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 (I)
Universal Lossy and Lossless Image Compression Using Optimal Wavelets (I)
計畫編號：NSC 89-2213-E-002-102
執行期限：88 年 8 月 1 日至 89 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一年. 本計畫提出一種用於梯形非正交小波的最小雜訊架構. 這類最佳MINLAB編碼器的編碼增益永遠大於一. 對於AR(1)及MA(1)訊號,兩個係數的MINLAB編碼器的效能比任何的正交編碼器要好. 此外, MINLAB編碼器的設計簡單及成本低,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
Abstract: In this first year of a three year project, we introduce a novel minimum noise structure for ladder-based biorthogonal wavlets. The coding gain of the proposed optimal MINLAB coder is always greater than unity. For both the AR(1) and MA(1) processes, the MINLAB coder with 2 taps outperforms the optimal orthonormal coders of ***any*** number of taps. Moreover, the optimal biorthogonal coder has a very low design and implementational cost. The proposed coder also enjoys many advantages that make it an attractive choice for lossy/lossless image compression.
Keywords: Image compression, wavelet, coder

2 緣由與目的

Recently there has been considerably interest in applying the ladder structure to data compression [1]–[3]. Fig. 1 shows a simple two-channel filter bank (FB) that uses only one ladder. In the absence of quantizers, the FB has perfect reconstruction, regardless of the choice of $P(z)$. The implementation and design of the biorthogonal system involve only $P(z)$, hence its complexity is very low. Even though the system is simple, its coding performance is comparable to that of orthonormal coders.

On the other hand, the class of orthonormal FB is known to have coding gain $\mathcal{CG} \geq 1$. There has been a lot of interest in finding the optimal orthonormal FB that yields a maximum coding gain for a given input statistics [4], [5]. The theory of optimal orthonormal coder is closely related to the principle component FB and its solution is given in [4]. The optimal FIR case is solved in [5].

In this first year of a three year project, a minimum noise structure is introduced for the ladder-based FBs shown in Fig. 1. The proposed MINLAB coder has the unity noise gain property. The coding gain of the optimal MINLAB coder is equal to the square root of the prediction gain and hence it is guaranteed to be greater than or equal to unity. The optimal biorthogonal coder can be solved using Levinson recursion. For both AR(1) and MA(1) processes, the proposed biorthogonal coder with 2 taps has a higher coding gain than ***any*** optimal orthonormal FB (with any number of taps).

3 結果與討論

Traditional Subband Coder We assume that the quantizers are scalar uniform quantizers and can be modelled as an additive noise source as indicated by the dashed line in Fig. 1. We assume that for a b_i-bit quantizer, the variance of quantization noise $q_i(n)$ satisfies $\sigma_{q_i}^2 = c\,2^{-2b_i}\sigma_{x_i}^2$.

In a traditional subband coder, quantizers Q_i are

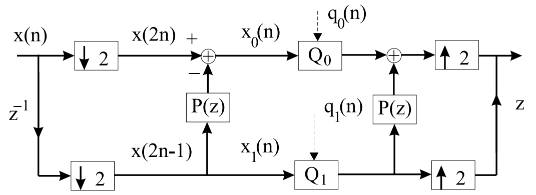

圖 1: Conventional subband coder using ladder.

placed directly after the subband signals $x_i(n)$ as shown in Fig. 1. The output noise $q_{out}(n)$ contains contribution from both $q_0(n)$ and $q_1(n)$. Due to the upsampler, the output noise is not a WSS process. To quantify the error, we use the average noise variance. Assume that $q_1(n)$ is white and uncorrelated with $q_0(n)$. Then one can show that the average noise variance is

$$\sigma^2_{q_{out}} = \frac{1}{2}\sigma^2_{q_0} + \frac{1}{2}\sigma^2_{q_1}(1+E_p),$$

where $E_p = \int_0^{2\pi} |P(e^{j\omega})|^2 \frac{d\omega}{2\pi}$ is the energy of the filter $P(z)$. The noise gain for $q_0(n)$ is unity while $q_1(n)$ is amplified by $1+E_p$.

The MINLAB Coder

In the traditional subband coder shown in Fig. 1, the input to $P(z)$ at the analysis end is $x(2n-1)$, while the input to $P(z)$ at the synthesis end is its quantized version, $\widehat{x}(2n-1)$. That means, in the reconstruction process $q_1(n)$ is added to the top branch through the filter $P(z)$. To avoid this, we can move the quantizer Q_1 to the left, as shown in Fig. 2. This has the dramatic effect of making the noise gain unity. We will refer to Fig. 2 as the ***MInimum Noise LAdder-based Biorthogonal*** (MINLAB) coder. To explain the unity noise gain property of this structure, note that from Fig. 2, we have the following relations:

$$\begin{aligned} x_0(n) &= x(2n) - v_0(n) \\ \widehat{x}_0(n) &= x_0(n) + q_0(n) \\ y_0(n) &= \widehat{x}_0(n) + v_0(n). \end{aligned}$$

From the above equations and Fig. 2, we can conclude that the errors on the top and bottom branches are respectively

$$y_0(n) - x(2n) = q_0(n), \quad y_1(n) - x(2n-1) = q_1(n).$$

Therefore the average variance of output error in the MINLAB coder is given by $\sigma^2_{q_{out}} = 0.5(\sigma^2_{q_0} + \sigma^2_{q_1})$. That means, the noise gain is ***always*** one even though the FB is never orthonormal. Using our noise model and applying the AM-GM inequality to the above equation, we get

$$\sigma^2_{q_{out}} \geq c\, 2^{-2b}\left[\sigma^2_{x_0}\sigma^2_x\right]^{1/2},$$

with equality if and only if the bits are allocated as:

$$b_i = b + \frac{1}{2}\log\sigma^2_{x_i} - \frac{1}{2}\log\left[\sigma^2_{x_0}\sigma^2_x\right]^{1/2}, \qquad (1)$$

where $b = 0.5(b_0+b_1)$ is the average bit rate. From the above derivation, we see that the average output noise variance $\sigma^2_{q_{out}}$ is minimized when the two quantizers have the same noise variance. The noise variances $\sigma^2_{q_i}$ and the quantization stepsize Δ_i are related as $\sigma^2_{q_i} = const * \Delta_i^2$. The MINLAB coder is optimal if the stepsizes of the quantizers are equal. Entropy coding can be applied to further compress the quantizer output. If we define the coding gain of the coder as the ratio of the error variance in PCM over that of the coder, $\sigma^2_{q_{out}}$. Then under the optimal bit allocation (1), the coding gain can be written as:

$$\mathcal{CG} = \frac{\sigma^2_x}{[\sigma^2_{x_0}\sigma^2_x]^{1/2}} = \sqrt{\frac{\sigma^2_x}{\sigma^2_{x_0}}}. \qquad (2)$$

Optimal $P(z)$

From (2), the coding gain $\mathcal{CG}$ is maximized if $\sigma^2_{x_0}$ is minimized. The optimal solution of $P(z)$ such that $\sigma^2_{x_0}$ is minimized can be obtained from the well-known linear prediction theory. To see this, let $P(z)$ be an FIR filter of the form $P(z) = \sum_{n=-N}^{N-1} p(n)z^{-n}$. Then the optimal solution is precisely the optimal predictor of $x(2n)$ based on the observations of $x(2n-2k-1)$, for $-N \leq k < N$. Noncausal predictor can be used here since we are predicting the even samples from the odd samples. A causal implementation of such a system is always possible by inserting enough delays at appropriate places in Fig. 2. Let $x(n)$ be a real-valued wide sense stationary process with autocorrelation coefficients $r(k)$. Then using the orthogonality principle, the optimal $p(n)$ that minimizes $\sigma^2_{x_0}$ is the

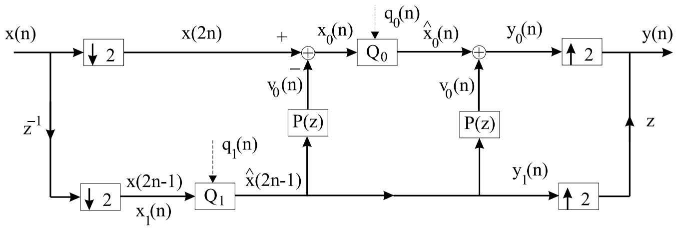

圖 2: The MINLAB coder.

solution of the following equation

$$\mathbf{R}_x \mathbf{p} = \mathbf{r}, \tag{3}$$

where $\mathbf{p} = [p(-N)\ p(-N+1)\ \ldots\ p(N-1)]^T$, $\mathbf{r} = [r(2N-1)\ r(2N-3)\ \ldots\ r(1)\ r(1)\ \ldots\ r(2N-1)]^T$, and $\mathbf{R}_x$ is the autocorrelation matrix of the signal $x(2n-1)$. The above normal equation can be solved in $\mathcal{O}(N^2)$ by using the Levinson fast algorithm. The optimal predictor $\mathbf{p}$ is given by $\mathbf{p}_{opt} = \mathbf{R}_x^{-1}\mathbf{r}$. And the minimum achievable variance $\sigma^2_{x0,min}$ is given by

$$\mathcal{E} = r(0) - \mathbf{r}^T\mathbf{R}_x^{-1}\mathbf{r} = r(0) - \sum_{k=-N}^{N-1} p_{opt}(k) r(2k+1).$$

And the prediction gain is $G_p = \sigma_x^2/\sigma^2_{x0,min} = \sigma_x^2/\mathcal{E} \geq 1$. The prediction gain is unity if and only if all the observations are uncorrelated to the target of prediction $x(2n)$. Using (2), the maximum coding gain of the MINLAB coder is $\mathcal{CG}_{max} = \sqrt{G_p}$.

Linear phase property. The optimal predictor $\mathbf{p}_{opt} = \mathbf{R}_x^{-1}\mathbf{r}$ has linear phase, i.e., $p_{opt}(n) = p_{opt}(-n-1)$. To see this, note that the matrix $\mathbf{R}_x$ satisfies $\mathbf{J}\mathbf{R}_x\mathbf{J} = \mathbf{R}_x$, where $\mathbf{J}$ is the reversal matrix. Since the vector $\mathbf{r}$ is symmetric, we have $\mathbf{J}\mathbf{r} = \mathbf{r}$. Using these properties, we can rewrite (3) as $\mathbf{R}_x(\mathbf{J}\mathbf{p}) = \mathbf{r}$. Comparing this equation and (3), we conclude that $\mathbf{J}\mathbf{p}_{opt} = \mathbf{p}_{opt}$. Hence $P(z)$ has linear phase.

Merits of MINLAB Coder

The MINLAB coder in Fig. 2 enjoys many advantages [6]. In the following, we list some of them:

1. ***Low design and implementational cost:*** The design of the optimal MINLAB coder is simple. Unlike the optimal orthonormal coder, no constrained optimization and no spectral factorization is needed. Optimal MINLAB coder can be obtained by using Levinson algorithm. To implement the analysis or synthesis bank, we need only one filter $P(z)$. Moreover the optimal $P(z)$ has ***linear-phase***. Therefore the complexity of the biorthogonal coder is roughly a quarter of that of an orthonormal coder of the same order.

2. ***Low delay:*** It is known that the delay of an orthonormal coder is proportional to the filter order. The longer the filters are, the larger the system delay is. In the MINLAB coder, if $P(z)$ is a causal filter, then the system delay is only one sample regardless of the filter order. As the prediction gain increases with filter order, so is the coding gain. Therefore we can improve the performance of such a biorthogonal coder without introducing extra system delay.

3. ***Lossy/lossless compression:*** Let the input $x(n)$ be a discrete amplitude signal with stepsize Δ_x. For many applications, the inputs are integers. Suppose the output of $P(z)$ is quantized using a quantizer Q_p. Then the MINLAB coder can be modified for lossless compression as follows:

   (a) Set the stepsize of Q_p be an integer multiple of Δ_x. That is, $\Delta_p = n\Delta_x$ for some integer n. Normally $n = 1$. And any type of quantizer (round off or truncation or ceiling) can be used as Q_p.

   (b) Set the stepsizes of the subband quantizers as $\Delta_0 = \Delta_1 = \Delta_x$. And use entropy coding to encode the outputs of Q_0 and Q_1.

Example 1. AR(1) Inputs. Let the input be an AR(1) process with $r(k)=\rho^{|k|}$ for $0<\rho<1$. We compare the performance of the following various coders:

1. Let $P(z)=p(-1)z+p(0)$. The optimal coding gain $\mathcal{CG}_{MINLAB}(2)=\sqrt{(1+\rho^2)/(1-\rho^2)}$,
2. Take $P(z)=p(0)$. The optimal coding gain is $\mathcal{CG}_{MINLAB}(1)=1/\sqrt{1-\rho^2}$.
3. Consider the coding gain for optimal orthonormal coders with infinite taps and 4 taps. It was shown in [4]–[5] that the coding gains are respectively $\mathcal{CG}_{ortho}(\infty)=\left(\sqrt{1-(16/\pi^2)(\tan^{-1}\rho)^2}\right)^{-1}$ and $\mathcal{CG}_{ortho}(4)=\sqrt{(1+1/3\rho^2)/(1-\rho^2)}$.
4. Consider DPCM of order one. Its coding gain is given by $\mathcal{CG}_{DPCM}(1)=1/(1-\rho^2)$.
5. Suppose we use the traditional biorthogonal coder in Fig. 1. The maximum coding gain for a two-tap filter $P(z)$ is given by $\mathcal{CG}_{tradit}(2)=\sqrt{(1+\rho^2)/(1-\rho^2)}\sqrt{1/(1+E_p)}$, where $E_p=2\rho^2/(1+\rho^2)^2$.

These gains are shown in Fig. 3. It is clear from the figure that $\mathcal{CG}_{DPCM}(1)>\mathcal{CG}_{MINLAB}(2)>\mathcal{CG}_{ortho}(\infty)>\mathcal{CG}_{ortho}(4)>\mathcal{CG}_{MINLAB}(1)$ for all possible ρ. Therefore we see that for AR(1) process, the optimal MINLAB coder with 2 taps (1 multiplier) outperforms the optimal orthonormal coder with infinite number of taps.

Example 2. MA(1) Inputs. Let the input be an MA(1) process with $r(0)=1$, $r(\pm1)=\rho$ for $0<\rho<0.5$, and $r(k)=0$ for all the other k. One can show [6] that there are closed form formulas for the coding gain of the five cases defined in Example 1. All these gains are shown in Fig. 4. The MINLAB coder with 2 taps outperforms all the other coders, including the DPCM and orthonormal coders.

4 計畫成果自評:

The result of this project is very satisfactory. We have successfully derived the optimal ladder-based biorthogonal coders. These biorthogonal coders outperform most orthonormal coders, as demonstrated in the above examples. So far, a conference paper and a transaction paper have been accepted.

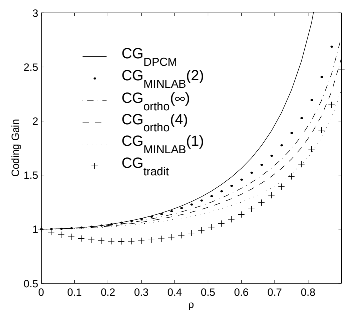

圖 3: Coding gain for AR(1) process.

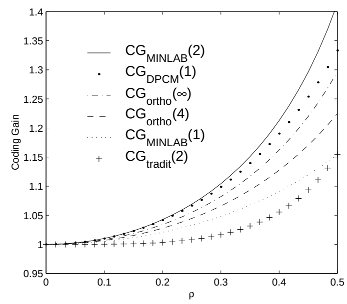

圖 4: Coding gain for MA(1) process.

參考書目

[1] P. Roos, A. Viergever, and M. van Dijke, "Reversible intraframe compression of medical images," *IEEE Trans. Med. Imaging,* Sep. 1988.

[2] A. Said and W. Pearlman, "An image multiresolution representation for lossless and lossy compression," *IEEE Trans. IP,* pp. 1303-11, Sep. 1996.

[3] A. R. Calderbank, I. Daubechies, W. Sweldens, B. Yeo, "Lossless image compression using integer to integer wavelet transforms," in *ICIP,* 1997.

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

[5] A. Kirac and P. P. Vaidyanathan, "Theory and design of optimal FIR compaction filters," *IEEE Trans. SP*, pp. 903-919, April 1998.

[6] S. Phoong and Y. Lin, "MINLAB: Minimum Noise Structure for Ladder-Based Biorthogonal Coders," *IEEE Trans. SP*, 2000.