New fast and robust stochastic algorithm of two stage vector quantization for joint source-channel speech coding for any transmission channel

M. S. García V.**, A. A. Ramírez A.*
* PILIMTEC R&D., 14 Blvd du Château, 35410 Châteaugiron, France
** CITEDI-IPN., Av. Del Parque 1340, Mesa de Otay, 25000 Tijuana, México
E-mail: mgarciav@citedi.mx, aramirez@pilimtec.fr

Abstract

A new algorithm of two stage vector quantization for joint source-channel speech coding for any transmission channels is presented. The computational complexity is only slightly higher than the most widely used Multi Stage Vector Quantization algorithm (MSVQ). This new algorithm improves the characteristics and the results of a sequential quantizer of two stages. The base of this algorithm is the modification of the well-known GS-RGSKA ε algorithm (Reduced complexity Generalized Stochastic K-means Algorithm -of- Great Speed) for a non-stationary channel. This new algorithm is optimal for the joint construction of two stages. The main features of the proposed algorithm are as follows:

• Due to its stochastic nature it avoids being trapped in poor local minimums.
• Initial codebooks are not needed; the codevectors move away from the gravity center of the training vectors towards their final position.
• Source coding and channel coding are jointly optimized to obtain robust codebooks for different levels of noise in transmission channels.
• The reduction of calculation time is based on geometric considerations and memory management.

This algorithm allows to design the codebook orderly due to its advantageous convergence properties. The results showed that the algorithm needs only 8 to 16% of the number of mathematical operations in comparison with the operations required by others propositions for full search of MSQV, either stationary or non-stationary channels.

1. Introduction

The vector quantization (VQ) [3] is an efficient method for the speech compression. The complexity of the vector quantization is expressed in three levels [8]:

1. The complexity of the learning algorithms. They allow to build the dictionary.
2. The processing time for the encoding of the entrance vectors.
3. The importance of the storage memory.

The vector quantization is designed for real transmissions, where the indexes of information are transmitted through a more or less channel noisy. These indexes they do not hold an intrinsic protection against the errors of transmission that the canal may present. To protect the transmitted indexes, the error-correcting codes [7] can be used; the disadvantage of such a process is that the bit rate increases. As the objective is of diminishing the bit rate, in this article is presented an intrinsic robustness against the transmissions errors it’s given to the technique of vector quantization. The construction of the dictionary is optimized at the same time for the source and for the channel instead of a traditional dictionary. This optimization diminishes the degradation caused by the transmission errors without increasing the bit rate [8].

Some methods of fast encoding as the vector quantization product or the vector quantization in cascade (multi-stages, MSVQ) [4,5,2] diminish simultaneously the complexity of construction of dictionary and the required memory for storage. Here, we proposed a new algorithm of joint construction codebooks in the two stage vector quantization case. The proposed algorithm contributes to the improvement in great measure of the three levels of complexity of vector quantization.
2. Multi stage vector quantization

In the quantizer, the available of $R$ bits to quantify a vector of $p$ components, are distributed on $M$ stages, $R_M$ bits for each stage, where the total resolution of the quantizer is $R_1 + R_2 + \cdots + R_M = R$. Each stage is formed by a dictionary $C_M = \left\{ \rightarrow c_{M,1}, \rightarrow c_{M,2}, \cdots, \rightarrow c_{M,2^M} \right\}$ of $2^M$ codeword of $p$ dimension. The quantization of an entrance vector is formed selecting a representative $\rightarrow c_{M,I}$ of each dictionary $C_M$, the quantized vector is then the sum of these representatives: $\rightarrow x = \rightarrow c_{1,I_1} + \rightarrow c_{2,I_2} + \cdots + \rightarrow c_{M,I_M}$. The description $MSVQ(R_1, \cdots, R_M)$ is used for the dimension of the stages. The $MSVQ$ needs to store $2^R + 2^R + \cdots + 2^R$ codewords instead of $2^R + R + \cdots + R$. The $MSVQ$ reduces considerably the storage cost, and it is also a sub-optimal method, that is to say, the distortion increases with the number of stages for the same bit rate. The figure 1 shows the main diagram of the two stages vector quantization.

![Figure 1. Two stages vector quantization.](image)

2.1 Encoding and sequential construction

In their classic form [1,5], the two dictionaries $C_1$ and $C_2$ are built sequentially and the encoding of an entrance vector is also sequential. An entrance vector $\rightarrow x$ is quantized by looking for its next neighbor $\rightarrow c_{1,I_1}$ for the first dictionary. Then a residual vector $\rightarrow e_1$ is calculated by the relationship $\rightarrow e_1 = \rightarrow x - \rightarrow c_{1,I_1}$. This error vector is quantized by the second quantizer in $\rightarrow c_{2,J_2}$, then $\rightarrow e_2 = \rightarrow c_{2,J_2}$, always using the next neighbor’s rule. To build the second dictionary, the learning sequence is quantized by the first stage, and for each learning vector their residual vector is calculated and keep it. This group of residual vectors is used as learning base for the dictionary of the second stage. The quantization of the vector $\rightarrow x$ is given by the sum of their two representatives: $\hat{\rightarrow x} = \rightarrow c_{1,I_1} + \rightarrow c_{2,I_2}$. Using the sequential search of the quantization, the complexity of the encoder is provided by the number of stored representatives $\left(2^R + 2^R\right)$.

2.2 Joint encoding and joint construction

It will see that the construction and the sequential encoding are sub-optimal [2]. The total error $\rightarrow e_t$ of the quantization to two stages is expressed as:

$$\rightarrow e_t = \rightarrow x - \hat{\rightarrow x} = \rightarrow x - \rightarrow c_{1,I_1} - \rightarrow c_{2,I_2} = \rightarrow e_1 - \rightarrow e_2$$

To find the representatives $\rightarrow c_{1,I_1}$ and $\rightarrow c_{2,I_2}$ that minimize the total distortion, it is necessary to make a search among all the $2^R 2^R = 2^R$ couples of possible indexes. The complexity of such a search is as high as that of a full search of a quantizer of one $R$ bits stage. An intermediary solution has been proposed in [1]. The sequential construction of a quantizer of to two stages is equally sub-optimal, since, when the first stage is built, the second stage, it does not keep in mind. It is carried out the sequential classification of the learning vectors $\rightarrow x_j$ because the load of calculations of the joint classification is too important.

3. Stochastic algorithm of joint encoding

The description of an iterative algorithm to improve to two stages quantizer is given in [2], it is a deterministic method.

Here, we generalize our recent suggested algorithm, GS-RGSKAε [8], to the two stage quantizer. The inverse of the temperature is initialized to the zero value and the supposed value of the error probability from the transmission channel to 0.5. The quantizer obtained by this new construction algorithm is much more balanced and has better characteristics that the one obtained by the deterministic method. Building the
two stages simultaneously and cooling the two
dictionaries after each classification, the algorithm had
ended up being quicker. This technical of joint
construction avoids to store the vectors of error
quantization of the first stage. The obtained
dictionaries are more ordered against transmission
errors. Using the quantizer obtained with the joint
search, the decrease of the distortion is significant.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Codebook Size</th>
<th>Number calculus</th>
<th>Algorithm</th>
<th>Codebook Size</th>
<th>Number calculus</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGSKAεMSVQ(9,8)</td>
<td>512</td>
<td>512</td>
<td>RGSKAεMSVQ(9,8)</td>
<td>256</td>
<td>256</td>
</tr>
<tr>
<td>GS-RGSKAεMSVQ(9,8)</td>
<td>512</td>
<td>78.91</td>
<td>GS-RGSKAεMSVQ(9,8)</td>
<td>256</td>
<td>108.05</td>
</tr>
<tr>
<td>RGSKAεMSVQ(10,7)</td>
<td>1024</td>
<td>1024</td>
<td>RGSKAεMSVQ(10,7)</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>GS-RGSKAεMSVQ(10,7)</td>
<td>1024</td>
<td>130.71</td>
<td>GS-RGSKAεMSVQ(10,7)</td>
<td>128</td>
<td>56.69</td>
</tr>
<tr>
<td>RGSKAεMSVQ(11,6)</td>
<td>2048</td>
<td>2048</td>
<td>RGSKAεMSVQ(11,6)</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>GS-RGSKAεMSVQ(11,6)</td>
<td>2048</td>
<td>219.82</td>
<td>GS-RGSKAεMSVQ(11,6)</td>
<td>64</td>
<td>31.48</td>
</tr>
</tbody>
</table>

Table 1. Average number of distance calculations per vector.

4. Performance

The experimental evaluation of the mentioned
algorithm is made by applying the Cosine Line
Spectrum Pairs (CLSP) [9] coefficients of order 10 for
the Code Excited Linear Prediction (CELP) coder
[10,6] . Codebooks for 1st stage of 2048, 1024 and 512
codewords, and codebooks for 2d stage of 64, 128, 256...
codewords have been generated using a training data base of 10098 CLSP vectors.

Table 1(a) and 1(b) show the average number of distance calculations required per vector for each of the algorithms mentioned previously. The new $GS - RGSKAVQ(R_1, R_2)$ algorithm requires less than 16% of the distance calculations required by the others algorithms based on the full search.

The figures 2(a) and 2(b) illustrate the results of the distortion of different vector quantizer in accordance the transmission error probability of a binary symmetric channel for codebooks 1st stage of 2048, 1024 and 512 codewords, and the codebooks 2d stage of 64, 128 and 256 codewords. The distortion of a codebook is the mean distance between a test vector and its codeword reconstructed by the receiver. We used 989 testing vectors. The distortions of the quantizers $GS - RGSKAVQ(R_1, R_2)$ are small for $\varepsilon \leq 1\%$, which is the case for real applications. The figures 2(c) and 2(d) illustrate the mean distance between the codewords in accordance with the Hamming distance of their indexes.

5. Conclusion

In this paper, we presented a new fast and simple method of a stochastic algorithm design of two stage vector quantization for joint source-channel speech coding for any transmission channels. In this new algorithm the initial codebook does not have any influence on the reached global minimum. The complexity of the $GS - RGSKAVQ(R_1, R_2)$ algorithm is comparable to KMA algorithm, and far less than GKMA, GKMAc or RGSKAc reordering algorithms [8].

It is recommended strongly the $GS - RGSKAVQ(R_1, R_2)$ algorithm in all applications that require a codebook with a small distortion average in noise channels and a high degree of robustness independently of channel error probability variations and initial conditions (as a wireless channel).

We aggregated new foundations for obtaining a new algorithm of joint source-channel speech coding, which generated a balanced codebooks for any transmission channel (noiseless or not). Our suggested algorithm diminishes by 80 to 92% the coding time compared to the full search one and improves significantly the recently proposed algorithms.

10. References


Figure 2(c). Order of the codebooks 1st stage.

Figure 2(d). Order of the codebooks 2d stage.