

Index assignment in vector quantisation for channels with memory

W.-W. Chang and H.-I. Hsu

Abstract: The study presents a formulation of a Hadamard framework for analysing the vector quantisation over channels with memory. In seeking faster response, classes of index assignments are defined in terms of the Hadamard transform of channel transition probabilities. An index assignment algorithm is developed that achieves high robustness against channel errors, and its performance in vector quantisation of Gauss–Markov sources under noisy channel conditions is illustrated.

1 Introduction

Vector quantisation is widely used in source coding applications [1]. The vector quantiser (VQ) operates by mapping a large set of input vectors into a finite set of representative codevectors. The transmitter transmits the index of the nearest codevector to the receiver, while the receiver decodes the codevector associated with the index and uses it as an approximation of the input vector. However, transmitting VQ data over noisy channels changes the encoded information and consequently leads to severe distortions in the reconstructed output. Although forward error control could be used to protect VQ data, it would be more efficient to mitigate the effects of channel errors without adding redundant bits. This has motivated investigation of ways to reduce the channel distortion by assigning suitable indices to the codevectors of nonredundant VQ systems [2–7].

Finding the best index assignment requires searching among all possible codebook permutations for that which yields the minimum distortion under noisy channel conditions. An exhaustive search is not feasible, and many practical index assignment algorithms are suboptimal. The linearity increasing swap algorithm (LISA) [3] has been shown to be effective for index assignment and may be considered as representing the state of the art. The LISA has certain practical advantages over the binary switching algorithm (BSA) [4] and other algorithms [5, 6]. First, its performance criterion is the linearity index instead of channel distortion, and hence the index assignment can be formulated as a problem of linearising the VQ codebook. Secondly, the Hadamard matrix proves effective in describing the codebook, and use of this framework facilitates the search for the index assignments yielding a high linearity index. However, the usefulness of the LISA may be restricted because it was originally derived for a

memoryless binary symmetric channel (BSC). Transmission errors encountered in most real communication channels exhibit various degrees of statistical dependencies. It is therefore believed that further improvement can be realised through a more precise characterisation of the channel on which index assignment design is based [8].

In this paper, an attempt is made to capitalise more fully on the properties of the Hadamard transform and then to develop a nonredundant VQ system with increased robustness against channel errors. First mathematical tools are developed for use with the Hadamard framework in optimising index assignment. It will be shown that the index assignments can be divided into classes with related features in terms of the Hadamard transform of channel-transition probabilities. This division becomes especially favourable when the complexity of searching the VQ for optimal indices is of primary concern. An index assignment algorithm is then proposed that can effectively reduce the channel distortion by taking into account channel error characteristics.

2 Preliminaries

In this Section, the basic system model and assumptions are presented and an examination is carried out of how the channel distortion depends on the channel transition probabilities.

2.1 Robust VQ

The design of a d -dimensional, m -bit/vector VQ whose output is to be transmitted over noisy channels with memory will be discussed. The VQ encoder searches through the codebook for the codevector c_i that best matches the input vector x , and then transmits the index i to the decoder in binary format. Here, the codebook consisting of $M = 2^m$ codevectors $C = \{c_0, c_1, \dots, c_{M-1}\}$ is designed for a noiseless channel using the generalised Lloyd algorithm [9]. Let $\Phi = \{b(0), b(1), \dots, b(M-1)\}$ denote the set of binary codewords, with $b(i) = (b_{m-1}(i), b_{m-2}(i), \dots, b_0(i))$ being the m -bit binary expansion of an integer index i . Assume that a channel's input $b(i)$ and output $b(j)$ differ by an error pattern $b(e)$, so that the output bit $b_l(j) = b_l(i) \oplus b_l(e)$, $l = 0, 1, \dots, m-1$, where \oplus denotes the bitwise modulo-2 addition. Consequently, the

© IEE, 2002

IEE Proceedings online no. 20020366

DOI: 10.1049/ip-vis:20020366

Paper first received 21st August 2000 and in revised form 7th February 2002

The authors are with the Department of Communications Engineering, National Chiao-Tung University, Hsinchu, Taiwan, Republic of China

decoder looks up the corresponding codevector in its codebook and releases d samples of the codevector c_j , instead of c_i , as the output. Let $\|c_i - c_j\|^2$ represent the squared error distortion and let $P_{j|i}$ represent the probability of receiving the index j given that the index i is transmitted. The overall distortion $E[\|x - c_j\|^2]$ can be viewed as the sum of quantisation distortion $D_q = E[\|x - c_i\|^2]$ and channel distortion $D_c = E[\|c_i - c_j\|^2]$. For this investigation, the index assignment problem is focused on for the purpose of minimising D_c . Most previous work [3–7] regarding analytical expressions for the channel distortion assumes that transmitted VQ indices are equiprobable, i.e. $P(i) = 1/M$, for $i = 0, 1, \dots, M-1$. This assumption applies in particular to large-dimensional VQs that have been optimised for minimum quantisation distortion, according to the asymptotic quantisation theory [10]. In practical applications the assumption of equally likely indices can be a coarse approximation, but the achievable performance gain is still very effective (see Section 5).

2.2 Channel model

To introduce the class of channels investigated, consider first a vector channel described by $\mathbf{b}(j_n) = \mathbf{b}(i_n) \oplus \mathbf{b}(e_n)$, where $\mathbf{b}(e_n)$ represents the error pattern occurring in the n th block of source samples. For channels with finite memory, throughout the paper it is assumed that: (i) the error process is independent of the channel input, (ii) successive error patterns are independent, and (iii) $\mathbf{b}(e_n)$ has intravector memory in the sense that its l th bit $b_l(e_n)$ depends on $\{b_{l-1}(e_n), b_{l-2}(e_n), \dots, b_0(e_n)\}$. Now, taking on assumptions (i)–(iii),

$$P(j_N, j_{N-1}, \dots, j_1 | i_N, i_{N-1}, \dots, i_1) = \prod_{n=1}^N P_{j_n | i_n} \quad (1)$$

where

$$P_{j_n | i_n} = P(\mathbf{b}(e_n)) = \prod_{l=0}^{m-1} P(b_l(e_n) | b_{l-1}(e_n), b_{l-2}(e_n), \dots, b_0(e_n)) \quad (2)$$

The special case of (2) that will be used in the simulation is a two-state Markov-chain model proposed by Gilbert [11]. This model is relatively simple and can characterise a large variety of channels, as evidenced by its applicability to performance analysis of various error-control schemes [12]. The model state-transition diagram is shown in Fig. 1. The Gilbert model consists of an error-free state G and a bad state B , in which errors occur with the probability $(1 - q)$. The state-transition probabilities are P and p for the G to B and B to G transitions, respectively. The effective bit-error rate (BER) produced by the Gilbert channel is $\varepsilon = (1 - q)P/(P + p)$. Its channel-transition probabilities can be formulated as [13]

$$P(\mathbf{b}(e_n)) = \pi \prod_{i=1}^m P_e(b_{m-i}(e_n)) \mathbf{1} \quad (3)$$

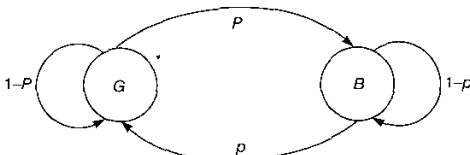


Fig. 1 Gilbert channel model

where the initial state probabilities $\pi = (p/(P + p), P/(P + p))$, $\mathbf{1}$ is a vector of ones, and $P_e(b_{m-i}(e_n))$ in the matrix form

$$P_e(0) = \begin{pmatrix} 1 - P & Pq \\ p & (1 - p)q \end{pmatrix} \\ P_e(1) = \begin{pmatrix} 0 & P(1 - q) \\ 0 & (1 - p)(1 - q) \end{pmatrix} \quad (4)$$

2.3 Channel distortion

The index assignment problem is addressed by using a Hadamard framework for VQ analysis [3]. Begin by defining a Sylvester-style Hadamard matrix $H = (\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_{M-1})$ with elements

$$h_{l,i} = h_{i,l} = (-1)^{W(b(l) \odot b(i))} \quad (5)$$

where \odot denotes the bitwise logical AND operation and W is the Hamming weight function. The codebook construction can be formulated as applying a mapping matrix T on \mathbf{h}_i to produce a codevector c_i ; i.e.

$$c_i = T \cdot \mathbf{h}_i = \sum_{l=0}^{M-1} t_l h_{l,i} \quad 0 \leq i \leq M-1 \quad (6)$$

where the mapping vector t_l denotes column l of $T = (t_0, t_1, \dots, t_{M-1})$.

The effectiveness of index assignment depends crucially on how the channel transition probabilities $P(\mathbf{b}(e))$ are specified in advance. Assuming that all the codewords are equiprobable, the channel distortion is expressed in the Hadamard framework by

$$D_c = \frac{1}{M} \sum_{i=0}^{M-1} \sum_{e=0}^{M-1} \sum_{l=0}^{M-1} P(\mathbf{b}(e)) (h_{i,l} - h_{i \oplus e, l}) t_l^T \cdot (c_i - c_{i \oplus e}) \\ = 2 \sum_{l=1}^{M-1} (1 - Q[\mathbf{b}(l)]) \|t_l\|^2 \quad (7)$$

where $\|t_l\|^2$ is the norm of t_l and

$$Q[\mathbf{b}(l)] = \sum_{e=0}^{M-1} P(\mathbf{b}(e)) (-1)^{W(\mathbf{b}(e) \odot \mathbf{b}(l))} \quad (8)$$

From this, it follows that $Q[\mathbf{b}(l)]$ can be viewed as the scalar Hadamard transform of $P(\mathbf{b}(e))$. A similar formula for D_c was derived in [7], but with the major difference of considering a memoryless BSC. For this specific case, $P(\mathbf{b}(e)) = \varepsilon^k (1 - \varepsilon)^{m-k}$ and $Q[\mathbf{b}(l)] = (1 - 2\varepsilon)^{W[\mathbf{b}(l)]}$, where ε is the channel BER and k is the number of ones occurring in $\mathbf{b}(e)$. Extension of these results to channels with memory requires that $Q[\mathbf{b}(l)]$ be carefully derived to account for the statistical dependencies between error occurrences. In this paper, a more general treatment of the channel distortion is provided by incorporating Gilbert model parameters into a new derivation of $P(\mathbf{b}(e))$ based on (3).

3 Decomposition of index assignments

Finding a good index assignment involves selecting a suitable channel model as well as an optimisation of the mapping matrix T for that specific channel. From the D_c in (7), it can be understood that the Hadamard transforms $Q[\mathbf{b}(l)]$ should be large for high-norm t_l vectors, as they cause large contribution to the channel distortion.

Definition: A binary sequence $\mathbf{b}^{ML}(l) = (b_{a+n}(l), b_{a+n-1}(l), \dots, b_{a+1}(l))$, which satisfies the following two conditions, is called the maximum-length burst of $\mathbf{b}(l)$:

- (i) $b_{a+n}(l) = b_{a+1}(l) = 1$
- (ii) The bits preceding $b_{a+n}(l)$ and following $b_{a+1}(l)$ are all 0s.

Following this, an arbitrary codeword $\mathbf{b}(l)$ can be rewritten in form of $0^r \mathbf{b}^{ML}(l) 0^s$, where 0^r denotes a run of r consecutive 0s. Depending on the pattern of $\mathbf{b}^{ML}(l)$, the set of nonzero indices can be categorised into $M/2$ disjunct classes as follows:

$$\Omega_k = \{l, \text{ such that } \mathbf{b}^{ML}(l) = \mathbf{b}^{ML}(k)\} \\ k = 1, 3, 5, \dots, M-1 \quad (9)$$

Table 1 shows details of the decomposition of index assignments, along with their corresponding maximum-length bursts for a few codebook sizes.

Below two lemmas will demonstrate that efficient computation of $Q[\mathbf{b}(l)]$ can be realised by decomposing the sequence $\mathbf{b}(l)$ into successively smaller subsequences. Begin by observing that $\mathbf{b}(l) = \mathbf{b}^L(l) b_0(l) = b_{m-1}(l) \mathbf{b}^R(l)$, where $\mathbf{b}^L(l) = (b_{m-1}(l), \dots, b_1(l))$ and $\mathbf{b}^R(l) = (b_{m-2}(l), \dots, b_0(l))$.

Lemma 1: When the least significant bit $b_0(l)$ is zero, $\mathbf{b}(l) = \mathbf{b}^L(l) 0$, the M -point Hadamard transform $Q[\mathbf{b}(l)]$ is equal to the $M/2$ -point Hadamard transform $Q[\mathbf{b}^L(l)]$.

Proof of lemma 1: One can consider computing $Q[\mathbf{b}(l)]$ by separating $P(\mathbf{b}(e))$ into two $M/2$ -point sequences consisting of the even-numbered points and odd-numbered points. With substitution of variables $e = 2r$ for e even and $e = 2r + 1$ for e odd, (8) can be written as

$$Q[\mathbf{b}(l)] = \sum_{r=0}^{(M/2)-1} P(\mathbf{b}^R(r) 0) (-1)^{W[(\mathbf{b}^R(r) 0) \odot (\mathbf{b}^L(l) 0)]} \\ + \sum_{r=0}^{(M/2)-1} P(\mathbf{b}^R(r) 1) (-1)^{W[(\mathbf{b}^R(r) 1) \odot (\mathbf{b}^L(l) 0)]} \\ = \sum_{r=0}^{(M/2)-1} P(\mathbf{b}^R(r)) (-1)^{W[(\mathbf{b}^R(r)) \odot (\mathbf{b}^L(l))]} \\ = Q[\mathbf{b}^L(l)] \quad (10)$$

Lemma 2: When the most significant bit $b_{m-1}(0)$ is zero, $\mathbf{b}(l) = \mathbf{b}^L(l) 0$, $Q[\mathbf{b}(l)]$ is equal to the $M/2$ -point Hadamard transform $Q[\mathbf{b}^R(l)]$.

Proof of lemma 2: With substitution of variables $e = r$ for $0 \leq e < M/2$ and $e = (M/2) + r$ for $M/2 \leq e \leq (M-1)$,

$$Q[\mathbf{b}(l)] = \sum_{r=0}^{(M/2)-1} P(0 \mathbf{b}^R(r)) (-1)^{W[(0 \mathbf{b}^R(r) 0) \odot (0 \mathbf{b}^L(l))]} \\ + \sum_{r=0}^{(M/2)-1} P(1 \mathbf{b}^R(r)) (-1)^{W[(1 \mathbf{b}^R(r)) \odot (0 \mathbf{b}^L(l))]} \\ = \sum_{r=0}^{(M/2)-1} P(\mathbf{b}^R(r)) (-1)^{W[(\mathbf{b}^R(r)) \odot (\mathbf{b}^L(l))]} \\ = Q[\mathbf{b}^R(l)] \quad (11)$$

For the more general case of $\mathbf{b}(l) = 0^r \mathbf{b}^{ML}(l) 0^s$, it is appropriate to proceed by decomposing the $M/2$ -point transforms in (10) and (11) into $M/4$ -point transforms and continue until left with only $Q[\mathbf{b}^{ML}(l)]$. This leads to the following theorem.

Theorem 1: The Hadamard transforms $Q[\mathbf{b}(l)]$ are equal for the set of indices having the same maximum-length burst $\mathbf{b}^{ML}(l)$.

This theorem may be interpreted as saying that, regardless of channel characteristics, the set of nonzero indices can be divided into $M/2$ disjunct classes and that all elements within each class Ω_k are assigned one unique $Q[\mathbf{b}(k)]$.

4 Index assignment algorithm

Finding the best index assignment requires searching $M!$ possible combinations of indices for a codebook of size M . This requires enormous computational complexity which, for even small codebooks, may be prohibitive. There is a clear need for techniques which lead to increased efficiency and robustness. For a memoryless BSC, it was observed in [3] that the channel distortion is dominated by the \mathbf{t}_l vectors associated with indices with Hamming weight 1 and hence linearising the VQ leads to minimum-distortion index assignment. However, for channels with memory, one must ensure that the index assignment is such that high $Q[\mathbf{b}(l)]$ corresponds to the \mathbf{t}_l vectors with large norms. To this end, the classes were by descending order of $Q[\mathbf{b}(k)]$ and then the class $\hat{\Omega}$ associated with the largest transform $\hat{\Omega} = \max_{\Omega} Q[\mathbf{b}(k)]$ was identified. To describe how dominant the \mathbf{t}_l vectors with indices $l \in \hat{\Omega}$ in the mapping matrix \mathbf{T} is, a measure

$$\lambda = \frac{1}{\sigma_{VQ}^2} \sum_{l \in \hat{\Omega}} \|\mathbf{t}_l\|^2 \quad (12)$$

was defined, where σ_{VQ}^2 is the sum of the norms $\|\mathbf{t}_l\|^2$; $l = 1, 2, \dots, M-1$. By separating the sum in (7) into one having indices $l \in \hat{\Omega}$ and one for the remaining values of l ,

Table 1: Decomposition of index assignments and their corresponding maximum-length bursts $\mathbf{b}^{ML}(l)$ for the codebook sizes $M = 4, 8, 16$

	Ω_1	Ω_3	Ω_5	Ω_7	Ω_9	Ω_{11}	Ω_{13}	Ω_{15}
$\mathbf{b}^{ML}(l)$	1	11	101	111	1001	1011	1101	1111
$M = 4$	{1, 2}	{3}						
$M = 8$	{1, 2, 4}	{3, 6}	{5}	{7}				
$M = 16$	{1, 2, 4, 8}	{3, 6, 12}	{5, 10}	{7, 14}	{9}	{11}	{13}	{15}

it is easily verified that the channel distortion is bounded by

$$2\sigma_{VQ}^2 \left[(1 - \hat{Q})\lambda + \left(1 - \max_{l \notin \hat{\Omega}} \{Q[\mathbf{b}(l)]\} \right) (1 - \lambda) \right] \leq D_c$$

$$\leq 2\sigma_{VQ}^2 \left[(1 - \hat{Q})\lambda + \left(1 - \min_{l \notin \hat{\Omega}} \{Q[\mathbf{b}(l)]\} \right) (1 - \lambda) \right] \quad (13)$$

The experimental results show consistency of high λ and low D_c , and hence finding a good index assignment may be transformed into a search for the one yielding a high λ .

Next, the effect on the mapping vectors that may result from performing an arbitrary pairwise swap of codevector indices were examined. Suppose that the codevectors \mathbf{c}_K and \mathbf{c}_L were swapped. Then the new \mathbf{t}_l becomes

$$\mathbf{t}'_l = \mathbf{t}_l - \frac{1}{M} \left[(\mathbf{c}_L - \mathbf{c}_K) \cdot ((-1)^{W[\mathbf{b}(L) \odot \mathbf{b}(l)]} - (-1)^{W[\mathbf{b}(K) \odot \mathbf{b}(l)]}) \right] \quad (14)$$

Substituting the norms $\|\mathbf{t}'_l\|^2$ into (12) leads to the new measure

$$\lambda' = \lambda + \frac{1}{\sigma_{VQ}^2 M^2} \sum_{l \in \hat{\Omega}} \left[(\Delta h_l)^2 \|\Delta \mathbf{c}\|^2 - 2M \Delta h_l \mathbf{t}'_l^T \Delta \mathbf{c} \right] \quad (15)$$

where $\Delta \mathbf{c} = \mathbf{c}_L - \mathbf{c}_K$ and $\Delta h_l = h_{L,l} - h_{K,l}$. Knowing that a high λ is favourable for increasing robustness, one may try to increase λ by successively swapping the codevectors so that the \mathbf{t}_l vectors with large norms correspond to the indices $l \in \hat{\Omega}$. Relative aspects of the proposed index assignment algorithm are summarised as follows:

Step 1: Compute the Hadamard transform \mathbf{T} of the codebook \mathbf{C} .

Step 2: Determine the \mathbf{t}_l vectors associated with indices in the class $\hat{\Omega}$.

Step 3: Perform pairwise swaps of the codevectors and, for each swap, calculate the difference $\Delta \lambda = \lambda' - \lambda$. If $\Delta \lambda > 0$, accept the swap and update \mathbf{t}_l ; otherwise, leave the previous assignment intact.

Step 4: Step 3 is repeated until convergence is reached where no increase in λ can result from swaps of any two codevectors.

5 Experimental results

Computer simulations were conducted to compare the performances of three different algorithms: the proposed algorithm (PA), the BSA [4], and the LISA [3]. All three algorithms are based on pairwise swaps to improve a given index assignment, but they employ different objective measures to determine when a local optimum is reached. The objective measure to be optimised for the PA is λ in (12), D_c in (7) for the BSA, and the linearity index for the LISA. The differences between these algorithms for transmitting VQ data over the Gilbert channel deserve comment. First, the LISA was originally derived for a memoryless BSC and hence experiences a channel mismatch between the design and evaluation assumptions. Secondly, the LISA gives an index assignment that is independent of the channel error characteristics. On the other hand, in using the BSA and PA, the channel-transition probabilities have to be combined with *a priori* knowledge of Gilbert model parameters which can be estimated once in advance using the gradient method [14].

Table 2: Channel distortion using various vector quantisers in five different runs of the BSA, LISA, and PA on a Gilbert channel with $\varepsilon = 0.01$

	$M=16, d=4$			$M=64, d=6$		
	BSA	LISA	PA	BSA	LISA	PA
$\rho=0$						
Run 1	0.0594	0.0956	0.0634	0.0919	0.1455	0.0970
Run 2	0.0604	0.1004	0.0632	0.0916	0.1445	0.0982
Run 3	0.0606	0.0999	0.0640	0.0910	0.1446	0.0969
Run 4	0.0585	0.1016	0.0626	0.0913	0.1435	0.0957
Run 5	0.0590	0.0948	0.0610	0.0930	0.1457	0.0966
Average	0.0596	0.0985	0.0628	0.0918	0.1448	0.0969
$\rho=0.5$						
Run 1	0.0675	0.1284	0.0720	0.0948	0.1771	0.0973
Run 2	0.0662	0.1281	0.0660	0.0944	0.1756	0.0979
Run 3	0.0651	0.1270	0.0654	0.0945	0.1790	0.0961
Run 4	0.0658	0.1280	0.0655	0.0945	0.1767	0.0983
Run 5	0.0647	0.1245	0.0659	0.0943	0.1759	0.0999
Average	0.0658	0.1272	0.0666	0.0945	0.1769	0.0979

The source is Gauss-Markov with $\rho=0.0$ and $\rho=0.5$

The input signals considered are Gauss-Markov sources described by $x(n) = \rho x(n-1) + w(n)$, where $w(n)$ is zero-mean, unit-variance white Gaussian noise, with correlation coefficients of $\rho=0$ and $\rho=0.5$. To compare the algorithms with respect to their consistency, each algorithm was run five times, each time beginning with a randomised index assignment. Table 2 presents the vector-quantisation results associated with various algorithms for the case where the bits in the codevector indices are subjected to error sequences typical of the Gilbert channel with effective BER $\varepsilon=0.01$. The performances were measured in terms of D_c for VQ having the following codebook sizes and dimension values (M, d) : (16, 4), (64, 6). The results obtained using the PA and BSA clearly demonstrate an improvement over those obtained using the LISA. Furthermore, the improvement has a tendency to increase for larger codebook sizes and for more heavily correlated Gaussian sources. Fig. 2 is presented to show consistency of high λ and low D_c , in which D_c is plotted as a function

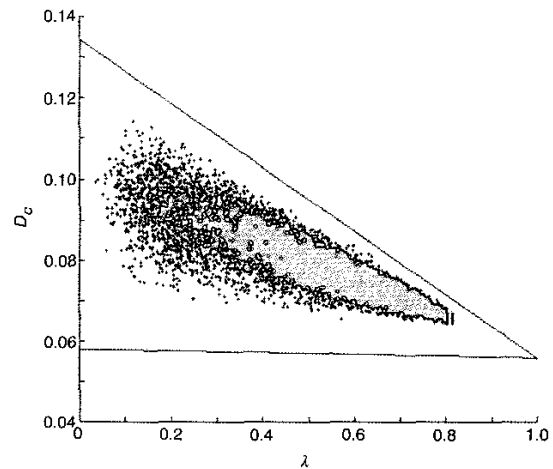


Fig. 2 Scatter plot of channel distortion D_c and its bounds as a function of λ for a 4-bit codebook and a Gilbert channel with $\varepsilon = 0.01$

Table 3: Complexity of the BSA and PA for various codebooks

		$\rho=0$			$\rho=0.5$		
		$d=4$	$d=6$	$d=8$	$d=4$	$d=6$	$d=8$
CPU time (s)	BSA	0.22	3.18	2345.71	0.16	13.24	5483.6
	PA	0	0.27	170.16	0	0.22	117.59
Number of tests	BSA	825	28 224	4 101 930	510	88 704	9 579 840
	PA	240	5952	326 400	240	4960	293 760
Number of accepted swaps	BSA	15	65	670	11	97	932
	PA	13	107	843	11	111	827

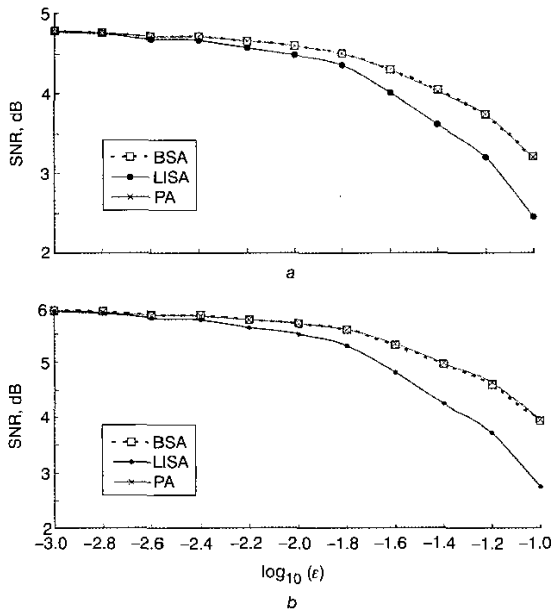


Fig. 3 SNR performance of vector quantisers with various index assignments on a Gilbert channel with effective bit-error rate ϵ
 a $M=256, N=8, \rho=0$
 b $M=256, N=8, \rho=0.5$

of λ , along with their bounds for a 4-bit codebook and a Gilbert channel with $\epsilon=0.01$.

To elaborate further, in Fig. 3 is shown the overall signal-to-noise ratio (SNR) of a VQ having $(M, d)=(256, 8)$ for the Gilbert channel with BER ranging from 10^{-3} to 10^{-1} . Numerical results for the BSA and PA are presented for the case where the running channel parameters agree exactly with those assumed in the design process. It is clear from Fig. 3 that the accuracy of the channel model used in developing the index assignment algorithm is extremely important to the performance of the VQ. A comparison of the BSA and PA also revealed that they yielded comparable performance, with a slight advantage favouring the BSA. However, the better performance of the BSA was achieved at the expense of higher computational complexity. Table 3 shows the complexity of the BSA and PA in terms of three different measures considered in [3]. They are the CPU time measured on a Pentium II-300 PC, the number of pairs that are tested for possible swap, and the number of accepted swaps. It can be seen that, compared with the PA, the BSA consumes much more CPU time and its complexity grows rapidly with increasing codebook size. In Table 4, the performance of the PA is examined under channel mismatch conditions. Here, ϵ_d

Table 4: SNR performance of the PA under channel mismatch conditions

	$\epsilon_d=0.001$	$\epsilon_d=0.01$	$\epsilon_d=0.1$
$\epsilon_a=0.001$	5.84	5.83	5.82
$\epsilon_a=0.01$	5.19	5.2	4.95
$\epsilon_a=0.1$	2.64	2.65	3.95

ϵ_d = design BER, ϵ_a = actual BER; $\rho=0.5, M=256, d=8$

refers to the BER value assumed in the design process, and ϵ_a refers to the true BER. It can be seen that the PA is not very sensitive to channel mismatch, particularly for low ϵ_a values.

6 Conclusions

This paper presents an index assignment algorithm for transmitting VQ data over channels with memory. First the rationale for matching the real channel behaviour with the channel model on which the index assignment design is based is presented. This task was accomplished by using finite-state Markov models to characterise the statistical dependencies in error occurrences. Also, a reduced-complexity algorithm is proposed in which pairwise swaps of VQ codevectors are arranged in accordance with the Hadamard transform of channel-transition probabilities.

7 Acknowledgment

This study was supported by the National Science Council, Taiwan, Republic of China, under contract NSC 90-2213-E-009-110.

8 References

- GRAY, R.M.: 'Vector quantization', *IEEE ASSP Mag.*, 1984, **1**, pp. 4–29
- HEDELIN, P., KNAGENHJELM, P., and SKOGLUND, M.: 'Theory for transmission of vector quantization data' in KLEIJN, W.B., and PALIWAL, K.K. (Eds.): 'Speech coding and synthesis' (Elsevier, Amsterdam, 1995)
- KNAGENHJELM, P., and AGRELL, E.: 'The Hadamard transform — A tool for index assignment', *IEEE Trans. Inf. Theory*, 1996, **42**, (4), pp. 1139–1151
- ZEGER, K., and GERSHO, A.: 'Pseudo-Gray coding', *IEEE Trans. Commun.*, 1990, **38**, pp. 2147–2158
- KNAGENHJELM, N.: 'A study of vector quantization for noisy channels', *IEEE Trans. Inf. Theory*, 1990, **36**, pp. 799–809
- PHAMDO, N., ALAJAJI, F., and FARVARDIN, N.: 'Quantization of memoryless and Gaussian sources over binary Markov channels', *IEEE Trans. Commun.*, 1997, **45**, pp. 668–675
- SKOGLUND, M.: 'On channel-constrained vector quantization and index assignment for discrete memoryless channels', *IEEE Trans. Inf. Theory*, 1999, **45**, (7), pp. 2615–2622

- 8 KANAL, L.N., and SASTRY, A.R.K.: 'Models for channels with memory and their application to error control', *Proc. IEEE*, 1978, **66**, (7), pp. 724-744
- 9 LINDE, Y., BUZO, A., and GRAY, R.M.: 'An algorithm for vector quantizer design', *IEEE Trans. Commun.*, 1980, **28**, pp. 84-95
- 10 GRAY, R.M.: 'Source coding theory' (Kluwer Academic Publishers, Norwell, MA, 1990, 2nd edn.)
- 11 GILBERT, E.N.: 'Capacity of a burst-noise channel', *Bell Syst. Tech. J.*, 1960, **39**, pp. 1253-1265
- 12 DRUKAREV, A.I., and YIU, K.P.: 'Performance of error-correcting codes on channels with memory', *IEEE Trans. Commun.*, 1986, **34**, (6), pp. 513-521
- 13 TURIN, W.: 'Performance analysis of digital transmission systems' (Computer Science Press, New York, 1990)
- 14 CHOUINARD, J.Y., LECOURS, M., and DELISLE, G.Y.: 'Estimation of Gilbert's and Fritchman's models parameters using the gradient method for digital mobile radio channels', *IEEE Trans. Veh. Technol.*, 1988, **37**, (3), pp. 158-166