APPLICATION OF UNIMODULAR MATRICES TO SIGNAL COMPRESSION

See-May Phoong
Dept. of EE & Grad. Inst. of Comm. Engr.
National Taiwan Univ.
Taipei, Taiwan, R.O.C.

Yuan-Pei Lin
Dept. Electrical and Control Engr.
National Chiao Tung Univ.
Hsinchu, Taiwan, R.O.C.

ABSTRACT

In the past decade, the application of orthonormal or paraunitary (PU) matrices to subband coding has drawn considerably attention. In this paper, we will study signal compression using unimodular matrices. Like PU matrices, unimodular matrices have FIR inverse. However unimodular matrices do not have the energy preservation property. Applying unimodular matrices to subband coding is not always a good choice as quantization noise might be amplified and coding gain can be less than one . We introduce a new structure for unimodular coders. The new structure consists of a closed-loop vector DPCM structure followed by a transform coder. Using such a structure, it is shown that the coding gain of unimodular coders is never less than 1. Simulation example shows that despite having the smallest system delay, unimodular coders have a better performance than subband coding with the same complexity for AR(1) input.

1. INTRODUCTION

Orthonormal or paraunitary (PU) matrices have found many applications in signal processing. In particular, they have been successfully applied to subband coding. PU matrices enjoy many desired features and have been widely studied (see [1] and references therein). In this paper, we will consider unimodular matrices for signal compression. Like PU matrices, unimodular matrices have the advantage that if they are used as polyphase matrices of filter banks (FB), perfect reconstruction (PR) can be obtained by FIR analysis and synthesis filters. In addition, unimodular matrices also enjoy the advantage of having the smallest system delay for all FBs. System delay is particular important in applications such as speech coding [2]. Though there are efficient design methods for low delay FBs [2][3], there are relatively few results on unimodular FBs.

The earliest paper that studied the relationship between unimodular matrices and FIR PR FB is [4]. The author introduces the most general degree one unimodular matrices and demonstrates that unlike PU matrices, there are unfactorizable unimodular matrices. In [5], we show that though there are unfactorizable higher order unimodular matrices, all first-order unimodular matrices (or lapped unimodular transforms, LUTs) can be minimally factorized into degree-one building blocks.

In this paper, we will apply the unimodular matrices to signal compression. It is well known that subband coding using non PU matrices suffers from quantization noise amplification problem. Therefore subband coding structure is not a good choice for implementing unimodular coders. In this paper, we introduce a new structure for the unimodular coders. The new structure consists of a closed-loop vector DPCM structure followed by a transform coder. It enjoys the unity noise gain property for any quantization noise model. Using the proposed structure, the coding gain of unimodular coders is never less than 1. Simulation example shows that for the coding of AR(1) process, the first-order unimodular coder outperforms LOT coder.

THIS WORK WAS SUPPORTED BY NSC 90-2213-E-002-097 AND 90-2213-E-009-108, TAIWAN, R.O.C.

2. REVIEW OF UNIMODULAR MATRICES

A causal matrix $\mathbf{A}(z)$ is unimodular if its determinant $\det[\mathbf{A}(z)] = c$ for some nonzero constant. Without loss of generality, we assume that $c = 1$ in this paper. The inverse of a causal unimodular matrix is also causal unimodular. If a causal unimodular matrix and its inverse are used as polyphase matrices of FBs, then PR is obtained with causal FIR analysis and synthesis filters. The system delay of an M-channel unimodular FB is always $(M-1)$, regardless of the filter length. Many useful properties of unimodular matrices can be found in [4] [6] [5]. In particular, for a causal unimodular matrix $\mathbf{A}_0 + \mathbf{A}_1 z^{-1} + \ldots + \mathbf{A}_N z^{-N}$, its first coefficient matrix $\mathbf{A}_0$ is always nonsingular. Moreover the most general degree-one unimodular matrix has the form $\mathbf{A}_0\mathbf{D}(z)$, where $\mathbf{A}_0$ is a nonsingular matrix and

$$\mathbf{D}(z) = \mathbf{I} + \mathbf{u}\mathbf{v}^\dagger z^{-1}, \quad \mathbf{v}^\dagger\mathbf{u} = 0. \tag{1}$$

The inverse of $\mathbf{D}(z)$ is given by $\mathbf{D}^{-1}(z) = \mathbf{I} - \mathbf{u}\mathbf{v}^\dagger z^{-1}$, which is also a degree-one unimodular system. Using $\mathbf{D}(z)$ as a building block, it is shown [5] that any first-order unimodular matrices can be decomposed into degree-one building blocks. Though there are higher order unfactorizable unimodular matrices, a broad class of useful unimodular matrices can be obtained by cascading the degree-one building block $\mathbf{D}(z)$.

3. REVIEW OF VECTOR DPCM SCHEME

The theory of DPCM is well known and can be found in many references. Let $\mathbf{x}(n)$ be a zero-mean WSS vector random process. Consider the Nth order prediction error vector $\mathbf{e}(n) = \mathbf{x}(n) - \mathbf{P}_1\mathbf{x}(n-1) - \ldots - \mathbf{P}_N\mathbf{x}(n-N)$. The prediction error polynomial matrix is given by

$$\mathbf{B}(z) = \mathbf{I} - \mathbf{P}(z),$$

where $\mathbf{P}(z) = \sum_{i=1}^{N} \mathbf{P}_i z^{-i}$. In a DPCM scheme, we quantize the prediction error $\mathbf{e}(n)$ instead of the original signal $\mathbf{x}(n)$. The reconstructed signal is obtained at the decoder whose transfer matrix is $\mathbf{B}^{-1}(z)$. An open loop DPCM structure will encounter noise amplification problem. Thus the closed loop DPCM structure shown in Fig. 1 is introduced for the implementation of the DPCM encoder. It is well known that the noise transfer function of DPCM scheme using closed loop structure is $\mathbf{I}$. Hence closed loop structure enjoys the unity noise gain property. The quantizer $\mathbf{Q}$ in Fig. 1 is in general a vector quantizer (VQ) as in the predictive VQ.

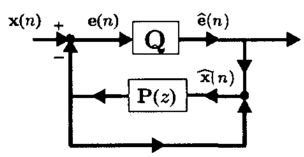

Figure 1: Closed loop structure for the vector DPCM encoder.

4. A NEW CODING STRUCTURE

Let the encoder be the M by M polynomial matrix $\mathbf{A}(z) = \sum_{k=0}^{N} \mathbf{A}_k z^{-k}$. To ensure that the inverse $\mathbf{A}^{-1}(z)$ has a causal realization, we assume that the matrix $\mathbf{A}_0$ is nonsingular. In the case when $\mathbf{A}(z)$ is unimodular, there is no loss of generality in making such an assumption because $\mathbf{A}_0$ is always nonsingular. Under this assumption, we can rewrite $\mathbf{A}(z)$ as

$$\mathbf{A}(z) = \mathbf{A}_0[\mathbf{I} - \mathbf{P}_1 z^{-1} \ldots - \mathbf{P}_N z^{-N}] = \mathbf{A}_0[\mathbf{I} - \mathbf{P}(z)]. \quad (2)$$

Using a technique similar to that in a closed loop DPCM system, we can implement the encoder $\mathbf{A}(z)$ as Fig. 2. As the constant term of $\mathbf{P}(z)$ is a zero matrix, there is no delay free loop in Fig. 2. The new structure is different from the vector DPCM structure in Fig. 1, it reduces to the vector DPCM structure when $\mathbf{A}_0 = \mathbf{I}$. In the absence of quantizers $\mathbf{Q}$, it can be shown that the transfer function from $\mathbf{x}(n)$ to $\mathbf{e}(n)$ is $\mathbf{A}(z)$. One can also show that the transfer function of the decoder is $[\mathbf{I} - \mathbf{P}(z)]^{-1}\mathbf{A}_0^{-1}$, which is equal to $\mathbf{A}^{-1}(z)$. In the absence of quantizers, the vector $\mathbf{r}(n)$ in Fig. 2 can be expressed as $\mathbf{r}(n) = \mathbf{x}(n) - \mathbf{P}_1\mathbf{x}(n-1) - \cdots - \mathbf{P}_N\mathbf{x}(n-N)$. Compared with the DPCM system, one immediately realizes that $\mathbf{P}(z)$ is in fact a predictor of $\mathbf{x}(n)$ and the prediction error is $\mathbf{r}(n)$. From the expression in (2), one can interpret the new structure in Fig. 2 as a predictive transform coder. The input $\mathbf{x}(n)$ is first passed through the prediction error polynomial $[\mathbf{I} - \mathbf{P}(z)]$ and the prediction vector $\mathbf{r}(n)$ is encoded using the transform coder $\mathbf{A}_0$.

In a predictive VQ scheme, a VQ is used to quantize the prediction error vector $\mathbf{e}(n)$. Here we assume that $\mathbf{Q}$ consists of a set of M scalar quantizers with different bit rate b_i so that the average bit rate is

$$b = \frac{1}{M}\sum_{k=0}^{M-1} b_k.$$

The use of scalar quantizers will not cause a major degradation on the performance. The reason is as follows. In a predictive VQ scheme, we have $\mathbf{A}_0 = \mathbf{I}$. Any two elements, say $e_i(n)$ and $e_j(n)$, in the prediction error vector $\mathbf{e}(n)$ are in general correlated. Hence a VQ is needed to encode $\mathbf{e}(n)$ efficiently. In the new coders, the nonsingular matrix $\mathbf{A}_0$ has the ability to decorrelate the error vector $\mathbf{e}(n)$. Hence scalar quantizers with suitable bit allocation can represent $\mathbf{e}(n)$ efficiently.

Remark: For the class of PU and CAFACAFI (CAusal Fir with AntiCAusal Fir Inverse) matrices, their first coefficient matrix $\mathbf{A}_0$ is always singular [6]. Therefore PU and CAFACAFI matrices can not be realized using the proposed structure in Fig. 2.

A. Noise Analysis

Let the quantization noise vector be $\mathbf{q}(n) = \hat{\mathbf{e}}(n) - \mathbf{e}(n) = [q_0(n) \ \ldots \ q_{M-1}(n)]^T$. Define the output noise vector as $\varepsilon(n) = \hat{\mathbf{x}}(n) - \mathbf{x}(n)$, where $\hat{\mathbf{x}}(n)$ is the output of the decoder as shown in Fig. 2. It is not difficult to show that the noise transfer matrix from $\mathbf{q}(n)$ to the output of the decoder is $\mathbf{A}_0^{-1}$. Therefore the output noise vector $\varepsilon(n)$ and the quantization error $\mathbf{q}(n)$ are related as $\varepsilon(n) = \mathbf{A}_0^{-1}\mathbf{q}(n)$. Using this relation, one immediately gets

$$\mathbf{R}_\varepsilon = E\{\varepsilon(n)\varepsilon^T(n)\} = \mathbf{A}_0^{-1}\mathbf{R}_\mathbf{q}\mathbf{A}_0^{-T}, \quad (3)$$

where $\mathbf{R}_\mathbf{q}$ is the autocorrelation matrix of $\mathbf{q}(n)$. Note that in the above derivation, we do not make any assumption on the quantization noise $\mathbf{q}(n)$. Therefore (3) holds for any additive noise model. The average output noise variance is given by $\sigma_{out}^2 = 1/M tr(\mathbf{R}_\varepsilon)$.

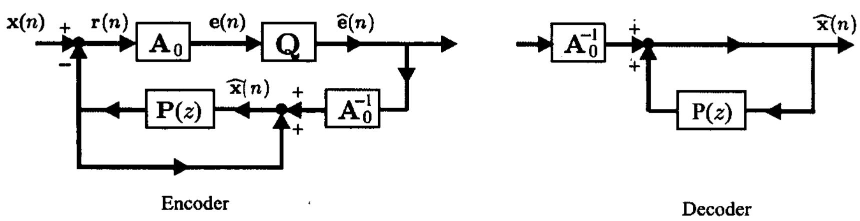

Figure 2: The new coding structure. The quantizers are a set of scalar quantizers.

B. Optimal Bit Allocation and Optimal A(z)
From the above discussion, the new coder in Fig. 2 can be viewed as a predictor followed by a transform coder. Therefore the optimization problem of b_i and nonsingular $\mathbf{A}_0$ (not necessary unitary) in the new coder is the same as that of a transform coder with input vector $\mathbf{r}(n)$. Under the high bit rate assumption, the autocorrelation matrix of the quantization noise $\mathbf{q}(n)$ is:

$$\mathbf{R}_q = diag[c2^{-2b_0}\sigma_{r_0}^2 \;\; c2^{-2b_1}\sigma_{r_1}^2 \;\; \dots \;\; c2^{-2b_{M-1}}\sigma_{r_{M-1}}^2],$$

where $\sigma_{r_k}^2$ is the variance of $r_k(n)$. Given this noise model, the transform $\mathbf{A}_0$ and the bit allocation b_i that minimize the trace of $\mathbf{R}_\epsilon$ in (3) is well-known. There are two optimal solutions which give the same minimum average output error variance σ_{out}^2:

S1. *Karhunen-Loeve Transform (KLT):* In the Appendix of [7], Vaidyanathan showed that the optimal nonsingular $\mathbf{A}_0$ that minimizes σ_{out}^2 is the well known KLT. The optimal bit allocation formula and the minimum achievable σ_{out}^2 are respectively given by [7]:

$$b_i = b + \log_2 \sigma_{e_i} - \frac{1}{M}\log_2 \prod_{i=0}^{M-1} \sigma_{e_i},$$

$$\sigma_{bit,\mathbf{A}_0}^2 = c\,2^{-2b}(\det \mathbf{R_r})^{1/M}, \qquad (4)$$

where $\mathbf{R_r}$ is the autocorrelation matrix of $\mathbf{r}(n)$. The subscript of '$bit, \mathbf{A}_0$' is chosen as a reminder that the error variance is obtained when the bit allocation and the matrix $\mathbf{A}_0$ are optimal. The optimal transform $\mathbf{A}_0$ is the unitary matrix that diagonalizes $\mathbf{R_r}$.

S2. *Prediction-based Lower triangular Transforms (PLT):* In [8], it was shown that though we cannot do better than the KLT, there are nonunitary transforms that can achieve the same coding performance. One of the solutions is a lower triangular transform with unity diagonal elements called PLT. Given an input vector $\mathbf{r}(n)$, the kth row of PLT matrix is formed by the optimal prediction coefficients when we predict $r_k(n)$ from $r_i(n)$ for $i < k$. As the PLT matrix is nonunitary, the quantization noise will be amplified at the reconstruction end. A unity noise gain structure called MINLAB [8] is introduced for the implementation of the PLT. By using the MINLAB structure, it is shown that the minimum achievable output noise variance is the same as $\sigma_{bit,\mathbf{A}_0}^2$ in (4). In addition to its excellent coding performance, the PLT enjoys many other attractive features [8], such as low design and implementation cost, structurally PR property, multiplier-less realization, and adaptability.

Under the optimal bit allocation and the optimal transform $\mathbf{A}_0$, the achievable lower bound on the average output noise variance is given by (4). Note that this lower bound depends on $\det \mathbf{R_r}$. From the expression of $\mathbf{r}(n)$, we know that $\mathbf{r}(n)$ can be viewed as the prediction vector of $\mathbf{x}(n)$. To minimize the average output noise variance $\sigma_{bit,\mathbf{A}_0}^2$, the predictor $\mathbf{P}(z)$ should be chosen such that $\det \mathbf{R_r}$ is minimized. This is different from the DPCM system where the predictor is chosen to minimize $tr(\mathbf{R_r})$. The optimization of $\mathbf{P}(z)$ such that $\det \mathbf{R_r}$ is minimized is nonlinear. A suboptimal solution is the conventional predictor where $tr(\mathbf{R_r})$ is minimized.

Performance Comparison of Proposed Coders and DPCM Scheme: In a DPCM scheme where $\mathbf{A}_0 = \mathbf{I}$, the predictor $\mathbf{P}(z)$ is chosen to minimize $tr(\mathbf{R_r})$. In general, bit allocation is not done in a DPCM scheme. That is, $b_i = b$. One can show that the coding gain of the DPCM scheme is given by:

$$\mathcal{CG}_{DPCM} = \frac{\sigma_x^2}{\frac{1}{M}\min\limits_{\mathbf{P}(z)} tr(\mathbf{R_r})}.$$

In the proposed coder as in Fig. 2, the predictor is chosen to minimize $\det \mathbf{R_r}$. The coding gain of the proposed coder is given by:

$$\mathcal{CG}_{new} = \frac{\sigma_x^2}{\min\limits_{\mathbf{P}(z)} \left(\det \mathbf{R_r}\right)^{1/M}}.$$

Comparing the above two equations, we have the ratio:

$$\frac{\mathcal{CG}_{new}}{\mathcal{CG}_{DPCM}} = \frac{\frac{1}{M}\min\limits_{\mathbf{P}(z)} tr(\mathbf{R_r})}{\min\limits_{\mathbf{P}(z)} \left(\det \mathbf{R_r}\right)^{1/M}}.$$

From the Hadamard inequality and the arithmetic-geometric inequality, we know that for any positive semidefinite matrix $\mathbf{R_r}$:

$$\det \mathbf{R_r} \leq \prod_{i=0}^{M-1} [\mathbf{R_r}]_{ii} \leq \left(\frac{1}{M} tr \mathbf{R_r}\right)^M.$$

The first inequality becomes equality if and only if $\mathbf{R_r}$ is a diagonal matrix. The second equality holds if and only if the diagonal entries of $\mathbf{R_r}$ are identical. Therefore we conclude that $\mathcal{CG}_{new} \geq \mathcal{CG}_{DPCM}$, with equality if and only if $\mathbf{R_r}$ is $\sigma_r^2\mathbf{I}$. That is, the DPCM and proposed coders have the same coding gain if and only if the elements $r_i(n)$ and $r_j(n)$ are uncorrelated. The additional gain of the proposed schemes in Fig. 2 comes from two modifications. One is the bit allocation which reduces the arithmetic mean, $\frac{1}{M} tr[\mathbf{R_r}]$, to the geometric mean $\prod_{i=0}^{M-1} \left([\mathbf{R_r}]_{ii}\right)^{1/M}$. One is the nonsingular transform $\mathbf{A}_0$ which reduces the product, $\prod_{i=0}^{M-1} [\mathbf{R_r}]_{ii}$, to the determinant $\det \mathbf{R_r}$.

FIR Encoder and FIR Decoder: In the proposed coder, we need to implement the inverse $\mathbf{A}^{-1}(z)$ at the decoder. The inverse $\mathbf{A}^{-1}(z)$ is in general IIR. In some applications, FIR systems might be preferred. Unlike scalar DPCM scheme where either the encoder or decoder has to be IIR, the encoder and decoder of a vector DPCM scheme can *both* be FIR. To achieve this, we can constrain $\mathbf{I} - \mathbf{P}(z)$ in (2) to be unimodular. One way to do this is to assume that

$$\mathbf{I} - \mathbf{P}(z) = \mathbf{D}_0(z)\mathbf{D}_1(z)\ldots\mathbf{D}_N(z), \qquad (5)$$

where $\mathbf{D}(z)$ is the degree-one unimodular matrix given in (1). This is in general a loss of generality as there are unfactorizable unimodular systems. In the special case of one-step prediction, $\mathbf{P}(z)$ becomes an LUT and the assumption of (5) is no loss of generality. To optimize $\mathbf{P}(z)$, the cascade $\mathbf{D}_0(z)\mathbf{D}_1(z)\ldots\mathbf{D}_N(z)$ should be designed such that $\det[\mathbf{R_r}]$ is minimized. In the following, we will provide an example for the comparison of the coding performance of LOT, BOLT [6] and LUT.

C. Numerical Example

In this example, the vector $\mathbf{x}(n)$ is taken as the blocked version of AR(1) process with correlation coefficients α. The number of channels is $M = 8$. For $0.85 \leq \alpha \leq 0.95$, we optimize the coding gain of (i) subband coding using LOTs and BOLTs [6]; and (ii) unimodular coder using LUTs. All three matrices are first-order with McMillan degree equal to 2. The results are plotted in Fig. 3. From the figure, we see that the LUTs always outperform the BOLTs while the BOLTs always outperform the LOTs. The system delays of the LUTs, BOLTs and LOTs are respectively 7, 23 and 15. We see that the LUTs have the highest coding gain and the smallest system delay when the input is an AR(1) process. The gain is substantial when the correlation coefficient α is close to 1.

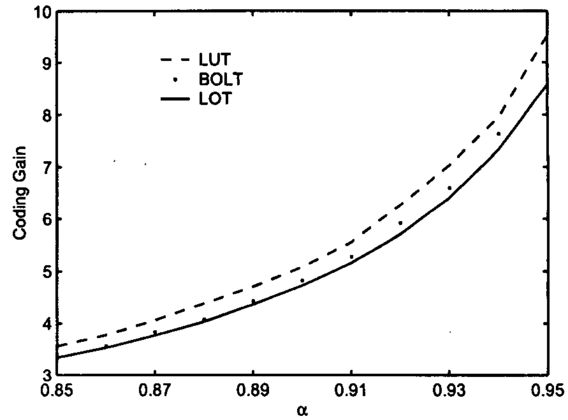

Figure 3: Comparison of coding gain of 8-channel degree-two LUT, BOLT and LOT for AR(1) input with correlation α.

5. REFERENCES

[1] P. P. Vaidyanathan, *Multirate systems and filter banks*, Prentice-Hall, 1993.

[2] K. Nayebi, T. Barnwell, M. Smith, "Design of low delay FIR analysis-synthesis filter bank systems," *Proc. Conf. on Info. Sci. and Sys.*, Mar. 1991.

[3] T. E. Tuncer and T. Q. Nguyen, "General analysis of two-band QMF banks," *IEEE Trans. SP*, Feb 1995.

[4] P. P. Vaidyanathan, "How to capture all FIR perfect reconstruction QMF banks with unimodular matrices," in *Proc. IEEE ISCAS*, May 1990.

[5] S. M. Phoong, and Y. P. Lin, "Minimal factorization of lapped unimodular transforms," *Proc. IEEE ICASSP*, 2000.

[6] P. P. Vaidyanathan and T. Chen, "Role of Anticausal Inverses in multirate filter banks – Part II: The FIR case, factorizations and biorthogonal lapped transforms," *IEEE Trans. SP*, May 1995.

[7] P. P. Vaidyanathan, "Theory of optimal orthonormal subband coders", *IEEE Trans. on SP*, 1997.

[8] S. M. Phoong, and Y. P. Lin, "Predictive Lower Triangular Transforms," *IEEE Trans. SP*, July 2000.