A SOFT DECODER VECTOR QUANTIZER FOR A RAYLEIGH FADING CHANNEL - APPLICATION TO IMAGE TRANSMISSION

Mikael Skoglund

Department of Information Theory, Chalmers University of Technology
S - 412 96 Göteborg, Sweden; email: skoglund@it.chalmers.se

ABSTRACT

A Hadamard-based framework for soft decoding in vector quantization over a Rayleigh fading channel is presented. We also provide an efficient algorithm for decoding calculations. The system has relatively low complexity, and gives low transmission rate since no redundant channel coding is used. Our image coding simulations indicate that the soft decoder outperforms its hard decoding counterpart. The relative gain is larger for bad channels. Simulations also indicate that encoder training for hard decoding suffices to get good results with the soft decoder.

1. INTRODUCTION

The demand for efficient wireless communication is continuously increasing. Since a radio channel imposes bandwidth limitations, we need powerful source coding algorithms to obtain a low transmission rate. According to rate-distortion theory, the most efficient way of compressing a source is to use Vector Quantization (VQ) with as high dimension as possible [1]. Increasingly complex VQ-based speech and image coding algorithms are permitted as computer technology advances, and a vector quantizer is becoming a standard tool in many practical systems. However, efficient data-compression gives sensitivity to channel errors, since error protecting redundancy is removed from the source signal. Furthermore, radio channels are poor, as they suffer from both multipath fading and additive noise. Consequently, some kind of error protection is necessary in wireless communication. Traditionally an error protection code has been used in tandem with a source code, and these codes have been optimized separately. This scheme would be optimal if an arbitrary large delay could be accepted. But this is of course not the case in a practical situation. This has led to an increasing interest in combined source-channel coding in recent research, since the combined approach gives good performance at moderate delay and complexity. In combined source-channel coding the VQ is designed to minimize the total distortion from the quantization *and* the channel errors, and no explicit channel code is used [2, 3].

Early work on VQ for combined source-channel coding assumed simple channel models, such as the binary symmetric channel (e.g. [2, 3]). Recent research in this area treats more complex, and realistic, channel models (e.g. [4-6]). General work on VQ for fading channels was presented in [4]. Examples of work on scalar quantization of images and speech over a fading channel can be found in [7] and [8].

In [6] we introduced a Hadamard transform approach to soft decision decoding for vector quantization. The Soft Hadamard Column Decoder (SHCD) of [6] is significantly easier to use than the general formulation of the soft decision decoder utilized in e.g. [5]. The SHCD was shown to be MSE-optimal in case of full encoder entropy. The present paper is based on [6]. New results introduced include the expression for the Hadamard transform based MMSE decoder for a Rayleigh fading channel, and the explicit modification of the SHCD to be optimal for all encoder entropies. We will demonstrate that a significant improvement over hard decisions can be obtained, without any notable increase in complexity. Results for image coding over a Rayleigh fading channel will be provided, but the method presented in this paper is general in nature and many other applications are thinkable.

2. MODELS AND ASSUMPTIONS

Here we present the basic assumptions of this work. We introduce our models and notation in general terms to illustrate the generality of the proposed system. The communication system model under consideration is depicted in figure 2.1, below.

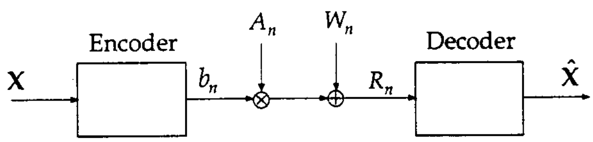

Figure 2.1 Communication system model.

2.1 The VQ Encoder and Decoder

The *encoder* is defined by means of the *encoder regions* $\{S_i\}_0^{N-1}$ which partition $\mathbf{R}^d$ in such a way that

$$\mathbf{X} \in S_i \Rightarrow I = i \tag{2.1}$$

where I denotes the random index chosen by the encoder. The integer i is transmitted to the decoder in binary form. Let $b_n(i) \in \{-1, +1\}$ denote the nth bit in the binary representation of i. The number of bits is $k = \log_2 N$. Also define the *encoder centroids* as $\mathbf{c}_i = E[\mathbf{X}|I = i]$.

The *decoder* is defined by a function $\hat{\mathbf{X}} = \hat{\mathbf{X}}(\mathbf{R})$ of the received channel outputs. This function is chosen such that the *distortion*

$$D = E\left\|\mathbf{X} - \hat{\mathbf{X}}(\mathbf{R})\right\|^2 \tag{2.2}$$

is minimized. Note that this paper treats a *soft decision decoder* in the sense that the decoder utilizes the unquantized channel outputs r_n. This approach was also taken in e.g. [6] and [5].

2.2 The Channel Model

We will restrict ourselves to a channel with binary modulation (e.g. BPSK, QPSK or BFSK), but our framework can be generalized to larger signal sets. We make the common assumptions that the additive noise $\{W_n\}$ is white, and the fading is frequency nonselective and sufficiently slow to allow for coherent detection. We also assume perfect interleaving in the transmission, making the amplitude process $\{A_n\}$ white. The noise is Gaussian with variance σ_w^2 and the amplitude is Rayleigh distributed with parameter σ_A^2. The received signal can be expressed as $R_n = A_n \cdot b_n + W_n$, where R_n denotes the output of the matched filter at the receiver. Let $\mathbf{R}$ denote the vector of received channel outputs corresponding to one source vector.

0-7803-2431-5/95 $4.00 © 1995 IEEE

3. DECODER AND ENCODER EXPRESSIONS

3.1 Optimal Decoder for a Fixed Encoder

The decoder function that minimizes the distortion D, for a fixed encoder, can be written as

$$\hat{\mathbf{X}}(\mathbf{r}) = E[\mathbf{X}|\mathbf{R}=\mathbf{r}] = E[\mathbf{c}_I|\mathbf{R}=\mathbf{r}] \tag{3.1}$$

This expression can be rewritten using a Hadamard transform approach. For this purpose we express the ith encoder centroid as $\mathbf{c}_i = \mathbf{T}\cdot\mathbf{h}_i$, where $\mathbf{h}_i$ is the ith column of an N by N Hadamard matrix $\mathbf{H}$ (c.f. [6]). The matrix $\mathbf{T}$ is fully specified by the encoder centroids. Thus the MMSE decoder can be written

$$\hat{\mathbf{X}}(\mathbf{r}) = \mathbf{T}\cdot\hat{\mathbf{h}}(\mathbf{r}) \tag{3.2}$$

where $\hat{\mathbf{h}}(\mathbf{r}) = E[\mathbf{h}_I|\mathbf{R}=\mathbf{r}]$. This Hadamard formulation is the key to our implementation of the MMSE decoder. Using (3.2) it can be shown that optimal soft decoding can be based on MMSE estimates, $\hat{b}(r_n) = E[b_n(I)|R_n = r_n, \Pr(b_n=+1)=1/2]$, of the individual index bits $b_n(I)$. These estimates will be real numbers in the interval $(-1,+1)$ instead of the "hard" values +1 and −1. More precisely, for the Rayleigh channel we have

$$\hat{b}(r_n) = r_n\cdot\left\{\sigma_W^2\sqrt{2/(\pi s)}\cdot\exp\left(\frac{-sr_n^2}{2\sigma_w^4}\right) + r_n\cdot\left[1-\operatorname{erfc}\left(\sqrt{s/2}\cdot\frac{r_n}{\sigma_W^2}\right)\right]\right\}^{-1}$$

where $s = \sigma_W^2\sigma_A^2/(\sigma_W^2+\sigma_A^2)$. To build a sufficient statistic for the decoding problem, we form every possible multiplicative combination of different estimates in a vector $\hat{\mathbf{p}}(\mathbf{r})$, according to

$$\hat{\mathbf{p}}(\mathbf{r}) = (1, \hat{b}(r_1), \hat{b}(r_2), \hat{b}(r_1)\cdot\hat{b}(r_2), \ldots, \hat{b}(r_1)\cdot\hat{b}(r_2)\cdots\hat{b}(r_k))^T \tag{3.4}$$

This expression is obtained by replacing the "hard bits" $b_n(i)$ in the definition of the Hadamard column $\mathbf{h}_i$ with the "soft" estimates $\hat{b}(r_n)$ (c.f. [6]). It can then be shown that the expression for the MMSE decoder becomes $\hat{\mathbf{X}}(\mathbf{r}) = \mathbf{T}\cdot\hat{\mathbf{h}}(\mathbf{r})$, with

$$\hat{\mathbf{h}}(\mathbf{r}) = f(\mathbf{r})\cdot\mathbf{R}_{\mathbf{hh}}\cdot\hat{\mathbf{p}}(\mathbf{r}) \tag{3.5}$$

where $\mathbf{R}_{\mathbf{hh}} = E[\mathbf{h}_I\mathbf{h}_I^T]$. The scalar function $f(\mathbf{r})$ is defined as $f(\mathbf{r}) = \{\mathbf{m}_{\mathbf{h}}^T\cdot\hat{\mathbf{p}}(\mathbf{r})\}^{-1}$, where $\mathbf{m}_{\mathbf{h}} = E[\mathbf{h}_I]$. We name this implementation of the MMSE decoder the Optimal Soft Hadamard Column Decoder (SHCD-OPT). Since the SHCD-OPT is MSE-optimal for all encoder entropies, we have found an optimal way of doing *joint amplitude estimation and VQ decoding*. In case of full encoder entropy (i.e. $\Pr(I=i) = 1/N, \forall i$), we have $\mathbf{R}_{\mathbf{hh}} = \mathbf{I}$ and $f(\mathbf{r}) = 1$, giving the simpler expression

$$\hat{\mathbf{X}}(\mathbf{r}) = \mathbf{T}\cdot\hat{\mathbf{p}}(\mathbf{r}) \tag{3.6}$$

for the decoder. This simpler form of the decoder is denoted the Soft Hadamard Column Decoder (SHCD) and was studied in [6], under the assumption of no channel fading. Since the SHCD is optimal for a VQ having no redundancy at the encoder output, the SHCD-OPT can be thought of as a modification of the SHCD to account for a priori information, or redundancy, which can be utilized to counteract channel errors.

Based on the recursive nature of the Hadamard matrix, an algorithm for computing $\hat{\mathbf{h}}(\mathbf{r})$ can be derived. We state this algorithm as follows: Let $\{\mu_m^n\}_{n=0}^{N/2^m-1}$ be a sequence of row-vectors of size 2^m. Also, let g_n denote $\hat{b}(r_n)$.

(0). *Initialization*: Set $\mu_0^n = \Pr(I=n)$, $\forall n$, and set m=1.

(1). WHILE $m \le k$, set

$$\mu_m^n = \begin{bmatrix}(1+g_m)\mu_{m-1}^{2n} + (1-g_m)\mu_{m-1}^{2n+1}\\(1+g_m)\mu_{m-1}^{2n} - (1-g_m)\mu_{m-1}^{2n+1}\end{bmatrix}$$

for $n = 0,\ldots,N/2^m-1$, and set $m+1 \to m$.

(2). Now $\hat{\mathbf{h}}(\mathbf{r}) = \mu_k^0\cdot\{\mu_k^0(1)\}^{-1}$, where $\mu_k^0(1)$ denotes the first entry of the vector μ_k^0.

This algorithm require an order of $N\log N$ operations, meaning that the calculation of $\hat{\mathbf{h}}(\mathbf{r})$ is not a heavy task, in the general case. What is then left, for decoding, is a matrix multiplication to obtain $\hat{\mathbf{X}}(\mathbf{r})$.

The Hadamard formulation of the optimal decoder has many advantages. One is the explicit dependence on the bit estimates. Soft information is often available in practical systems, and a modification to the soft decoder will become reasonably straightforward. Another major benefit is the introduction of the transform matrix $\mathbf{T}$. Studies of robust VQ for hard decision channels have illustrated that the complexity can be significantly decreased using a Hadamard approach (e.g. [9]). This reduction is obtained if elements of $\hat{\mathbf{h}}$, and columns of $\mathbf{T}$, are removed, giving a constrained linear mapping formulation of the decoder. The methods of [9] are fully applicable to our soft decoding problem. In this paper, however, we will study the unconstrained mapping.

3.2 Encoder Design for a Fixed Decoder

There are several possible encoder structures. In this paper, we treat two different approaches for the encoder design: The Robust VQ (RVQ) approach, and the Channel Optimized VQ (COVQ) approach [3].

In the *Robust VQ* approach the Voronoi regions of a VQ codebook, trained for a noiseless channel, are used as encoder regions. The encoder is given a good *index assignment* (e.g. [2]) to obtain channel robustness.

In the *channel optimized* approach the encoder/decoder pair is *jointly trained* for a specific channel. The expression for the MSE-optimal encoder regions are given by

$$S_i = \{\mathbf{x} : \operatorname{tr}(\mathbf{R}_i - \mathbf{R}_j) \le 2\mathbf{x}^T(\mathbf{m}_i - \mathbf{m}_j), \forall j\} \tag{3.7}$$

where $\mathbf{R}_i = E[\hat{\mathbf{X}}(\mathbf{R})\cdot\hat{\mathbf{X}}(\mathbf{R})^T|I=i]$ and $\mathbf{m}_i = E[\hat{\mathbf{X}}(\mathbf{R})|I=i]$. Using an approach analogous to the generalized Lloyd algorithm for ordinary VQ training, an encoder/decoder pair can be iteratively designed [3, 6]. For the SHCD-OPT, and the general decoder of e.g. [5], the expectations of (3.7) have to be calculated using Monte Carlo integration. This must be done in each iteration of a joint encoder/decoder design, making encoder design using (3.7) quite complex in practice. It was illustrated in [6] that the training is simplified if full encoder entropy is assumed and the SHCD is used as decoder. Another approach simplifying the training is to use the encoder of a COVQ optimized for hard decisions, with the SHCD-OPT as decoder. In that case the COVQ should be optimized for the expected value of the bit error probability [2]. This latter approach will be compared to the optimal encoder design of (3.7) in section 6.

4. APPLICATION TO CHANNEL ROBUST IMAGE CODING

We will now demonstrate the methods of section 3 for image transmission. Since we have applications in wireless communication in mind, we want to use an image coding algorithm giving a low transmission rate. It is a well known result that scalar quantization in tandem with an entropy code, such as the Huffman code, can perform close to the rate-distortion limit. This scheme is utilized in JPEG-like algorithms. However, entropy coding is preferably avoided for noisy channels, because of error propagation. Hence we use VQ to achieve low bit rate without entropy coding. The method of *vector transform image coding* (VTIC) is a good candidate for our purposes. This method was introduced by Li and Zhang in [10]. We extend the VTIC approach of [10] to combined source-channel coding in the framework of section 3. Other approaches for VQ in image coding can be found in [11].

Since we have a frequency domain coding approach, we can obtain a lower bit rate by using different VQs for different frequency regions. Hence, the frequency domain is divided into three different regions: Low, medium and high frequency. These regions are then coded with different VQs with the highest rate for low frequencies. The frequency domain can be further subdivided to obtain a more efficient utilization of the energy packing of the transform. We use a Greedy approach for the *bit allocation* ([1] p. 234). Combined with a split training ([1] p. 361) this approach gives a joint training and bit allocation. The simple procedure is: (0) Each VQ is initialized with the mean of its training set; (1) allocate one more bit to the VQ giving the highest distortion, by splitting the codevectors of this VQ, and retrain; repeat (1) until all available bits are allocated.

In our evaluations *all* frequency components, including the DC-components, are coded and transmitted. We have chosen to use *overlapping windows* to avoid "blockiness" due to erroneous transmission of low frequency components. The price for the windowing is a higher bit rate. Without windowing we end up at a transmission rate of 0.34 bits per source pixel (bpp). With windowing the equivalent transmission rate becomes 0.43 bpp.

As is usual in VQ design, expectations over the source statistics have to be replaced with averages over a training set. We used a training set consisting of nine 512 by 512 pixel images. The images were stored at 8 bpp (gray scale).

5. RESULTS

In this section we present some results for transmitting images over the Rayleigh fading channel. For PSNR-curves we have used an evaluation set consisting of four 512 by 512 pixel images, and for subjective comparisons we have used the image "Lenna". All test images were outside the training set. The average bit error rate q for hard decisions can be expressed, in terms of the channel SNR (CSNR) $\gamma = \sigma_A^2 / \sigma_W^2$, as

$$q(\gamma) = \frac{1}{2}\left[1 - \sqrt{\frac{\gamma}{1+\gamma}}\right]$$

We will specify channel quality either in terms of CSNR or by the average bit error probability q.

5.1 Robust VQ

Three frequency domain VQs were trained assuming a noiseless channel. The VQs were given good index assignments by means of the FBS algorithm of [12]. Figures 5.1 and 5.2 below illustrate the performance of our system when these VQs were utilized to define the encoder, and the SHCD-OPT was used as decoder.

An evaluation over the set of 4 images is illustrated in figure 5.1. We can see that the gain of the SHCD-OPT is large, on the order of 10 dB in CSNR, for bad channels. Turning to fig. 5.2. and comparing images (3) and (4), we can see that the soft decoding counteracts large errors. The SHCD-OPT also introduces a large general improvement over the hard decoding. Note the entire quality gain is due to the modified decoder, since the same encoder is used in the comparison. And, as we have discussed earlier, this improvement can be obtained without notably higher complexity by using the SHCD-OPT instead of hard decoding. Also note that the transmission rate is the same in both cases, and there is no redundancy of a channel code added.

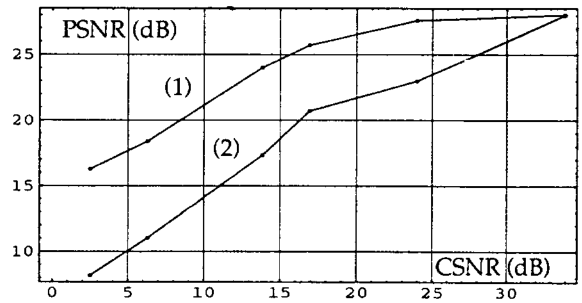

Figure 5.1. *Evaluation of hard decoding versus the SHCD-OPT, with the same VQ-codebook defining the encoder. The upper curve (1) illustrates RVQ encoding and the SHCD-OPT as decoder. The lower curve (2) illustrates RVQ encoding with hard decision decoding.*

Figure 5.2. *Illustration of the subjective performance when using RVQ encoding with hard or soft decoding. Overlapping windows have been used in images (3) and (4). Top left: (1) Original image. Top right: (2) Coded image over a noiseless channel. Rate 0.34 bpp. PSNR 29.1 dB. Bottom left: (3) The SHCD-OPT as decoder. Rate 0.43 bpp. q=0.001. PSNR=26.7 dB. Bottom right: (4) Hard decoding. Rate 0.43 bpp. q=0.001. PSNR=21.8 dB.*

5.2 Channel Optimized VQ

In this section we present results for COVQ-encoding with hard and soft decoding. We have chosen to include an evaluation of COVQ using an encoder optimized for hard decoding. In this way we get an illustration of the effects of the decoding alone, since we use the same encoder for hard and soft decisions. Training of a COVQ with hard decoding is described in e.g. [3]. We also provide a simulation of the SHCD-OPT as decoder with an encoder optimized for soft decoding (the expression for the optimal encoder regions was given in (3.7)). In all our simulations the encoder/decoder pair was trained for the CSNR for which it was evaluated.

Figure 5.3 below illustrates the subjective performance by means of decoded images. In this simulation an encoder optimized for hard decoding was used. The CSNR corresponds to a bit-error rate of q=0.05. In comparing images (1) and (2) of figure 5.3, we can observe the ability of the soft decoding to counteract large errors. A general quality improvement, with the SHCD-OPT, can also be noticed. For a CSNR this low, the hard decisions of (1) make the image quality to bad for most applications. However, the quality of the SHCD-OPT of image (2) is acceptable, and most of the important features of the image are regenerated correctly.

Figure 5.3. *Illustration of the subjective performance for a COVQ encoder with hard and soft decoding. The same encoder was used in both cases. The encoder was trained for hard decoding at the error rate q=0.05. Overlapping windowing has been used in both cases, giving the bit rate 0.43 bpp. Left: (1) Hard decoding, q=0.05, PSNR=23.09 dB. Right: (2) The SHCD-OPT as decoder, q=0.05, PSNR=25.50 dB.*

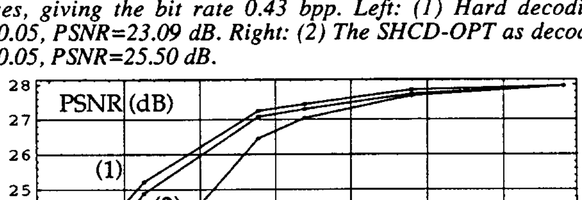

Figure 5.4. *Evaluation of COVQ encoding with hard and soft decoding. The upper curve (1) illustrates the SHCD-OPT as decoder, with an encoder trained for that decoder. The middle curve, (2), illustrates the SHCD-OPT as decoder, with an encoder trained for hard decisions. The lower curve, (3), shows the performance of the same encoder as in (2), but with hard decoding. In all cases the system was designed for the same CSNR as for which it was used.*

Figure 5.4. above illustrates the differences in objective performance between COVQ with hard and soft decoding. Included in this simulation are: (1) An encoder trained for the SHCD-OPT with soft decoding; (2) an encoder trained for hard decisions with the SHCD-OPT as decoder, and; (3) a COVQ with hard decoding. We can see that the relative gain of the soft decoding is smaller than that of figure 5.1. But at small CSNRs the difference becomes significant, a fact that was also illustrated in figure 5.3. For bad channels the gain of the soft decoding is approximately 5 dB of CSNR. Note that the performance of the SHCD-OPT degrades smoothly for low CSNR, while the hard decisions show a quite rapid performance degradation below a CSNR of about 15 dB. Also note that there is no great difference between (1) and (2), suggesting that an encoder trained for hard decisions suffices in most cases. This is a notable result for practical applications, since there is a large gap in complexity between training COVQ encoders for hard decisions and for the SHCD-OPT. Observe that same encoder is used in cases (2) and (3), consequently the difference between (2) and (3) is only due to the different decoders. The proposed scheme of (2) is relatively simple to train, and to use, since the COVQ is easy to train for hard decoding. Thus we have presented a moderate complexity method for COVQ over a Rayleigh channel that performs well for quite severe channel imperfections.

6. CONCLUSIONS

The major conclusion of this work is that we have found a method for combined source-channel image coding over a Rayleigh fading channel, with a low transmission rate since there is no redundancy of a channel code. With this method we obtain an acceptable quality for as low CSNR as that corresponding to a average bit error probability of 5 percent.

7. REFERENCES

[1] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Kluwer Academic Publishers, 1992.

[2] N. Farvardin, "A Study of Vector Quantization for Noisy Channels," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 799-809, 1990.

[3] N. Farvardin and V. Vaishampayan, "On the Performance and Complexity of Channel-Optimized Vector Quantizers," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 155-159, 1991.

[4] F.-H. Liu, P. Ho, and V. Cuperman, "Sequential Reconstruction of Vector Quantized Signals Transmitted Over Rayleigh Fading Channels," in *Proc. ICC '94*, pp. 23-27, New Orleans, 1994.

[5] F.-H. Liu, P. Ho, and V. Cuperman, "Joint Source and Channel Coding Using a Non-Linear Receiver," in *Proc. ICC '93*, pp. 1502-1507, 1993.

[6] M. Skoglund and P. Hedelin, "Vector Quantization over a Noisy Channel using Soft Decision Decoding," in *Proc. ICASSP*, pp. V605-V608, Adelaide, Australia, 1994.

[7] D. G. Daut and J. W. Modestino, "Two-Dimensional DPCM Image Transmission Over Fading Channels," *IEEE Transactions on Communications*, vol. COM-31, no. 3, pp. 315-328, 1983.

[8] D. J. Goodman and C. E. Sundberg, "Combined Source and Channel Coding for Variable Bit Rate Speech Transmission," *Bell Syst. Tech. J.*, vol. 62, no. 7, pp. 2017-2036, 1983.

[9] R. Hagen and P. Hedelin, "Robust Vector Quantization in Spectral Coding," in *Proc. ICCASP*, Minneapolis, Minnesota, 1993.

[10] W. Li and Y.-Q. Zhang, "New Insights and Results on Transform Domain VQ of Images," in *Proc. ICASSP*, pp. V609-V612, 1993.

[11] N. M. Nasrabadi and R. A. King, "Image Coding using Vector Quantization: A Review," *IEEE Transactions on Communications*, vol. 36, pp. 957-971, 1988.

[12] P. Knagenhjelm and E. Agrell, "Hadamard Classes - A Tool for Index Assignment," *Submitted to IEEE Transactions on Information Theory*, December 1993.