行政院國家科學委員會專題研究計畫成果報告

計劃名稱：適應性餘弦轉換影像編碼系統之壓縮率與失真度關係分析
Rate-Distortion Analysis for Adaptive Cosine Transform Image Coding System
計劃編號：NSC 88-2218-E-212-003
執行期限：87 年 08 月 01 日至 88 年 07 月 31 日
主持人：吳家琪, jcwu@aries.dyu.edu.tw
執行機構及單位名稱：大葉大學資訊工程系

一、　中文摘要

本研究探討非靜態 (nonstationary) Gaussian 空間變化 autoregressive (SVAR)程序的傳輸率與失真度函數(RDF, $R(D)$)，並對此一隨空間變化，非靜態二維數位影像傳輸系統，作理論及實驗性的研究。本研究針對非靜態 Gaussian SVAR 程序，先推導其傳輸率與失真度函數。並藉由此傳輸率與失真度的計算，我們在此研究中發展建立起一完整的適應性餘弦轉換 (ADCT) 影像傳輸系統。本計劃以變動式影像方塊 SVAR 非靜態 Gauss-Markov 影像模式，來更精確估算真實世界影像的非靜態統計性。所建立的適應性餘弦轉換(ADCT) 影像編碼系統，是以指定每一個不同尺寸之影像方塊，具有相同的平均失真度，並以變動每一影像方塊的編碼壓縮率，來達成此一固定失真度的要求。系統模擬結果，比較具有 SVAR 影像模式之適應性餘弦轉換影像編碼系統與傳統式具有固定編碼壓縮率的靜態餘弦轉換影像編碼系統。此適應性系統的執行效能，並與提出之傳輸率與失真度的理論值比較。結果明顯證明，應用精確的統計性影像模式及適應性影像編碼系統的設計，將大幅減低影像傳輸所需之資訊傳輸量，並能夠保有影像資料重組的品質。

關鍵詞：傳輸率與失真度函數、變動式影像方塊分割、空間變化 Autoregressive 程序、適應性餘弦轉換

Abstract

Rate-distortion function (RDF, $R(D)$) for nonstationary Gaussian space-varying autoregressive (SVAR) process and two-dimensional space-varying nonstationary digital transform image transmission system are investigated theoretically and experimentally. We have derived the $R(D)$ function for nonstationary SVAR Gaussian processes. The rate-distortion measure then serves as a guide to develop an adaptive transform image transmission system. An adaptive discrete cosine transform (ADCT) image transmission system is established in this study. To accurately estimate the nonstationary statistics of a real-world image, we proposed a variable block size SVAR Gauss-Markov image model. The proposed ADCT image coding system is implemented with the same average distortion designated for each variable size image block and the code rate for each image block allowed to vary in order to meet this fixed distortion criterion. Simulation results are given to compare the reconstructed images coded by the ADCT scheme, based on the SVAR image model with the fixed-rate, stationary cosine transform coding scheme. The theoretical rate-distortion measure serves as an ultimate bound on practically achievable system performance. Overall performance for the transform image transmission system based on an SVAR image model is studied and compared with the theoretical R-D bound. The results demonstrates the efficacy of employing an accurate statistical image model and adaptive image coders design to minimize the overall transmission rate requirements

while preserving the quality of the reconstructed imagery data.

Keywords: Rate-distortion function, Time-Varying Autoregressive (TVAR) process, Variable Blocksize Segmentation, Adaptive Discrete Cosine Transform (ADCT).

二、 緣由與目的

The subject of rate-distortion theory is concerned with the determination of the minimum encoding rate with which data generated by a source of certain type can be compressed via a code from a class of codes, subject to a constraint on the distortion in reconstruction of the data from its encoded form. For fairly simple measures of image degradation, it is possible to specify the functional relationship between the channel transmission rate and the numerical values of the criterion of quality. Thus, denoting the value of distortion as D and the rate as R, we are able to specify the rate-distortion function $R(D)$, and so determine either the necessary rate for a given level of degradation or, conversely, the amount of distortion produced by transmission at a predetermined rate.

In earlier attempts to develop mathematical models for images, image data were considered to be samples from a wide-sense stationary random field. However, the assumption of stationarity produces unsatisfactory results when dealing with real-world images. Almost always the stationarity assumption of a real-world image will hold only over very small regions. Therefore, to mitigate the nonstationary behavior in structure, a given image to be processed is partitioned into smaller, hopefully stationary, regions. In order to explore the nonstationarity of real-world images, we propose a non- stationary image model which consists of:

- Variable blocksize segmentation to partition a given gray-level image into smaller image blocks with nearly homogeneous regions;
- A distinguished Gaussian space-varying autoregressive process representation for each of the image subblocks.

The main purpose of this research is to investigate the nonstationary adaptive cosine transform image coding problem in the context of rate-distortion theory. Due to the mathematical complexity associated with the description of nonstationary processes, we apply the time-varying (or, space-varying in the image processing domain) process to approximate the characterization of nonstationary real-world sources. Therefore, we first investigate the *R(D)* function for the nonstationary time- varying autoregressive (TVAR) Gaussian process. The establishment of the *R(D)* function for the time-varying/space-varying Gaussian process then serve as a guide to develop adaptive image transmission systems, which involve the development of a mathematical space-varying nonstationary image model and design strategies for adaptive cosine transform image coding scheme. Based upon a more accurate digital image model, we can apply an appropriate designed adaptive cosine transform coding algorithm to implement a more efficiency image transmission system.

三、結果與討論

The *R(D)* function for nonstationary Gaussian space-varying autoregressive (SVAR) process is given as:

$$D_\theta = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{0}^{1}\int_{0}^{1} \min\left[\theta, \frac{1}{g(u_1,u_2;\omega_1,\omega_2)}\right] du_1 du_2 d\omega_1 d\omega_2$$

$$R(D_\theta) = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{0}^{1}\int_{0}^{1} \max\left[0, \frac{1}{2}\log\frac{1}{\theta g(u_1,u_2;\omega_1,\omega_2)}\right] du_1 du_2 d\omega_1 d\omega_2$$

which serves as a guide to develop the adaptive space-varying nonstationary image transmission systems. An adaptive discrete cosine transform (ADCT) image transmission system is presented as shown in Fig. 1. For a given test image, x_{ij}, we first conduct image segmentation to obtain the variable blocksize *M(k)* x *M(k)* image subblocks, x^k_{ij}, m, n=1, 2, …, *M(k)*. The variable blocksize segmentation is performed using a quadtree data structure by dividing the perceptually more important regions of an image into smaller size blocks compared to the size of blocks containing lesser amounts of spatial activity. Each image subblock is subtracted from its local mean value, μ_k, to construct the residual image, $\bar{x}^k_{mn}$, which is then described by a 2-D SVAR Gauss-Markov random field, according to:

$$\bar{x}^k_{mn} = -a_{10}(m,n)\bar{x}^k_{m-1,n} - a_{01}(m,n)\bar{x}^k_{m,n-1} + a_{11}(m,n)\bar{x}^k_{m-1,n-1} + \varsigma^k_{mn}$$

We proposed a new adaptive discrete cosine transform (ADCT) image coding system where the adaptivity is achieved by considering the variable blocksize DCT according to the image contents and by modeling each image blocks with a different SVAR process to adapt the quantizer stepsize. The proposed variable blocksize ADCT coding scheme as shown in Fig. 1 is implemented with the same average distortion designated for each block. This constant block distortion designation not only has perceptual advantages, but also allows the rate to vary and thereby adjust to the changing spectral characteristics among the various blocks within the image.

The *MSE* rate-distortion performance results for the proposed adaptive cosine transform coding system with constant distortion assignment are summarized in Table 1. Table 1 presents results for the ADCT system with variable blocksize segmentation *H*=5, where *H* is the quadtree segmentation threshold, and with an 8x8 fixed blocksize segmentation. Table 2 is the MSE rate-distortion results for conventional cosine transform image coding system with fixed rate assignment and fixed blocksize segmentation 16 x16.

Rate-distortion performance curves (*R-D* curves) for both the ADCT system and a conventional stationary-based system have been obtained. For clarity we have plotted the normalized distortion versus the overall rate (*R* in bits/pixel). Results for the ADCT system with variable blocksize segmentation threshold *H*=5 and fixed 8x8 blocksize segmentation compared with the conventional stationary DCT coding scheme with fixed blocksize, 16x16, are illustrated in Fig. 2, for the "LENA" image. We have included the theoretical rate-distortion function *R(D)* of the corresponding 2-D Gauss-Markov source in each figure. In all cases, the best rate-distortion performance was shown to be the proposed adaptive DCT system developed in this study. Both *R-D* curves for the adaptive DCT with variable blocksize *H*=5 system and the ADCT with 8x8 blocksize system, closely approach the rate-distortion theoretical limits. The optimum performance is obtained by employing the adaptive DCT system with constant distortion assignment and fixed 8x8 blocksize segmentation. This has confirmed that for nonstationary source encoding, the *R(D)* function for a nonstationary source should be applied as the optimum information-theoretic bound for the purpose of assessing source coding efficiency. The case of stationary AR modeling for real-world images shows the worst performance which is reasonable since the stationary AR image model is not suitable for real-world image as we had described earlier.

The superior rate-distortion performance of the adaptive DCT system shown in Fig. 2 is due to a more appropriate image modeling as well as an improved adaptive DCT source encoding strategy. The proposed adaptive DCT system benefits from the more uniformly distributed distortion among the blocks than in fixed-rate transform coding systems. In this way the bit rate assignment is more consistent with the global spatial activity present within a given image.

四、參考文獻

[1] G. Alengrin, M. Barlaud, and J. Menez. "Unbiased parameter estimation of nonstationary signals in noise". IEEE Trans. on Acoust., Speech, Signal Processing, Vol. ASSP-34, pp. 1319-1322, October 1986.

[2] M. Barlaud, L. Blanc-Feraud, J. Menez, and P. Mathieu. "Image coding using 2-D space-varying linear predictive filters excited by a 2-D multipulse field". Signal Processing, Vol. 16 (No. 1), pp. 13-19, January 1989.

[3] T. Berger. *Rate Distortion Theory: A Mathematical Basis for Data Compression.* Prentice Hall, Englewood Cliffs, N.J., 1971.

[4] R. Charbonnier, M. Barlaud, G. Alengrin, and J. Menez. "Results on AR-modeling of non-stationary signals". Signal Processing, Vol. 12, pp. 143-151, March 1987.

[5] C.-T. Chen. "Adaptive transform coding via quadtree-based variable blocksize DCT". Proceedings of IEEE ICASSP, Vol. III, pp. 1854-1857, Glasgow, Scotland, May 1989.

[6] D. G. Daut and J.-C. Wu. "Adaptive cosine transform image coding with variable block size and constant block distortion". In SPIE Proceedings VCIP'96, pp. 1104-1115, Orlando, FL, March 1996.

[7] Dinstein, K. Rose, and A. Heiman. "Variable block-size transform image coder". IEEE Trans. on Commun., Vol. COM-38, pp. 2073-2078, November 1990.

[8] Grenander and G. Szego. *Toeplitz Forms and Their Applications*. University of California Press, CA, 1958.

[9] Y. Grenier. "Time-dependent ARMA modeling of nonstationary signals". IEEE Trans. On Acoust., Speech, Signal Processing, Vol. ASSP-31, pp. 899-911, August 1983.

[10] G. Mobasseri. "Stochastic quadtree and image segmentation". In Proceedings of SPIE, Vol. 829, pp. 206-213, 1983.

[11] N. Netravali and B. G. Haskell. *Digital Pictures: Representation and Compression.* Plenum, New York, N.Y., 1988.

[12] W. A. Pearlman. "Adaptive cosine transform image coding with constant block distortion". IEEE Trans. on Commun., Vol. COM-38, pp. 698-703, May 1990.

[13] K. R. Rao and P. Yip. *Discrete Cosine Transform: Algorithms, Advantages, and Applications*. Academic Press, San Diego, CA, 1990.

[14] J. Sakrison and B. Algazi. "Comparison of line-by-line and two-dimensional encoding of random images". IEEE Trans. on Inform. Theory, Vol. IT-17, pp. 386-398, July 1971.

[15] E. Shannon. "A mathematical theory of communication". Bell Syst. Tech. Journal, Vol. 27, pp. 379-423 and pp. 623-656, 1948.

[16] P. Strobach. "Quadtree-structured recursive plane decomposition coding of images". IEEE Trans. on Signal Processing, Vol. SP-39, pp. 1380-1397, June 1991.

[17] J. Vaisey and A. Gersho. "Variable block-size image coding". In Proceedings of IEEE ICASSP, pp. 25.1.1-4, Dallas, TX, April 1987.

[18] J. Vaisey and A. Gersho. "Image compression with variable block size segmentation". IEEE Trans. on Signal Processing, Vol. SP-40, pp. 2040-2060, August 1992.

[19] J.-C. Wu and D. G. Daut. "Adaptive nonstationary DPCM image coding with variable blocksize". In SPIE Proceedings VCIP'97, pp. 447-458, San Jose, CA, February 1997.

[20] J.-C. Wu, *Rate distortion Analysis and Digital Transmission of Nonstationary Images*. Ph.D. Thesis, ECE Dept., Rutgers University, Piscataway, New Jersey, December 1997.

Table 1. MSE *R-D* performance for ADCT image transmission system with constant block distortion and variable blocksize for segmentation threshold *H*=5.

"LENA"	Variable blocksize		Fixed 8x8 blocksize	
D_t	R	D/σ^2_x	R	D/σ^2_x
0.1	4.07	0.0016	3.57	0.0018
0.5	2.94	0.0028	2.45	0.0029
2	2.01	0.0054	1.55	0.0055
10	1.10	0.0130	0.79	0.0132
15	0.92	0.0165	0.65	0.0165
20	0.79	0.0192	0.57	0.0193

Table 2. MSE *R-D* performance for conventional, stationary cosine transform image transmission system with fixed rate and fixed blocksize of 16x16.

	D/σ^2_x ("LENA" Image)	
R	*Real world image data*	AR image model design

3	0.0082	0.0299
2	0.0152	0.0368
1.1	0.0275	0.0512
0.79	0.0332	0.0610

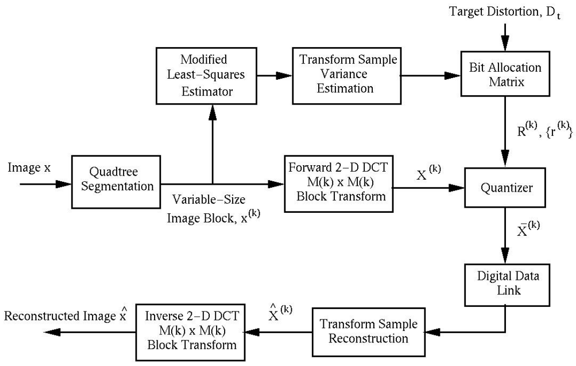

Figure 1. Block diagram of 2-D variable blocksize ADCT image transmission system.

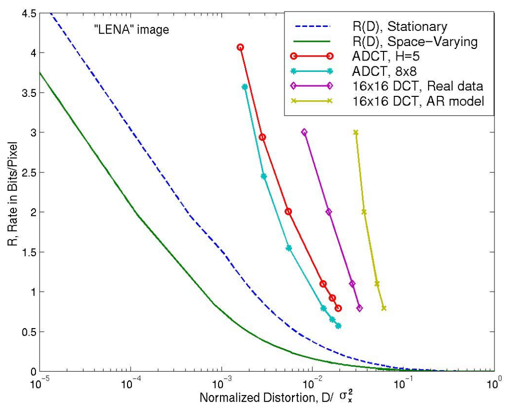

Figure 2. *R-D* performance of 2-D ADCT coding system versus 16x16conventional DCT system, "LENA" image.