A Data Modem for GSM Adaptive Multi Rate Voice Channel

M. Boloursaz, A. H. Hadavi, R. Kazemi, Student Member, IEEE and F. Behnia, Senior Member, IEEE

Department of Electrical Engineering
Sharif University of Technology
Tehran, Iran
Email: Boloursaz@ee.sharif.edu

Abstract

This paper considers the problem of digital data transmission through the Global System for Mobile communications (GSM). A data modem is presented that utilizes codebooks of Speech-Like (SL) symbols to transmit data through the GSM Adaptive Multi Rate (AMR) voice codec. Using this codebook of finite alphabet, the continuous vocoder channel is modeled by a Discrete Memory less Channel (DMC). A heuristic optimization algorithm is proposed to select codebook symbols from a database of observed human speech such that the capacity of DMC is maximized. Simulation results show that the proposed data modem achieves higher data rates and lower symbol error rates compared to previously reported results while requiring lower computational complexity for codebook optimization.

1. Introduction

Nowadays, with the spread of mobile networks, voice channels are widely available ubiquitously. Previous research [1], [2], [3] has shown that it is possible to use such channels for data communication in low-rate applications. Although high speed data dedicated channels are currently available, using compressed voice dedicated channels for data communication is still interesting because of wider coverage, better service availability, better QoS, more reliability and robustness.

A data modem of the type described here has been suggested for transmission of encrypted end-to-end secure voice or data over the GSM voice channel in [1], [4]. This modem has also been proposed for real-time and secure transmission of Point-of-Sale (POS) transaction information from the POS terminal to the financial host through the GSM voice channel [5].

Providing data connection over voice dedicated channels is a challenging problem due to unpredictable distortions caused by voice codecs used in such channels to the transmitted waveform as described in [2].

The previously proposed methods for the problem of data communication through compressed voice channels can be classified to three different groups due to the conceptual idea they are based on [6]. These three groups are referred to as "Parameter Mapping", "Codebook Optimization" and "Modulation Optimization" methods. In this research, the "Codebook Optimization" approach is taken because it proved to achieve higher throughputs and lower error rates [3]. This method [2], [3], [6] is to encode the input bit stream onto a codebook of predefined waveforms (symbols) that are optimized for efficient communication through the desired voice codec.

In this research, a heuristic codebook design algorithm is demonstrated to maximize the capacity of this approach from an information theoretic point of view. Simulation results on the GSM Adaptive Multi Rate (AMR) voice codec [7] shows that in comparison with previous works, the proposed algorithm converges to codebooks with higher data rates and lower error rates in less number of iterations.

The rest of this paper is organized as follows. Section 2 outlines the general modem structure and the encoding/decoding process. Section 3 presents the codebook symbol generation process. Section 4 describes the results of numerical simulations. Section 5 describes an optimization method for further performance improvement. Finally, section 6 concludes the paper.

2. General Modem Structure

The codebook $\mathrm{C}_{2^{nR}\times n}$ is a table containing 2^{nR} symbols of n samples each. The encoder and decoder are mappings $\mathrm{E}:\{1:2^{nR}\} \rightarrow S^n$ and $\mathrm{D}: \hat{S}^n \rightarrow \{1:2^{nR}\}$ that map the input data onto the codebook symbols and extract the output data from the received signals respectively. So the transmitter converts the input data bit stream into decimal messages m_i of nR bits each

978-1-4799-2096-9/13/$31.00 ©2013 IEEE

and transmits $S_i^n = \mathrm{E}(m_i)$ through the vocoder channel. On the receiver side, message $\widehat{m}_i$ is detected from the received distorted symbol $\hat{S}_i^n$ by $\widehat{m}_i = \mathrm{D}(\hat{S}_i^n)$. Finally, the estimate of the transmitted bits are extracted from $\widehat{m}_i$. The overall block diagram of this system is shown in Fig. 1. Therefore, the main problem is to devise the codebook C and the mappings E and D for efficient data communication through the desired vocoder.

Previous studies have modeled the total channel distortion caused by long memory, lossy compression and anticausality of the vocoder channel by an effective Gaussian random noise not necessarily centered on zero. LaDue et al. [2] supports this model by stating that $P(\hat{S}^n|S^n)$ proved to have a bell-shaped distribution of finite variance by simulation experiments of modulating statistically sufficient amount of random data on symbols, passing them through the vocoder and observing the output symbol histograms.

Applying this model along with Maximum a Posteriori (MAP) detection, [2] shows that the index i of the transmitted symbol $S_i^n = \mathrm{E}(m_i)$ can be selected by

$$\hat{\imath} = arg\,max_i\left[< \hat{S}^n, S_i^n >\right]. \quad (1)$$

where <,> represents vector dot product. But (1) is true if and only if the assumed noise variance is equal for all codebook symbols, however simulation results show the opposite. Neglecting this fact, the same matched filter detector as in (1) is used in this research.

3. Codebook Optimization Problem

3.1. Objective Function

The general modem structure described above is depicted in Fig. 2. As seen in this figure, the transmitter codebook symbols are represented by a random variable X with alphabets from the set

$$X_i \in \{X_i : X_i = S_i^n = \mathrm{E}(m_i), i = 1, \ldots, 2^{nR}\}. \quad (2)$$

Similarly, the channel output symbols are represented by the random variable Y taking values from the set

$$Y_i \in \{Y_i : Y_i = \hat{S}_i^n = D^{-1}(\widehat{m}_i), i = 1, \ldots, 2^{nR}\}. \quad (3)$$

In fact, applying the modem structure demonstrated in section II and using the above notation, the vocoder channel is modeled by a Discrete Memory less Channel (DMC) (Fig. 2) with its probability distribution $P(X|Y)$ dependent on the choice of codebook symbols.

The capacity for this channel is obtained by

$$C = max_{P(X)}\, I(X;Y) = max_{P(X)}\, \big(H(X) - H(X|Y)\big). \quad (4)$$

Assuming $H(X) = nR$ is constant, maximization of $I(X;Y)$ leads to minimization of $H(X|Y)$. Therefore, the goal of optimization is to find a set of codebook symbols that minimize $H(X|Y)$ or equivalently maximize the DMC channel capacity C'.

The cost function $H(X|Y)$ depends on joint probability distribution of X & Ywhich is derived by simulation experiments on the vocoder channel simulator [7]. This histogram is represented by a $2^{nR} \times 2^{nR}$ matrix P whose elements p_{ij} denote the empirical probability of a transmitted X_i symbol that is decoded as Y_j.

Previous research suggest the empirical Symbol Error Rate (SER) as the cost function for minimization, but clearly minimization of $H(X|Y)$ is more general and leads to minimized SER.

3.2. Search Space

The codebook symbols are chosen from TIMIT which is a data base of human speech.

Shahbazi et al. [3] uses the same TIMIT database as the search space. The optimization process addressed in that work consists of a preprocessing stage followed by a symbol selection procedure. The preprocessing stage itself consists of pitch modification, filtering, power normalization, phase continuity adjustment and removal of similar symbols.

In this work, the search space is further reduced using a similar preprocessing phase consisting of the same pre filtering, power normalization, phase continuity adjustment and similar symbols removal. The pith modification phase has been omitted due to the fact that it was observed to ruin the dynamic spectrum of the modulator output which in turn makes further modifications necessary to bypass the VAD.

Applying this space limitation and the mentioned preprocessing, a symbol set consisting of $K \cong 100 \times 2^{nR}$ symbols each having n samples is acquired.

3.3. Proposed Optimization Algorithm

Before proceeding to the optimization algorithm definition, the term *"most error causing symbol"* should be defined. It is known from the classic information theory that the objective function $H(X|Y)$ can be written as

$$H(X|Y) = -\textstyle\sum_i \sum_j p_{ij} \times \log\left(\frac{p_{ij}}{p_j}\right). \quad (5)$$

In which p_j is the probability of Y_j obtained by

$$p_j = \textstyle\sum_i p_{ij}. \quad (6)$$

Contribution of the $l'th$ codebook symbol to $H(X|Y)$ is denoted by k_l and can be written as:

$$k_l = -\sum_{i=1}^{2^{nR}} p_{il} \times \log\left(\frac{p_{il}}{p_l}\right) + \sum_{j \neq l} p_{lj} \times \log\left(\frac{p_{lj}}{p_j}\right). \quad (7)$$

The index $\hat{l}$ of the *"most error causing symbol"* is given by

$$\hat{l} = arg\,max_l(k_l). \quad (8)$$

The proposed minimization algorithm is an iterative method that starts from an initial random codebook and tries to find a codebook with a reduced $H(X|Y)$ in each iteration. Iterations involve 4 steps as demonstrated below:

Step1: *Find the "most error causing symbol" of the current codebook using its histogram.* This histogram has been calculated in 3^{rd} step of the last iteration and stored in P matrix.

Step2: *Generate N new codebooks by substituting the most error causing symbol of the current codebook obtained in step 1, by N random symbols from the search space.*

Step3: *Estimate the histogram matrix P for N codebooks generated in step2 by simulation experiments on the vocoder channel simulator.*

Step4: *Calculate $H(X|Y)$ for these N codebooks using the histograms estimated in step3. Now, if the child with the least $H(X|Y)$ improves the objective function compared to its parent codebook, it is set as the current codebook, else return to step2. End if the stopping criterion is met, otherwise go to step1.*

The stopping criterion is set such that the algorithm stops if a predefined number of iterations is exceeded or $H(X|Y)$ is below a preset threshold.

4. Simulation Results

To assess the performance of the proposed codebook design algorithm, codebooks with different dimensions providing various Modem Data Rates (MDR) were generated for each Codec Data Rate (CDR) of the AMR voice coder applying the proposed algorithm. The test results along with some codebook and algorithm parameters are reported in Table1. The Modem Data Rate (MDR) is calculated by

$$MDR = 8000 \times R. \quad (9)$$

The reported MAEFR parameter represents the Maximum Achievable Error Free Rate. It can be interpreted as a lower bound on the vocoder channel capacity and is calculated by

$$MAEFR = \frac{8000}{n} \times C'. \quad (10)$$

In the above equation $\frac{8000}{n}$ is the maximum number of symbols transmitted in a second and C' represents the maximum number of bits transmitted on each codebook symbol.

It should be noted that simulation results show that the optimum value of n is 40 samples, hence $n = 40$ for all the results reported in Table 1.

5. Further Performance Improvement

In order to achieve lower Bit Error Rates (BER) by the proposed data modem, symbols with higher misdetection probability ($p_{ij}, i \neq j$) must be mapped to m-ary codes with less hamming distance. The m-ary code assignment is optimized in an iterative process that starts from an initial random code assignment and tries to find an assignment with a reduced BER in each iteration.

The *"code distance matrix"* is a $2^{nR} \times 2^{nR}$ matrix B whose elements b_{ij} denote the hamming distance between binary representations of $i-1$ and $j-1$.

Now suppose that Q is the Hadamard (entrywise) product of the matrices P and B. It can be easily shown that

$$\sum_{i,j} q_{ij} = nR \times BER \quad (11)$$

Therefore, the sum of all elements of Q is taken as the cost function to minimize. Similar to the proposed algorithm for codebook optimization, the $l'th$ symbol's contribution to the cost function is

$$c_l = \sum_{i=1}^{2^{nR}} q_{li} + \sum_{j=1}^{2^{nR}} q_{jl}. \quad (12)$$

Similarly, the *"most BER increasing symbol"* is the symbol with maximum c_l. Each algorithm iteration involves 3 steps as demonstrated below:

Step1: *Find the "most BER increasing symbol" of the current code assignment using Equ. 12.*

Step2: *Generate $2^{nR}-1$ new code assignments by swapping the codes of "most BER increasing symbol" and the other $2^{nR}-1$ codebook symbols.*

Step3: *Calculate the cost function for these $2^{nR}-1$ code assignments. Now, if the swapped assignment with the least cost function improves the objective function compared to its parent assignment, it is set as the current assignment and the algorithm returns to step 1 with this new assignment, else the current assignment does not change and the algorithm returns to step1 but for the next one iteration swaps the code for the "second BER increasing symbol" and so on.*

To prevent the algorithm from being trapped in a local minimum the algorithm starts from 1000 initial random and consequently different assignments.

The BER results of applying this algorithm on codebooks obtained in previous sections are reported in Table2. It should be noted that in this table, the term *"ABIRA"* denotes the Average BER of 1000 Initial Random Assignments.

6. Conclusion

In this paper a data transfer technique for digital data communication through GSM voice channel is presented. Compared to the previous works, a main contribution of the authors is modeling the AMR vocoder channel by a Discrete Memory less Channel (DMC) and setting the capacity of this DMC channel as the objective function for maximization instead of the empirical Symbol Error Rate (SER) that has been optimized in most previous works. The other contribution of the authors is the heuristic algorithm proposed for codebook optimization that proved to achieve better rate and SER performance compared to previous works while having faster convergence rate. Finally, the BER performance of this data modem is further improved using another heuristic optimization method.

7. References

[1] N. Katugampala, S.Villette, and A. M. Kondoz, "Secure voice over GSM and other low bit rate systems," *IEE Secure GSM and Beyond: End to End Security for Mobile Communications,* pp. 3/1- 3/4, London, Feb. 2003.

[2] C. K. LaDue, V. V. Sapozhnykov, and K. S. Fienberg, "A data modem for GSM voice channel," *IEEE Transactions on Vehicular Technology*, vol.57, no.4, pp.2205-2218, July 2008

[3] A. Shahbazi, A. H. Rezaei, A. Sayadiyan, and S. Mosayyebpour, "Data Transmission over GSM Adaptive Multi Rate Voice Channel Using Speech-Like Symbols,". *International Conference on Signal Acquisition and Processing ICSAP*, pp. 63-67, Feb. 2010

[4] N. N. Katugampala, K. T. Al-Naimi, S. Villette, and A. M. Kondoz, "Real time data transmission over GSM voice channel for secure voice and data applications," *Secure Mobile Communications Forum: Exploring the Technical Challenges in Secure GSM and WLAN*, pp. 7/1-7/4, 23 Sept. 2004.

[5] B. Kotnik, Z. Mezgeca, J. Svecko, and A. Chowdhury, "Data transmission over GSM voice channel using digital modulation technique based on autoregressive modeling of speech production" *Digital Signal Processing*, vol.19, no.4, pp. 612-627, July 2009

[6] V. Sapozhnykov and S. Fienberg, "A Low-rate Data Transfer Technique for Compressed Voice Channels," *Journal of Signal Processing Systems*, vol. 68, no. 2, pp. 151-170, 2012

[7] ETSI, Digital Cellular Telecommunications Systems (Phase 2+), Adaptive Multi-Rate (AMR) speech transcoding, ANSI-C code for AMR speech codec, (GSM 06.90 version 7.2.1 Release 1998).

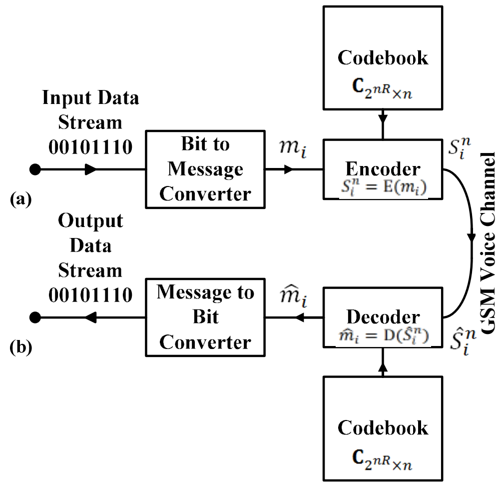

Figure 1. General modem structure. (a) transmitter, (b) receiver

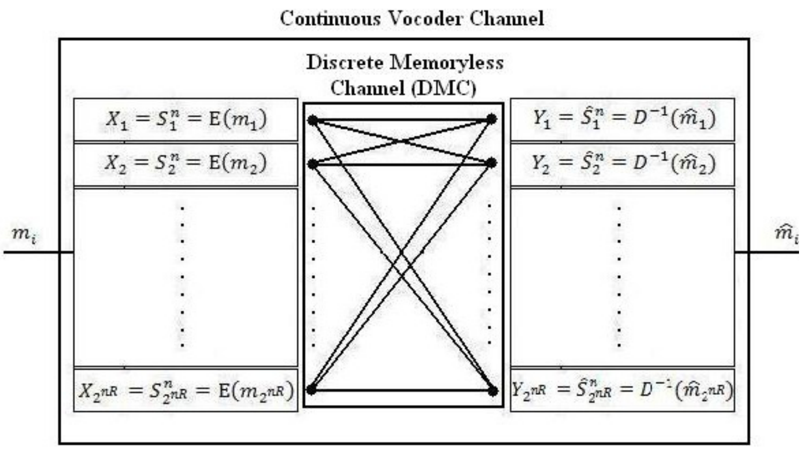

Figure 2. Discrete memoryless model of the vocoder channel

Table 1 Performance of the proposed optimization algorithm over different compression levels of the AMR voice codec

CDR (Kbps)	Codebook Size 2^{nR}	MDR (bps)	SER	MAEFR (bps)	Algorithm Iterations
10.2	64	1200	8.8×10^{-5}	1199.7	110
	32	1000	7.1×10^{-6}	999.9	45
	16	800	6.4×10^{-6}	799.9	20
7.4	64	1200	3.4×10^{-3}	1199.3	110
	32	1000	5.3×10^{-4}	998.6	45
	16	800	2.1×10^{-4}	799.4	20
5.9	64	1200	5.2×10^{-2}	1152.9	135
	32	1000	1.6×10^{-2}	970	100
	16	800	3.4×10^{-3}	792.6	120
4.75	32	1000	6.8×10^{-2}	895.6	85
	16	800	1.7×10^{-2}	769.4	85

Table 2 Performance of the proposed BER optimization algorithm over different compression rates of the AMR voice codec

CDR (Kbps)	Codebook Size 2^{nR}	MDR (bps)	ABIRA	Optimized BER	Number of Iterations
10.2	32	1000	3.6×10^{-6}	1.4×10^{-6}	40
7.4	32	1000	2.8×10^{-4}	1.4×10^{-4}	265
	16	800	1.2×10^{-4}	7.4×10^{-5}	55
6.7	64	1200	1.1×10^{-2}	5.9×10^{-3}	635
	32	1000	5.3×10^{-3}	2.7×10^{-3}	230
	16	800	6.6×10^{-4}	3.8×10^{-5}	45
5.9	64	1200	2.7×10^{-2}	1.6×10^{-2}	695
	32	1000	8.6×10^{-3}	5×10^{-3}	175
	16	800	1.9×10^{-3}	1.2×10^{-3}	45
4.75	32	1000	3.5×10^{-2}	2.2×10^{-2}	150
	16	800	9.5×10^{-3}	6.1×10^{-3}	60