unser and A. Aldroubi, "A general sampling theory for non-ideal acquisition devices," *IEEE Trans. Signal Proc.*, pp. 2915–2925, 1994.

[11] E. P. Simoncelli and E. H. Adelson, "Subband transforms," in *Subband Image Coding*, J. W. Woods, Ed., pp. 143–192. Kluwer Academic Publishers, Inc., Boston, MA, 1991.