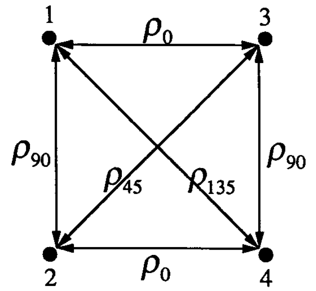

Fig. 1. Pairwise interactions between the four components of the local feature vector.

is a slightly more restrictive constraint:

$$R_x = \sigma^2 \begin{bmatrix} 1 & \rho_{90} \\ \rho_{90} & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & \rho_0 \\ \rho_0 & 1 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 1 & \rho_{90} & \rho_0 & \rho_0\rho_{90} \\ \rho_{90} & 1 & \rho_0\rho_{90} & \rho_0 \\ \rho_0 & \rho_0\rho_{90} & 1 & \rho_{90} \\ \rho_0\rho_{90} & \rho_0 & \rho_{90} & 1 \end{bmatrix}. \quad (2)$$

We note that this latter case corresponds to the direct separable extension of the results reported in [1], since the 2×2 DHT can be constructed from the outer product of horizontal and vertical one-dimensional sum and difference operators. The ability of the 2×2 DHT to decorrelate real texture images is usually excellent and is sufficient for most practical purposes. For instance, it was shown in [4] that the 2×2 DHT as applied to each of 12 Brodatz textures of an experimental set almost identical to that used by Cohen *et al.*, reduces the nondiagonal energy of the covariance matrix to less than 1% (with one exception at 2.2%) of its initial contribution.

References

[1] M. Unser, "Sum and difference histograms for texture classification," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. PAMI-8, pp. 118-125, Jan. 1986.
[2] M. Unser, "Local linear transforms for texture analysis," in *Proc. Seventh Int. Conf. Patt. Recog.* (Montreal, Canada), July 30-Aug. 2, 1984, pp. 1206-1208.
[3] M. Unser, "Local linear transforms for texture measurements," *Signal Processing*, vol. 11, pp. 61-79, July 1986.
[4] M. Unser, "Description statistique de textures: Application à l'inspection automatique," Ph.D. dissertation 534, Swiss Federal Inst. Technology, Lausanne, Switzerland, 1984.

On Coefficient-Quantization and Computational Roundoff Effects in Lossless Multirate Filter Banks

P. P. Vaidyanathan

Abstract—**FIR lossless transfer matrices have found recent applications in multirate analysis/synthesis systems having the perfect reconstruction property. It is shown in this correspondence that these lossless systems can be implemented such that regardless of coefficient quantization, the lossless property (and hence perfect reconstruction) can be retained. Such a result was shown to be true in the past only for two-channel filter banks. This correspondence also presents a noise gain result for lossless systems which finds application in analyzing roundoff noise in multirate filter banks.**

Manuscript received September 9, 1989; revised July 23, 1990. This work was supported in part by the National Science Foundation under Grants DCI8552579 and MIP8604456.
The author is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.
IEEE Log Number 9042289.

I. Introduction

Multirate filter banks with the perfect reconstruction property have received considerable attention recently (see [1]-[7] and references therein). In this correspondence we are concerned with the technique reported in [3]-[7], in which the polyphase matrix of the set of analysis filters is FIR and lossless. Our purpose here is to present two results, one concerning coefficient quantization and the other concerning roundoff noise propagation.

Unless mentioned otherwise, the notations used are the same as those in any one of the [3]-[7]. As a review, $H_k(z)$ and $F_k(z)$, $0 \le k \le M-1$ represent the M analysis and synthesis filters of an M band maximally decimated filter bank [6, fig. 1]. The analysis and synthesis banks are associated with two $M \times M$ matrices $\boldsymbol{E}(z)$ and $\boldsymbol{R}(z)$ called polyphase component matrices [6, fig. 2]. The methods employed in [3]-[7] are such that $\boldsymbol{E}(z)$ is (FIR and) lossless. This means that $\boldsymbol{E}(z)$ satisfies

$$\tilde{\boldsymbol{E}}(z)\,\boldsymbol{E}(z) = c^2\boldsymbol{I}, \quad \forall z, \quad c \text{ real.} \quad (1)$$

With $\boldsymbol{E}(z)$ satisfying this property, the maximally decimated system has perfect reconstruction if and only if the synthesis filters are chosen as $F_k(z) = \alpha z^{-L}\tilde{H}_k(z)$ where $\alpha \ne 0$ and L is an arbitrary integer.

If $c^2 = 1$ in (1), we say that $\boldsymbol{E}(z)$ is normalized-lossless. Note that losslessness can also be defined for rectangular matrices [5]. For $M \times 1$ stable systems, losslessness is same as the power-complementary property [7].

In order to design and implement filter banks with lossless $\boldsymbol{E}(z)$, it is necessary to obtain a structure for $M \times M$ FIR lossless systems. Structures for arbitrary real-coefficient FIR lossless systems were presented in [5] based on a state-space approach, with planar rotations as building blocks. More recently, a method was developed in [6] which leads to a different representation of lossless systems. This method does not involve rotations, but is entirely in terms of diadics (i.e., matrices of the form $\boldsymbol{v}\boldsymbol{v}^\dagger$ where $\boldsymbol{v}$ in a column vector). This is simpler in terms of derivation as well as implementation. Also, it covers both real and complex coefficient FIR lossless systems. In this correspondence, we shall show that the diadic based structure retains losslessness in spite of coefficient quantization (for arbitrary M), whereas such a property is not true for the rotation based structure except in the $M = 2$ case.

It should be noticed that the above discussion does not take into account errors due to quantization of signals (such as state variables) in the structure. The noise due to this must be separately analyzed; thus in Section IV we present a result on noise amplification by lossless systems, and indicate applications in filter bank structures.

II. Coefficient Quantization in Rotation-Based Structures

First consider a degree-0 lossless system $\boldsymbol{R}$ (i.e., a constant unitary matrix). The unitary property is equivalent to

$$\boldsymbol{R}_k^\dagger \boldsymbol{R}_m = 0, \quad k \ne m \quad (2)$$

$$\boldsymbol{R}_m^\dagger \boldsymbol{R}_m = d > 0 \quad (3)$$

where $\boldsymbol{R}_m$ is the mth column of $\boldsymbol{R}$. Such matrices can be expressed in terms of planar rotations [8]. For example, in the 3×3 real-coefficient case

$$R = \sqrt{d}\begin{bmatrix} \mu_0 & 0 & 0 \\ 0 & \mu_1 & 0 \\ 0 & 0 & \mu_2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & c_0 & s_0 \\ 0 & -s_0 & c_0 \end{bmatrix} \cdot \begin{bmatrix} c_1 & 0 & s_1 \\ 0 & 1 & 0 \\ -s_1 & 0 & c_1 \end{bmatrix}\begin{bmatrix} c_2 & s_2 & 0 \\ -s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

where $c_m = \cos\theta_m$, $s_m = \sin\theta_m$ and $\mu_i = \pm 1$. When the multipliers c_m and s_m are quantized to $Q[c_m]$ and $Q[s_m]$, then $Q^2[c_m] + Q^2[s_m] \neq 1$ any more. So each nondiagonal factor in (4) is such that it has unequal column norms. In other words, the factor satisfies (2) but not (3). As a result, one can verify that the product (4) satisfies neither (2) nor (3), i.e., is not unitary. One could consider replacing the entry "1" in each nondiagonal factor of (4) with the norm of the remaining columns after quantization, but norms involve square roots which usually require infinite precision.

For similar reasons, structures for lossless systems of arbitrary degree based on rotations [5] do not retain the lossless property when the cosines and sines are quantized. (An exception to this is the 2×2 case, as elaborated in [4]. In this case there is no difficulty because the column-norms of 2×2 orthogonal matrices are equal even after quantization.)

III. Coefficient Quantization in Dyadic-Based Structures

First consider constant unitary matrices with possibly complex entries. It is well known [9], [10] that such a matrix can be factorized as

$$R = \sqrt{d}\,G_1 \cdots G_{M-1} D \quad (5)$$

where G_m are Householder matrices, i.e., $M \times M$ matrices of the form

$$G_m = I - 2u_m u_m^\dagger \quad (6)$$

where u_m is a unit-norm column vector. It is easily verified that $G_m^\dagger G_m = I$. The quantity D in (5) is diagonal with $D_{ii} = e^{j\phi_i}$. If we implement a unitary matrix in factored form as in (5), then the quantization of the components of u_m results in loss of the unit-norm property of u_m so that G_m (and hence R) does not remain unitary. However, a simple trick can be used to overcome this difficulty. For this, note that the modified matrix $H_m \triangleq \|u_m\|^2 I - 2u_m u_m^\dagger$ satisfies $H_m^\dagger H_m = \|u_m\|^4 I$, i.e., is unitary for any nonzero vector u_m. We can therefore express any unitary R as

$$R = H_1 \cdots H_{M-1}\hat{D} \quad (7)$$

where $\hat{D} = \sqrt{d}D/\alpha$, $\alpha \triangleq \Pi\,\|u_m\|^2$. When the components of u_m are quantized, the matrices H_m remain unitary and so does the product $H_1 \cdots H_{M-1}$. But we have to be careful about the diagonal matrix $\hat{D}$ (which can be complex). Under unquantized conditions the diagonal elements have equal norm ($= \sqrt{d}/\alpha$). Upon quantization of the diagonal entries, the norms can become unequal. So the product (7) after quantization satisfies (2) but not necessarily (3). In other words, the columns of R are still pairwise orthogonal but may have unequal norms. This is, of course, better than losing both the properties (2) and (3) as in the rotation-based structure. Finally, for real matrices the diagonal elements of $\hat{D}$ are real and have the form $\pm\sqrt{d}/\alpha$ and have equal norm even after quantization so that both (2) and (3) hold.

Now consider lossless systems with arbitrary degree. Every $M \times M$ FIR lossless system of degree N (with possibly complex coefficients) can be represented [6] as $E(z) = V_N(z) \cdots V_1(z)S$ where $V_m(z) \triangleq I - v_m v_m^\dagger + z^{-1}v_m v_m^\dagger$. Here v_m is $M \times 1$ (so that $V_m(z)$ has degree $= 1$), with $\|v_m\| = 1$. The unit-norm constraint on v_m ensures that $V_m(z)$ is lossless. The quantity S here is $M \times M$ unitary. Now consider quantization. When v_m is quantized, the unit-norm property is usually lost so that $V_m(z)$ (and hence $E(z)$) does not remain lossless. To avoid this difficulty we define

$$U_m(z) \triangleq \|v_m\|^2 I - v_m v_m^\dagger + z^{-1}v_m v_m^\dagger \quad (8)$$

and rewrite $E(z) = U_N(z) \cdots U_1(z)R$ where $R = S/\beta$, $\beta = \Pi\,\|v_m\|^2$. Equation (8) represents a lossless system for any nonzero v_m. The unit-norm constraint on v_m is not required any more. When v_m are quantized, $U_m(z)$ still remains lossless so that $U_N(z) \cdots U_1(z)$ remains lossless. The matrix R, which is unitary, can be represented as in (7) so that it remains unitary (except that in the complex case the column norms can become unequal). Summarizing, $E(z)$ continues to satisfy $\tilde{E}(z)E(z) = J$ inspite of quantization, where J is a constant diagonal matrix. In the real-coefficient case v_m and R real so that $J = c^2 I$, i.e., $E(z)$ remains lossless. As a result one can implement a M-band FIR perfect reconstruction system which retains the perfect reconstruction property in spite of coefficient quantization. It is true that the multipliers $\|v_m\|^2$ require higher precision than the components of v_m but these precisions remain finite. Notice that the modified structure does not require multipliers involving square roots.

Similar results are easily derived [11] for IIR lossless systems (both $M \times M$ and $M \times 1$) but are not of immediate relevance here.

IV. Roundoff Noise Considerations

All structures for lossless systems reported in [5] and [6] have the form shown in Fig. 1 where $V_m(z)$ are degree-one lossless systems and R is constant unitary satisfying (2), (3). Assume $d = 1$ for simplicity. Let the computation inside each building block $V_m(z)$ and R be performed without quantization. Assume that quantizers are inserted only at the outputs of these building blocks so that the noise model is as shown in Fig. 1. Here each $e_m(n)$ is a vector of M noise sources. We shall make standard assumptions about the noise model [12]. Thus the quantizers are b-bit fixed point roundoff devices generating zero-mean white noise with variance σ_e^2, and the noise components in $e_m(n)$ are mutually uncorrelated. The following result helps to simplify noise analysis in lossless systems.

The Noise Gain Lemma: Let $H(z)$ be a system such that $H^T(z)$ is normalized-lossless. Let the input vector $u(n)$ be wide-sense stationary, with zero mean and

$$E[u(n)\,u^\dagger(n')] = \sigma^2\delta(n - n')I. \quad (9)$$

Then the output $y(n)$ has zero mean and satisfies

$$E[y(n)\,y^\dagger(n')] = \sigma^2\delta(n - n')I. \quad (10)$$

Proof: Let $H(z) = \Sigma_n\, h(n)z^{-n}$ so that $y(n) = \Sigma_m\, h(m)\,u(n - m)$. Losslessness of $H^T(z)$ implies $\Sigma_m\, h(m)\,h^\dagger(k + m) = \delta(k)I$. By using this and (9) it can be verified that $E[y(n)\,y^\dagger(n')]$ indeed simplifies to the form (10). ■

According to the Lemma, if the components of $u(n)$ are zero mean, white, and uncorrelated with equal variance, then so are the components of $y(n)$. Returning to Fig. 1, each of the noise sources $e_m(n)$ faces a noise transfer function which is $M \times M$ normalized-lossless. Since $e_m(n)$ and $e_{m'}(n')$ are assumed to be uncorrelated for $m \neq m'$, the total noise $f(n)$ at the output terminal is zero mean with $E[f(n)\,f^\dagger(n')] = (N + 1)\,\delta(n - n')\sigma_e^2 I$. In other words, each component $y_i(n)$ of the output $y(n)$ is contaminated with additive white noise of variance $(N + 1)\sigma_e^2$.

Note that the noise gain Lemma holds also for rectangular matrices $H(z)$. This is useful in analyzing the noise gain due to a power-complementary set of M synthesis filters $F(z) = [F_0(z) \cdots F_{M-1}(z)]$ (Fig. 2). Such a set satisfies $\Sigma_m |F_m(e^{j\omega})|^2 = 1$ so that $F^T(z)$ is lossless. Assume that at the input of each filter $F_m(z)$ we have a white noise component with variance σ_e^2, from some source. If the noise components are uncorrelated, then the noise at the output is white with variance σ_e^2.

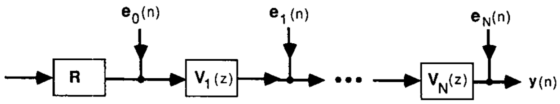

Fig. 1. Noise model for the FIR lossless structure.

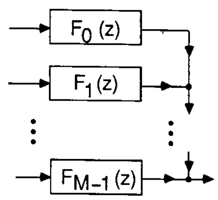

Fig. 2. A power complementary synthesis bank.

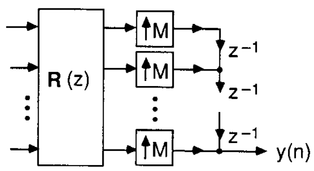

Fig. 3. Polyphase implementation of a synthesis bank.

As a final example, consider Fig. 3 which is a polyphase implementation of the synthesis bank. Here $\boldsymbol{R}(z)$ has noise components at the input (possibly from quantizers at the analysis end). If the noise sources satisfy the usual assumptions and if $\boldsymbol{R}(z)$ is normalized-lossless (as in the perfect reconstruction systems in [3]) then the noise components at the output of $\boldsymbol{R}(z)$ are white and uncorrelated, with equal variance σ_e^2. Since the noise contaminating $y(n)$ is merely a time-multiplexed version of these components, it is white with variance σ_e^2. In practice, the noise generated due to quantization inside $\boldsymbol{R}(z)$ should be added to this; it can be analyzed in a similar way.

V. Concluding Remarks

Note that in many multirate filter bank applications, a major source of "noise" is the coding error, introduced at the output of the analysis bank. This noise does not satisfy the standard assumptions stated above, so the above results do not apply. The noise gain Lemma is primarily useful for analyzing roundoff noise generated in the filter structure.

It can be shown that the lossless nature of building blocks also results in automatic $\mathcal{L}_2$-scaling of internal nodes. Detailed analysis of this can be found in [13].

The fact that unitary matrices can be implemented without loss of unitary property under quantization appears to have escaped the attention of researchers (even though the usefulness of Householder matrices are very well known [9], [14]). This property might be valuable in implementation of numerical techniques for least squares problems [14].

References

[1] M. Vetterli, "Filter banks allowing for perfect reconstruction," *Signal Processing*, vol. 10, pp. 219–244, Apr. 1986.

[2] M. J. T. Smith and T. P. Barnwell, III, "Exact reconstruction techniques for tree-structured subband coders," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 434–441, June 1986.

[3] P. P. Vaidyanathan, "Theory and design of *M*-channel maximally decimated quadrature mirror filters with arbitrary *M*, having perfect reconstruction property," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, pp. 476–492, Apr. 1987.

[4] P. P. Vaidyanathan and P.-Q. Hoang, "Lattice structures for optimal design and robust implementation of two-channel perfect reconstruction QMF banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 81–94, Jan. 1988.

[5] Z. Doğanata, P. P. Vaidyanathan, and T. Q. Nguyen, "General synthesis procedures for FIR lossless transfer matrices, for perfect-reconstruction multirate filter bank applications," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1561–1574, Oct. 1988.

[6] P. P. Vaidyanathan, T. Q. Nguyen, and Z. Doğanata, and T. Saramaki, "Improved approach for design of perfect reconstruction FIR QMF banks with lossless polyphase structures," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1042–1056, July 1989.

[7] P. P. Vaidyanathan, "Quadrature mirror filter banks, *M*-band extensions and perfect-reconstruction techniques," *IEEE ASSP Mag.*, vol. 4, pp. 4–20, July 1987.

[8] F. D. Murnaghan, *The Unitary and Rotation Groups*. Washington, DC: Spartan, 1962.

[9] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins University Press, 1989.

[10] J. N. Franklin, *Matrix Theory*. Englewood Cliffs, NJ: Prentice Hall, 1968.

[11] Z. Doganata and P. P. Vaidyanathan, "New minimal structures for the implementation of PC IIR systems," in *Proc. ISCAS*, May 1989, pp. 1660–1663.

[12] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1975.

[13] P. P. Vaidyanathan, "Passive cascaded lattice structures for low sensitivity FIR filter design, with applications to filter banks," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 1045–1064, Nov. 1986.

[14] A. O. Steinhardt, "Householder transforms in signal processing," *IEEE ASSP Mag.*, pp. 4–12, July 1988.

Automatic Segmentation of Speech

Jan P. van Hemert

Abstract—**This correspondence describes a method for automatic segmentation of speech. First the incoming utterance is split up into more or less stationary parts, then these stationary parts are labelled as phones using the phonetic transcription of the utterance.**

I. Introduction

For the purpose of recognition or synthesis, speech often needs to be segmented into phonetic units. Manual segmentation is tedious and time consuming, and the results lack reproducibility because of the subjective decisions involved. This calls for automatic segmentation methods. In this correspondence we describe such an automatic method. An important application of the method is the automatic generation of segment libraries for speech synthesis. With this method, diphone libraries for Dutch, German, and British English have been made in only a fraction of the time a manual method would have taken. A perceptual test has shown that isolated Dutch words constructed with diphones generated by this method compare favorably, both in intelligibility and in speech quality, to words synthesized using a manually prepared inventory.

Manuscript received August 29, 1988; revised January 31, 1990.

The author was with the Institute for Perception Research, 5600 MB Eindhoven, The Netherlands. He is now with the Philips Innovation Center, D-5100 Aachen, Germany.

IEEE Log Number 9042286.