行政院國家科學委員會專題研究計畫成果報告

利用最佳小波作有損耗及無損耗的影像壓縮 (II)
Universal Lossy and Lossless Image Compression Using Optimal Wavelets (II)
計畫編號：NSC 89-2213-E-002-163
執行期限：89年8月1日至90年7月31日

主持人:馮世邁　　國立台灣大學電信工程研究所
email: smp@cc.ee.ntu.edu.tw

1 中文摘要

這是三年期計畫的第二年.在第一年我們提出一種用於雙次頻帶濾波器組的最小雜訊架構.在本計畫中,我們將此結果推廣到M次頻帶濾波器組.這類最佳編碼器稱爲預測式下三角轉換(PLT). PLT具有與KLT一樣好的編碼增益但其複雜度遠低於KLT,且它還享有許多好處,使它適用於有損耗及無損耗的影像壓縮.

關鍵詞：影像壓縮,小波,編碼器

摘要

This is the second year of a three year project. In the first year of a three year project, we introduce a novel minimum noise structure for 2-channel filter banks. In this report, we generalize the results of the first year to the M-channel case. The resultant optimal transform coder is called prediction based lower triangular transform (PLT). The PLT has the same coding gain as the Karhunen-Loeve transform (KLT). Compared with the KLT, the PLT has a much lower cost. In addition, it enjoys many other advantages that make it an attractive candidates for lossless and lossy image compression.
Keywords: Image compression, wavelet, coder

2 緣由與目的

Transform coding has found many applications in various areas of signal processing and communication [1]. Fig. 1 shows a transform coder implemented using multirate building blocks. It is well known that the optimal unitary transform that yields the maximum coding gain is the KLT. Due to its signal dependence and computational cost, the KLT is often only used as a benchmark for performance comparison. In many applications, suboptimal signal independent transforms like discrete cosine transform (DCT) are often used.

Recently there are interests in applying the KLT to universal transform coding. By estimating the statistics from quantized data, the authors in [2] derive a class of universal transform coders using KLT. No side information is needed because both encoder and decoder can access the quantized data. In [3], the authors introduce a classification based method using KLT. The signal space is divided into a number of classes and a fixed transform is designed for each class. In [4], it was shown that under the assumption that the quantization noise is white, the coding gain of the best non unitary transform cannot be better than that of KLT. In this report, we introduce a class of optimal nonunitary transform that has the same coding performance as KLT. In addition to its excellent coding performance, the new transform has many other features that make it an attractive choice for signal compression. Many results in this paper will be stated without proof. Readers are referred to [5] for details.

3 結果與討論

Consider Fig. 1. Let $\mathbf{x}(n)$ and $\mathbf{y}(n)$ be respectively the input and output vectors of the transform $\mathbf{T}$. Assuming that the input is real and WSS, then the $M \times M$ autocorrelation matrices are related as

$$\mathbf{R}_y = \mathbf{T}\mathbf{R}_x\mathbf{T}^T. \quad (1)$$

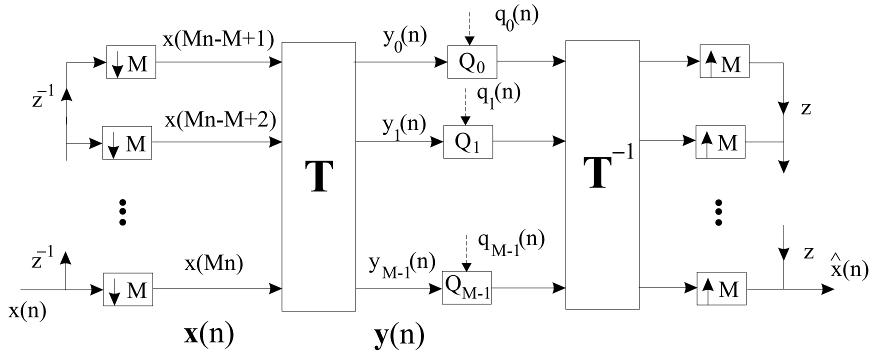

圖 1: An M-dimensional transform coder.

Since $\mathbf{R}_x$ is symmetric, its eigenvectors are orthonormal. By choosing these eigenvectors as the column vectors of $\mathbf{T}$, $\mathbf{R}_y$ is diagonal. In other words, the transform coefficients $y_i(n)$ are uncorrelated. Such a unitary transform $\mathbf{T}$ is the well-known KLT. One can show [1] that KLT maximizes the coding gain of the transform coder. Let σ_x^2 be the input variance and $\sigma_{KLT}^2(k)$ be the variances of $x_k(n)$ for KLT. Then under optimal bit allocation, the coding gain of KLT is given by [1]

$$\mathcal{CG}_{KLT} = \frac{\sigma_x^2}{\prod_{k=0}^{M-1}\left[\sigma_{KLT}^2(k)\right]^{1/M}} = \frac{\sigma_x^2}{[\det \mathbf{R}_x]^{1/M}}. \tag{2}$$

However the KLT is not the only transform that has this decorrelation property. In fact, there exists a lower triangular matrix $\mathbf{T}$ such that the transform coefficients $y_k(n)$ are decorrelated. To see this, we need to use the following lemma from matrix theory [6]:

Lemma 1 ***The LU decomposition of matrices [6]:*** *Let* $\mathbf{A}$ *be an* M *by* M *nonsingular matrix. Suppose that all of its principle submatrices* $\mathbf{A}_K$ *are nonsingular. Then* $\mathbf{A}$ *can be written as*

$$\mathbf{A} = \mathbf{LDU}, \tag{3}$$

in which $\mathbf{L}$ *(respectively* $\mathbf{U}$*) is a lower (respectively upper) triangular matrix with all diagonal entries equal to 1, and* $\mathbf{D}$ *is a diagonal matrix. Moreover the matrices* $\mathbf{L}$, $\mathbf{U}$ *and* $\mathbf{D}$ *are unique. In particular,* $\mathbf{D}$ *is determined by* $\det[\mathbf{D}_K] = \det[\mathbf{A}_K]$, *where* $K = 1, \ldots, N$.

It is not difficult to show that if the matrix $\mathbf{A}$ is symmetric, then the unique matrices $\mathbf{U}$ and $\mathbf{L}$ in Lemma 1 are related as $\mathbf{U} = \mathbf{L}^T$. Applying this fact to (1), we immediately see that there is a unique lower triangular matrix $\mathbf{P}$ such that $\mathbf{P}\mathbf{R}_x(M)\mathbf{P}^T$ is a diagonal matrix. Such a transform $\mathbf{P}$ will be called Prediction-based Lower triangular Transform (PLT). The reason for this name will become clear later. The geometric mean (GM) of the subband variances σ_{PLT}^2 becomes

$$\prod_{k=0}^{M-1}\left[\sigma_{PLT}^2(k)\right]^{1/M} = \det[\mathbf{P}\mathbf{R}_x\mathbf{P}^T]^{1/M} = \det[\mathbf{R}_x]^{1/M}. \tag{4}$$

The GM of $\sigma_{PLT}^2(k)$ is the same as that of $\sigma_{KLT}^2(k)$. However the PLT is not unitary and one can show [5] that the quantization noise $q_i(n)$ in Fig. 1 will be amplified by the inverse PLT at the decoder. Therefore the coding gain of the PLT is less than that of KLT in (2) if the traditional transform coding structure in Fig. 1 is used. To solve this problem, we will introduce a novel minimum noise structure for the PLT in the next section.

Given any input with autocorrelation matrix $\mathbf{R}_x$, PLT can be obtained by using the Gaussian elimination in $\mathcal{O}(M^3)$. However since $\mathbf{R}_x$ is Toeplitz, the computation of PLT can be done in $\mathcal{O}(M^2)$. To see this, let $p_{k,i}$ be the coefficients of $\mathbf{P}$ in the kth row and let

$$P_k(z) = 1 + p_{k,k-1}z^{-1} + \ldots + p_{k,0}z^{-k}, \tag{5}$$

for $k = 1, \ldots, M-1$. If we take $P_k(z)$ as the kth order prediction filter of $x(n)$, then the transform coefficients $y_k(n)$ are the kth order prediction error $e_k(n - M + k + 1)$. Using the orthogonality principle from linear prediction theory [1], one can show that $E\{y_k(n)y_j(n)\} = 0$ for $k \neq j$. From Lemma 1, we know that the lower triangular matrix with such a decorrelation property is unique. Therefore the matrix $\mathbf{P}$ formed by the prediction filter coefficients is the PLT. Hence $\mathbf{P}$ is called the *prediction-based* lower triangular matrix. Using the Levinson-Durbin fast algorithm, all the kth order prediction filters (for $k = 1, \ldots, M-1$) can be obtained in $\mathcal{O}(M^2)$. The kth subband variance for PLT, $\sigma^2_{PLT}(k)$, is equal to the kth order prediction error variance, $\mathcal{E}(k)$.

4 Ladder-based and MINLAB Structures for PLT

Using two different factorization forms of lower triangular matrices, we are able to find two structurally PR implementations using ladder structure for PLT [5]. In this paper, we will discuss only one of the structures and readers are referred to [5] for the other. Note that the lower triangular matrix $\mathbf{P}$ can be decomposed as

$$\mathbf{P} = \mathbf{P}_1\mathbf{P}_2 \ldots \mathbf{P}_{M-1}, \tag{6}$$

where the elementary matrix $\mathbf{P}_k$ is

$$\mathbf{P}_k = \begin{pmatrix} & \mathbf{I}_k & & & \mathbf{0}_{M-k-1\times k} & \\ p_{k,0} & \cdots & p_{k,k-1} & 1 & \cdots & 0 \\ & \mathbf{0}_{k\times M-k-1} & & & \mathbf{I}_{M-k-1} & \end{pmatrix}.$$

Using (6), we have $\mathbf{P}^{-1} = \mathbf{P}^{-1}_{M-1} \cdots \mathbf{P}^{-1}_1$. The inverses of these elementary matrices, $\mathbf{P}^{-1}_k$ can be obtained by replacing the nontrivial elements $p_{k,i}$ in $\mathbf{P}$ with $-p_{k,i}$. Fig. 2 shows the ladder-based implementation obtained from this factorization for the case of $M = 4$. The advantages of the ladder structure are two-fold: (1) In the absence of quantizers, the PLT coder continues to have PR even when all the multipliers $p_{k,i}$ in the structure are quantized to a finite precision. (2) The inverse transform also has $p_{k,i}$ as its multiplier and it can be obtained by inspection.

However, the PLT is a nonunitary matrix, so is its inverse. Hence the PLT coder does not have the energy preservation property. In general, the quantization noise generated in the subbands will be amplified at the decoder. To see this, consider Fig. 2. The inputs to the multipliers $p_{k,i}$ at the encoder are the unquantized data while the inputs to the multipliers $p_{k,i}$ at the decoder are quantized data. That means, the predictors at the encoder use unquantized data as their observations while the predictors at the decoder use the quantized data. It is this mismatch that causes the noise amplication. To avoid the mismatch of observations, one can modify the structure so that the inputs to the multipliers $p_{k,i}$ at the encoder are the quantized data instead of the original unquantized values. The encoder of the modified structure for $M = 4$ case is shown in Fig. 3 and the decoder is the same as Fig. 2. From the figure, one can verify that the structure has the unity noise gain property. This property holds even for correlated and colored quantization noise. The implementation in Fig. 3 will be referred to as MInimum Noise LAdder-based Biorthogonal (MINLAB) structure for PLT. Using the unity noise gain property and the fact that the GM of the subband variances is equal to $\det[\mathbf{R}_x]$, one can prove that the coding gain for the MINLAB PLT coder is the same as KLT.

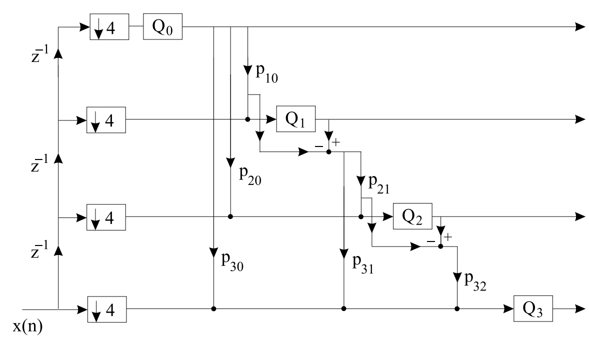

圖 3: The encoder of a PLT MINLAB structure, the decoder is the same as Fig. 2.

The PLT has many other advantages. For example, they can be used to implement a universal transform coders that have the ability to adapt to the input statistics. Lossless transform coders can also be derived from PLT after the simple addition of some rounding operation in the MINLAB structure.

AR(1) Inputs: If the input is an AR(1) process with correlation ρ, then all the prediction error polynomials $P_k(z)$ in (5) will have the same form $(1 - \rho z^{-1})$. The PLT in this case has the following closed form

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho & 1 & 0 & \cdots & 0 \\ 0 & -\rho & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -\rho & 1 \end{pmatrix}. \tag{7}$$

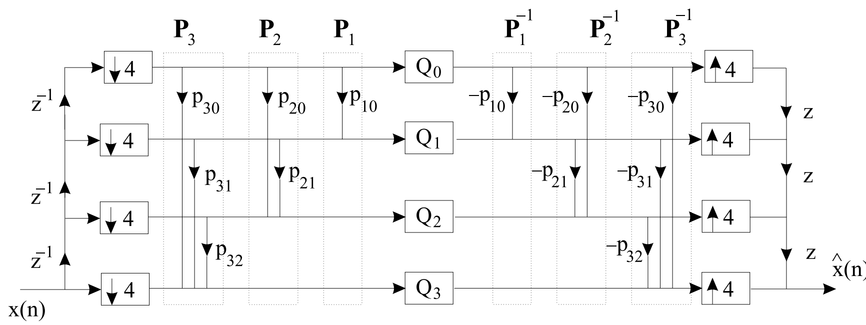

圖 2: Ladder-based implementation of a four-channel PLT coder.

Once ρ is known, we can find $\mathbf{P}$ by inspection, no computation is needed. Therefore the optimal universal PLT coder introduced in Section 4 becomes very simple, we need to estimate only one parameter ρ. The coding gain in this case becomes

$$\mathcal{CG}_{PLT,MIN}(M) = \left[\frac{1}{1-\rho^2}\right]^{\frac{M-1}{M}}. \qquad (8)$$

When M is large, the above gain approaches the coding gain of a DPCM coder. Also note that the transform in (7) is almost independent of the input signal. An $M \times M$ PLT for AR(1) process needs only (M-1) multiplications and additions. Thus its complexity is lower than the DCT which has a complexity of $\mathcal{O}(M \log M)$. Moreover the PLT in (7) is optimal for all AR(1) processes, unlike DCT which is optimal only when ρ approaches 1.

5 計畫成果自評:

The result of this project is very satisfactory. We have proposed a class of low cost optimal transform coders. The following comparison shows that the proposed PLT outperforms the KLT in many aspects:

1. PLT has the same coding performance as the KLT.
2. The design cost of PLT is much lower than that of KLT. The implementational of PLT is less than one half of KLT.
3. Unlike KLT, PLT has a structurally PR implementation using simple building blocks.
4. PLT coders can implement both lossy and lossless compression while KLT in general cannot be used for lossless coding.
5. Unlike KLT, PLT has the simple form in (7) for AR(1) inputs. In this case, the M-dimensional PLT takes only $M-1$ multiplications and additions for each input block. Moreover it is almost signal independent.

參考書目

[1] N. S. Jayant, and P. Noll, *Digital coding of waveforms,* Prentice Hall, 1984.

[2] V. Goyal, J. Zhuang, M. Vetterli, and C. Chan, "Transform coding using adaptive bases and quantization," in *Proc. IEEE Int. Conf. Image Proc.,* pp. 365-368, 1996.

[3] M. Effros, and P. A. Chou, "Weighted universal transform coding: Universal image compression with the Karhunen-Loeve transform," in *Proc. IEEE Int. Conf. Image Proc.,* pp. 61-64, 1995.

[4] P. P. Vaidyanathan, "Theory of optimal orthonormal coder," *IEEE Trans. Signal Proc.,* June 1998.

[5] S. M. Phoong, and Y. P. Lin, "Predition-based lower triangular transform," *Preprint,* National Taiwan University, Oct. 1998 (submitted to IEEE Trans. SP).

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