FreeDV-036 Radio Autoencoder (RADE) V1 Introduction and Waveform Description

David Rowe VK5DGR

October 4, 2024 Git: `dc699d7` on branch `dr-doc`

1 Introduction

The purpose of the Radio Autoencoder (RADE) V1 is to send speech over a HF radio channel. The speech signal has an audio bandwidth of 8kHz, but the RADE V1 signal requires just 1500Hz of RF bandwidth. The Peak to Average Power Ratio (PAPR) is less than 1dB, allowing efficient use of transmitter power amplifiers. Our testing indicates RADE works well on low and high SNR HF radio channels, and has impressive speech quality compared to SSB and traditional digital voice over radio systems. RADE V1 requires more memory and CPU than a traditional digital voice system, but will run just fine with the resources of a typical PC.

You can run RADE on a Windows PC or laptop using the FreeDV-GUI application V2.0 and above.

This document is an introduction to Version 1 of RADE, and a description of the waveform. The target audience is the Radio Amateur and regulatory organisations that govern Amateur Radio.

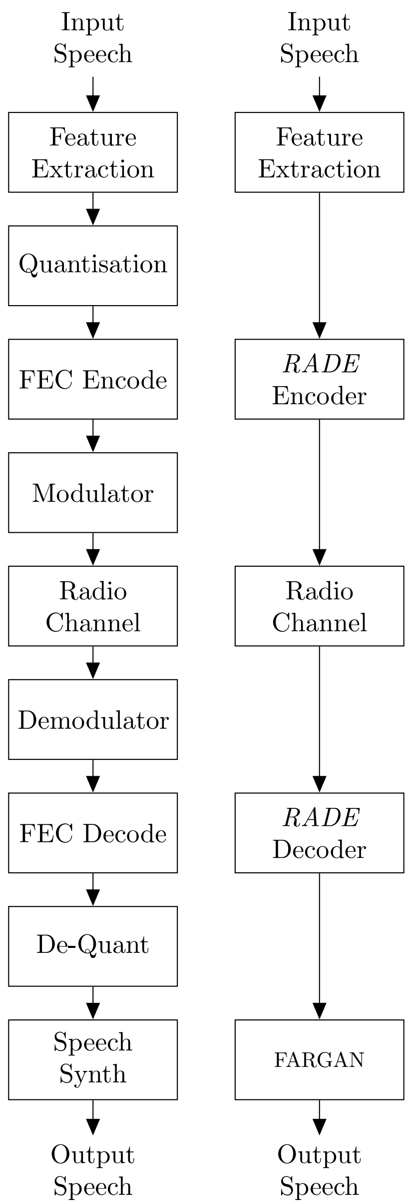
1.1 Acknowledgements

The RADE concept evolved from a discussion between Jean-Marc Valin and David Rowe, after which Jean-Marc quickly put together an initial proof-of-concept. Over a period of several months David built on this work to develop a practical over the air waveform for speech over HF radio channels. Mooneer Salem is handling integration of RADE into the FreeDV GUI application. The FreeDV Project Leadership Team and many others have helped with support and testing over the course of 2024. The contributions from David, Mooneer and the FreeDV PLT was generously supported by a grant from Amateur Radio Digital Communications (ARDC).

2 Radio Autoencoder

Figure 1 compares a traditional digital speech over radio system to RADE. In conventional digital speech systems, the speech encoder extracts features like

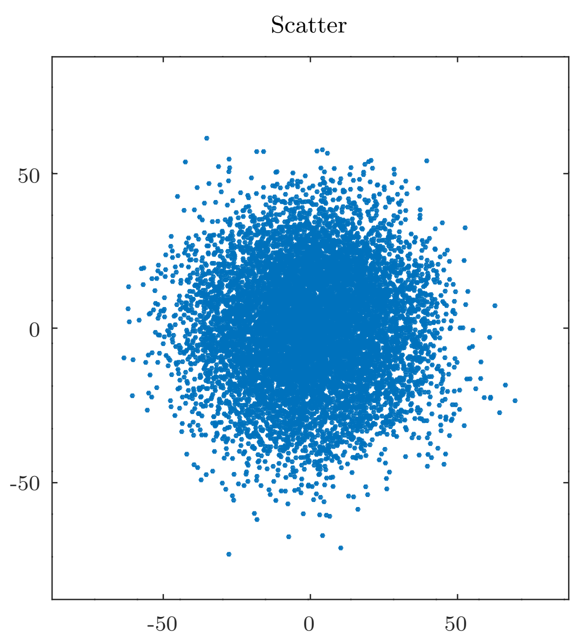
Figure 1: Traditional Digital Voice at left, RADE at right.



pitch, voicing, and short term spectrum, and quantises them to a fixed number of bits, e.g. 700 bits/s. Forward Error Correction adds extra bits to protect the encoded speech from bit errors. The FEC encoded bit stream is then passed to a modulator that generates an analog signal we can send through a radio transmitter over the channel. The demodulator takes the received signal, and converts it to a stream of bits. Some of these bits will have errors, which the FEC decoder will attempt to correct . Finally, the bits are converted back to vocoder features (De-quantised), and speech is synthesised.

RADE takes a novel twist – the Encoder converts vocoder features directly to Phase Shift Keyed (PSK) symbols. It effectively combines quantisation, FEC coding, and modulation. The RADE Decoder converts received PSK symbols back to features that are synthesised using the high quality FARGAN synthesis engine. The RADE encoder, decoder, and FARGAN synthesiser are built using modern machine learning techniques. RADE has been trained to produce good quality speech even with the distortion of the HF radio channel. Not shown on Figure 1 is some traditional DSP that converts the PSK symbols to and from an OFDM signal, and house keeping tasks like synchronisation. The PSK symbols are sent over the channel at 2000 symbols/second.

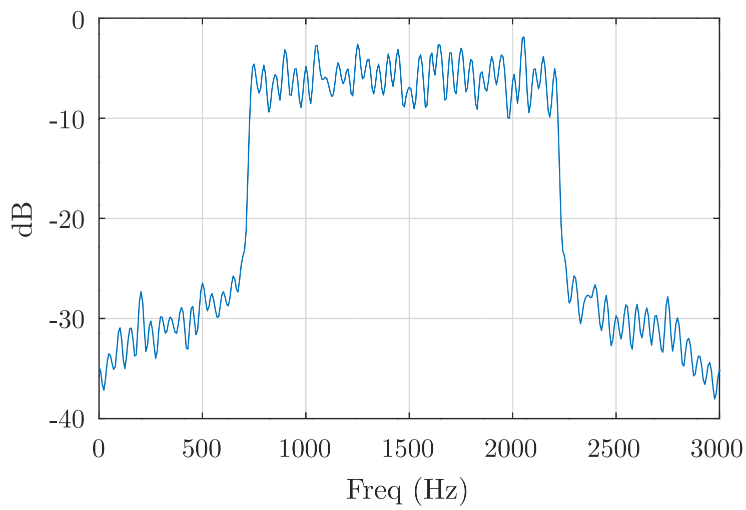
Figure 2: Constellation plot of RADE Encoder PSK symbols. Compared to traditional QPSK, the constellation looks like noise.



As shown in Figure 2 the PSK symbols from RADE are not discrete constellation points like traditional digital modems, instead they appear to be positioned at random. This constellation was "designed" by training the Autoencoder using many examples of speech and HF channels. Interestingly, there are no "bits" anywhere in the RADE system. The values from the features extractor, PSK symbols, through to synthesis are floating point numbers. The RADE signal can be seen as a form of sampled, analog PSK, built with a combination of machine learning and classical DSP techniques.

Figure 3 is the spectrum of the RADE V1 signal. This is similar to other OFDM waveforms, with a RF bandwidth of 1500 Hz. The spectral "grass" at the low and high frequency edges is relatively high at around -25dB from the peak, suppression of this has not been optimised in the V1 release.

Figure 3: Spectrum of RADE V1 Signal.



3 Waveform Description

Parameter	Value	Comment
Audio Bandwidth	100-7900 Hz	
RF Bandwidth	1500 Hz	-6dB from peak
Modulation	OFDM	Discrete time, continuously valued PSK symbols
Frame size	120ms	Algorithmic latency
Vocoder	FARGAN	Low CPU complexity ML vocoder
Payload symbol rate	2000 Hz	Payload data symbols, all subcarriers combined
Number of Subcarriers	30	
Subcarrier Symbol rate	50 Hz	
Subcarrier Spacing	50 Hz	
Cyclic Prefix	4ms	
Tx Peak Average Power Ratio	< 1dB	
Threshold SNR (AWGN)	-3dB	AWGN channel, 3000 Hz noise bandwidth
Threshold C/No (AWGN)	32 dBHz	AWGN channel
Threshold SNR (MPP)	0 dB	Multipath Poor (MPP) channel (1Hz Doppler, 2ms path delay), 3000 Hz noise bandwidth
Threshold C/No (MPP)	35 dBHz	
Worst case channel	MPD	Multipath Disturbed (MPD) channel 2Hz Doppler spread, 4ms path delay, two path Watterson model
Mean acquisition time	< 1.5s	0dB SNR MPP channel
Acquisition frequency range	+/- 50 Hz	
Acquisition co-channel interference tolerance	-3dBC	Interfering sine wave level for $<$ 2s mean acquisition time
Auxilary text channel	25 bits/s	Note all aux bits used for sync on RADE V1, no bits available for text
SNR measurement	No	

Table 1: RADE V1 waveform and performance parameters