Analog Transmission of Vocoder Features

February 10, 2024

Given a vector of vocoder features $\mathbf{f}$, use an autoencoder E to map them to a dimension d latent vector $\mathbf{z}$ where d is even. The magnitude of each element z_i of $\mathbf{z}$ is constrained to a maximum of 1, but unlike digital modulation is continuously valued and not constrained to a discrete set of points. For bandwidth efficient transmission over the channel the elements of $\mathbf{z}$ are mapped to $d/2$ complex symbols $\mathbf{q}$. Compared to classical digital modulation, the elements of $\mathbf{z}$ can be considered BPSK symbols (continuously valued, analog bits), and the elements of $\mathbf{q}$ analog QPSK symbols.

Our goal is to determine if reasonable speech quality can be obtained over a channel of bandwidth $B < 3000$ Hz and SNR between 0 and 6dB, roughly the lower limit of Single Side Band (SSB) - a common power and bandwidth efficient form of analog radio communication.

1 Noise Simulation

The autoencoder output $\mathbf{z}$ is updated every $T_z = 1/R_z$ seconds, giving a BPSK symbol rate of:

$$R_b = d/T_z \tag{1}$$

For example with $T_z = 0.04, d = 80, R_b = 2000$ symbols/s. The QPSK symbol rate is given by:

$$R_q = \frac{d}{2T_z} \tag{2}$$

For example with $T_z = 0.04, d = 80, R_q = 1000$ symbols/s.

We wish to simulate a channel with a user-defined E_b/N_0, where E_b is the energy of each BPSK symbol, and N_0 is the noise power per unit bandwidth. The energy of each BPSK symbol E_b is the signal power S divided by the symbol rate $R_b = 1/T_b$. The noise per unit bandwidth is the total noise power N divided by the bandwidth B of the system. If we are simulating at one sample

per symbol, $B = R_b$:

$$\begin{aligned}
\frac{E_b}{N_0} &= \frac{S/R_b}{N/R_b} \\
&= \frac{S}{N} \\
&= \frac{A^2}{\sigma^2}
\end{aligned} \tag{3}$$

where A is the amplitude of each BPSK symbol and $\sigma^2 = N$ is the noise variance (mean noise energy per sample). If the noise is zero mean, we can estimate σ^2 over K noise samples r_i as:

$$\sigma^2 = \frac{1}{K} \sum_{i=0}^{K-1} r_i^2 \tag{4}$$

We generate noise samples r_i by sampling a unit variance, zero mean Gaussian noise source $\mathcal{N}_i(0,1)$ that is scaled by σ:

$$\begin{aligned}
r_i &= \sigma \mathcal{N}_i(0,1) \\
\sigma &= \frac{A}{\sqrt{Eb/N_0}}
\end{aligned} \tag{5}$$

In practice we simulate the channel by adding noise to the QPSK symbols. Due to orthogonality, this can be simulated as two independent channels that have the same noise power:

$$r_i = \sigma \mathcal{N}_{2i}(0,1) + j\sigma \mathcal{N}_{2i+1}(0,1) \tag{6}$$

2 SNR Measurement

In order to compare with other methods of speech communication, it is useful to formulate expressions for estimating SNR from the BPSK and QPSK symbols. The Signal to Noise ratio (SNR) is given by:

$$\begin{aligned}
\frac{S}{N} &= \frac{E_b R_b}{N_0 B} \\
&= \frac{E_q R_q}{N_0 B}
\end{aligned} \tag{7}$$

At one sample per symbol, the power, the mean energy of each QPSK symbol over a window of K samples is given by:

$$E_q = Var(q_i) = \frac{1}{K} \sum_{i=0}^{K-1} |q_i|^2 \tag{8}$$

As each QPSK symbol contains 2 BPSK symbols, the energy is split evenly:

$$E_b = E_q/2 \tag{9}$$

2

For example if the symbol amplitude is $A = 1$, $E_b = A^2 = 1$, then $E_q = 1 + 1 = 2$.

For transmission over multipath channels we arrange the QPSK symbols as N_c parallel carriers, each running at a symbol rate of R_s symbols/s, where R_s is chosen based on delay spread considerations. A typical value for HF modems is $R_s = 50$ Hz. However the OFDM carriers are arranged such that the total QPSK symbol rate over the channel R_q remains constant, so for a given signal power S, $E_q = S/R_q$ remains constant.

The expressions above can be used as a check for the noise injection process:

1. Set a target E_b/N_0 for the simulation run.

2. Establish the target SNR from (7) evaluated using the target E_b/N_0.

3. Measure $Eq = Var(q_i)$ over a sample of transmitted symbols.

4. Calculate measured SNR using (7) and compare.

3   Glossary

Symbol	Explanation	Units
B	noise or signal bandwidth	Hz
d	dimension of latent vector bfz	
E_b/N_0	energy per BPSK symbol on spectral noise density	
E_q/N_0	energy per QPSK symbol on spectral noise density	
N_c	Number of carriers	
$\mathbf{q}$	vector of QPSK symbols	
q_i	single QPSK symbol, element of $\mathbf{q}$	
R_b	BPSK symbol rate	symbols/second
R_q	QPSK symbol rate	symbols/second
R_s	per carrier QPSK symbol rate	symbols/second
R_z	latent vector update rate	Hz
T_b	BPSK symbol period	seconds
T_q	QPSK symbol period	seconds
T_s	per carrier QPSK symbol period	seconds
T_z	time between latent vector updates	seconds
r_i	noise sample	
SNR	signal to noise Ratio	
S	total signal (carrier) power	Watts
N	total noise power	Watts
$\mathbf{z}$	Autoencoder output latent vector	
z_i	single latent vector element of $\mathbf{z}$, a BPSK symbol	

Table 1: Glossary of Symbols