

An Improved Vector Quantisation Algorithm for Speech Transmission Over Noisy Channels

G.C. Cawley

School of Information Systems,
University of East Anglia, Norwich, U.K.

ABSTRACT

Vector quantisation (VQ) is a method widely used in low bit-rate coding and transmission of speech signals. Unfortunately, a single bit error in the transmitted index, due to noise in the transmission channel, could degrade perceived speech quality at the receiver quite dramatically, as the reference vector retrieved by the corrupted index may differ greatly from the vector corresponding to the intended index. The index assignment (IA) process (an NP-complete combinatorial optimisation problem) attempts to re-order the code book to minimise the effects of single-bit errors, but generally only at considerable computational expense. This paper presents an improved vector quantisation algorithm, based on Kohonen's Self-Organising Feature Map (K-SOFM), that jointly optimises the quantisation error and resistance to noise in the transmission channel. This is achieved using a neighbourhood function based on the Hamming distance between code book indices, rather than the normal Euclidean distance across a two dimensional feature map. As a result, similar reference vectors are recalled by indices with similar binary patterns, minimising the effect of errors in the transmitted index introduced by noise in the transmission channel.

1. INTRODUCTION

Given p -dimensional, unlabelled data $X = \{\mathbf{x}_1, \dots, \mathbf{x}_M\} \subset \mathbb{R}^p$ representative of a data manifold $V \subseteq \mathbb{R}^p$, the process of vector quantisation [1] attempts to partition V into a number of subregions V_i using a finite set of reference or "code book" vectors $W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_N\} \subset \mathbb{R}^p$. such that an input vector $\mathbf{v} \in V_i$ is mapped onto the most similar reference vector \mathbf{w}_i . The values of reference vectors are chosen so as to minimise the quantisation error, measured according to a distance metric, between a training vector \mathbf{x}_i and the best matching reference vector $\mathbf{w}_{(x_i)}$ over all vectors in X . The incoming vector can then be represented approximately by the index i , resulting in a considerable reduction in the required bit rate. Vector quantisation is however very sensitive to errors in the transmitted codeword, due to noise in the transmission channel, as the reference vectors recalled by the intended and corrupted indices might be very different.

The index assignment (IA) process attempts to re-order the code book, such that similar reference vectors are recalled by indices with similar binary patterns [2]. This minimises the effect of an error in the transmitted code word, as the reference vector corresponding to the corrupted index is made as similar as possible to the reference vector represented by the intended index.

2. METHODOLOGY

The Kohonen self-organising feature map (K-SOFM) is a data visualisation algorithm (Kohonen [3, 4]), typically consisting of a two dimensional array of neurons, each of which computes the Euclidean distance between the input vector and the neuron's "weight" vector. During each cycle of the iterative training procedure, an input pattern is chosen at random from the training set and presented to each neuron in turn. The neuron that responds most strongly is declared the *winner*. The weights of the winning neuron are then updated according to the following update rule:

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \alpha(t)[\mathbf{x}(t) - \mathbf{w}_j(t)].$$

In addition, the weights of the neurons in the spatial *neighbourhood* of the winning neuron are also updated according to the same rule. In this way, the network forms a topology preserving map of the input space, where neurons in similar positions in the map represent regions that are close together in the input space.

The Hamming learning vector quantiser [5, 6] (Algorithm 1) is an extension of the Kohonen self-organising feature map, in which the neurons comprising the neighbourhood of the winning neuron are selected according to the Hamming distance between their indices and the index of the winning neuron. This is equivalent to a Kohonen self-organising feature map, with 2^n neurons, arranged as an n -dimensional hypercube with a neuron at each vertex. The resulting feature map is organised such that similar input vectors are mapped onto neurons represented by indices with similar binary patterns. The weight vectors of each neuron represents a reference vector of an ordered codebook, suitable for use in vector quantisation for transmission over a noisy channel.

1. Given initial code book estimate
 $W = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N\} \subset \mathfrak{R}^p$,
 p dimensional training data
 $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\} \subset \mathfrak{R}^p$,
and neighbourhood size η
2. For each $i \in \{1, 2, \dots, M\}$
3. For each $j \in \{1, 2, \dots, N\}$
4. Calculate the output of neuron j for training vector \mathbf{x}_i :

$$\mathbf{o}_j = \|\mathbf{w}_j - \mathbf{x}_i\|$$

$$= \sqrt{\sum_{k=1}^p (w_{jk} - x_{ik})^2}$$
5. Select neuron with the greatest output:
 $\mathbf{o}_{\text{winner}} = \underset{k}{\operatorname{argmax}} (\mathbf{o}_k)$
6. For each $j \in \{1, 2, \dots, N\}$
7. if $d_{\text{Hamming}}(\text{winner}, j) \leq \eta$
 $\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \alpha(\mathbf{x}_i - \mathbf{w}_j(t))$
8. if $\left| \frac{\partial}{\partial t} \frac{1}{M} \sum_{i=0}^M \underset{j}{\operatorname{argmin}} (\mathbf{x}_i - \mathbf{w}_j) \right| \leq \varepsilon$
 $\eta = \eta - 1$
9. if $\eta \geq 0$ return to step 2

Algorithm 1: The Hamming learning vector quantiser.

3. RESULTS

For evaluation, codebooks of 256 reference vectors were compiled, from data obtained from line spectral pair (LSP) [7] analysis of speech from the DARPA/TIMIT [8] speech corpus, using the Hamming learning vector quantiser and a two-dimensional Kohonen self-organising feature map. The reference vectors found using the Hamming learning vector quantiser were also re-ordered randomly, to destroy any correlation between reference vectors and their indices, and also re-ordered using a random pairwise interchange algorithm with simulated annealing [6]. The normalised sample cross correlation of the Hamming distance between code book indices and the Euclidean distance between the corresponding reference vectors \mathbf{w}_i , provides a measure of the order within the codebook.

$$\text{Correlation} = \frac{\overline{he} - \bar{h}\bar{e}}{\sqrt{(\bar{h}^2 - (\bar{h})^2)(\bar{e}^2 - (\bar{e})^2)}}$$

where

$$h_{ij} = \text{hamming distance between } i \text{ and } j$$

$$e_{ij} = \|\mathbf{w}_i - \mathbf{w}_j\|$$

and $\bar{\cdot}$ denotes the mean, for example,

$$\bar{h} = \frac{\sum_{i=1}^n \sum_{j=1}^n h_{ij}}{n^2}$$

$$\bar{he} = \frac{\sum_{i=1}^n \sum_{j=1}^n h_{ij} e_{ij}}{n^2}.$$

Table 1 shows the normalised sample cross-correlation for each codebook. The Hamming learning vector quantiser achieves a substantially higher correlation than the two-dimensional Kohonen self-organising feature map, and also higher than the codebook re-ordered using a random pairwise interchange algorithm with simulated annealing, even though a very lengthy cooling schedule was employed.

Method	Correlation
Random	0.027907
K-SOFM	0.303807
Simulated Annealing	0.320980
HLVQ	0.360483

Table 1: Normalised sample cross correlation for randomly ordered code books, and for codebooks reordered using simulated annealing, HLVQ and K-SOFM

Figure 1 shows a graph of mean error power between reference vectors against the Hamming distance of codebook indices. The graph was created by simulating 1, 2, ... 8 bit errors in every possible bit position, for each reference vector and noting the error power. It can be seen that, as expected, the randomly ordered codebook shows very little change in error power as the Hamming distance increases, so single bit errors in the transmitted index lead to a similar amount of spectral distortion as eight bit errors. Each vector quantisation method, that includes an index assignment process, substantially reduces the error power for single bit errors, which are by far the most common type of transmission error. It can be seen that the error power for single bit errors is lowest for the codebook generated using the Hamming learning vector quantiser.

4. ON THE UNDERLYING DIMENSIONALITY OF THE VECTOR SPACE

To determine the effect of the dimensionality of the data on the correlation in conventional 2-D Kohonen self-organising feature maps and the Hamming learning vector quantiser, data sets were generated from randomly placed Gaussian clusters in 1, 2, 3, ... 10 dimensions. The Kohonen self-organising feature map and Hamming learning vector quantiser algorithms were then used to generate codebooks containing 256 reference vectors for each data set. It can be seen from **Figure 2** that the highest correlation is achieved when the dimensionality of the data space coincides with the dimensionality of the codebook (two-dimensional for the K-SOFM and eight-dimensional for the H-LVQ).

It would be sensible then, to employ a self-organising vector quantiser of the same dimensionality as the underlying dimensionality of the data. Adler *et al.* [9] report the underlying dimensionality of the speech space to be approximately four, i.e. the speech space can be viewed as the projection of a noisy four-manifold into the vector space formed

by the speech parameters. Clearly a four dimensional self-organising feature map would be best suited to vector quantisation of speech signals. The Hamming learning vector quantiser is better suited to vector quantisation of signals where the underlying dimension of the data is not known in advance, or is very high, for instance in vector quantisation image data.

5. CONCLUSIONS

A new combined vector quantisation and index assignment algorithm is presented, the Hamming learning vector quantiser, based on Kohonen's self-organising feature map. The algorithm is applied to the task of robust vector quantisation of speech data for transmission over a noisy transmission channel. The index assignment achieved by the new algorithm is shown to be better than that obtained using a random pairwise interchange algorithm with simulated annealing, or using the conventional Kohonen self-organising feature map. The Hamming self-organising feature map however, due to the high dimensionality of the codebook, is better suited to quantisation of data with a high underlying dimensionality, such as image data.

6. REFERENCES

1. R. M. Gray. Vector quantization. *I.E.E.E. Acoustics, Speech and Signal Processing Magazine*, pages 4–29, April 1984.
2. N. Farvardin and V. Vaishampayan. On the performance and complexity of channel-optimized vector quantizers. *IEEE Transactions on Information Theory*, 37(1):155–160, January 1991.
3. T. Kohonen. *Self-Organization and Associative Memory*, 3rd Edition. Springer Verlag, 1989.
4. T. Kohonen. The self-organizing map. *Proceedings of the I.E.E.E.*, 78(9):1464–1480, 1990.
5. G. C. Cawley and P. D. Noakes. The use of vector quantisation in neural speech synthesis. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN-94)*, volume 3, pages 2227–2230, Nagoya, Japan, 1994.
6. G. C Cawley. *The Application of Neural Networks to Phonetic Modelling*. PhD thesis, Department of Electronic Systems Engineering, University of Essex, 1996.
7. N. Sugamura and F. Itakura. Speech analysis and synthesis methods developed at ECL in NTT -from LPC to LSP-. In *Speech Communication*, volume 5, pages 199–215, 1986.
8. The TIMIT acoustic-phonetic continuous speech corpus. CD-ROM, October 1990.
9. M. D. Adler, R. Togneri, and Y. Attikiouzel. Dimension of speech space. *IEE Proceedings-I Communication, Speech and Vision*, 138(3):207–214, June 1991.

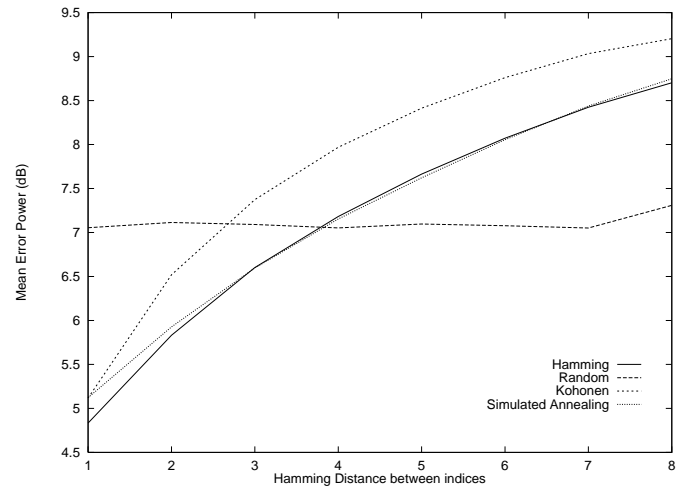


Figure 1: Graph of mean error power (dB) against Hamming distance, for code books generated by a conventional K-SOFM and the Hamming learning vector quantiser, an unordered code book, and a code book re-ordered using simulated annealing.

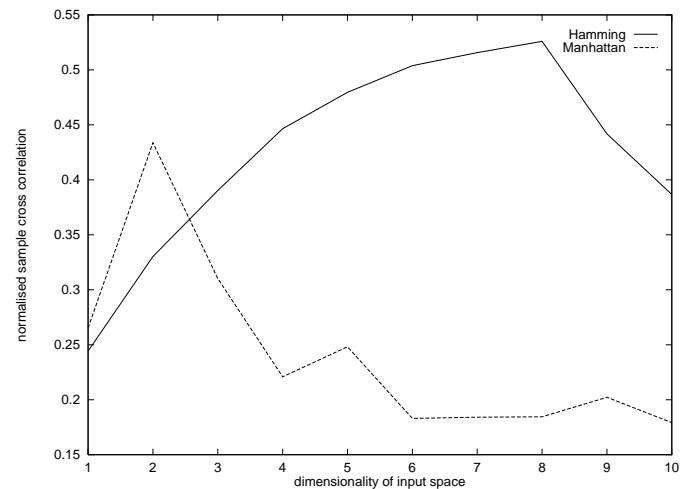


Figure 2: Graph of normalised sample cross-correlation against the intrinsic dimensionality of the training data, for a conventional K-SOFM and Hamming learning vector quantiser.