

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

David Rowe VK5DGR

Sat Oct 5 09:21:06 2024 Git: (in *packed-refs*) 7 on branch

Japanese Translation by RADE test team Yuichi Matsui, JH0VEQ. Nov.1. 2024 draft5

1 はじめに(Introduction)

ラジオ・オート・エンコーダ (RADE) V1の目的は、HF無線周波数で音声を送信することです。音声信号の帯域幅は8kHzですが、RADE V1信号のRF帯域幅は1500Hz(-6dB値)で済みます。ピーク対平均電力比 (PAPR) は1dB未満で、送信機のパワーアンプを効率的に使用できます。

私たちのテストによると、RADEは低SNRおよび高SNRのHF無線周波数で良好に動作し、SSBや従来のデジタル音声システムと比較して、印象的なスピーチ品質を持っています。RADE V1は、従来のデジタル音声システムよりも多くのメモリとCPUリソースを必要としますが、一般的なPCで問題なく動作します。

RADEは、FreeDV-GUIアプリケーションV2.0以上を使用して、WindowsデスクトップPC、ラップトップPC、または、MAC上で動作させることができます。

このドキュメントは、RADE のバージョン 1 の紹介と波形の説明です。対象読者は、アマチュア無線家およびアマチュア無線を管理する規制機関です。

1.1 謝辞(Acknowledgements)

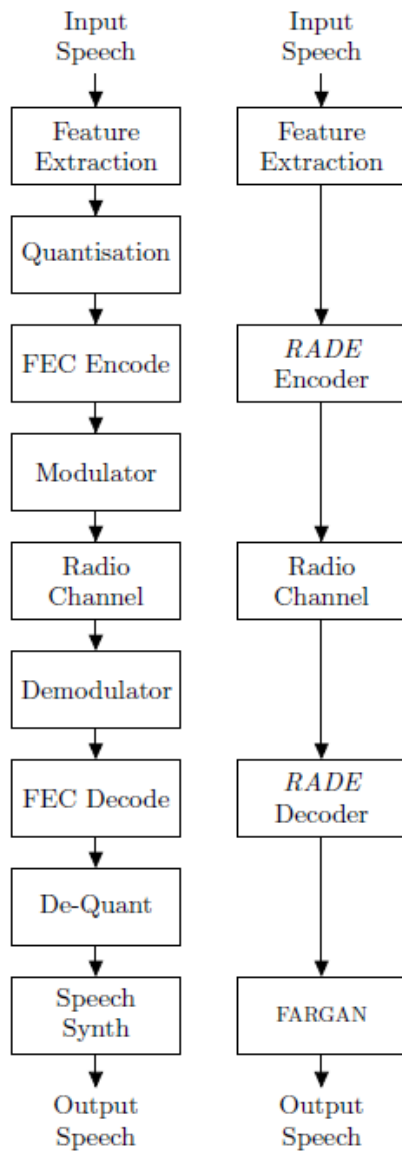
RADEのコンセプトは、Jean-Marc ValinとDavid Roweのディスカッションから発展しました。Davidは、数ヶ月かけ、このコンセプトを基に、HF無線周波数でのスピーチ用の実用的なOver-the-Air波形を開発しました。Mooneer Salem は FreeDV GUI アプリケーションへの RADE の統合を担当しています。FreeDV Project Leadership Team をはじめとする多くの人たちが、2024 年の間、サポートやテストに協力してくれました。David、Mooneer、FreeDV PLT の貢献は、Amateur Radio Digital Communications (ARDC) からの助成金によって支援されています。

2 RADE概要 (Radio Autoencoder)

図 1 は、従来からの無線デジタル・スピーチ・システムとRADEを比較したものです。従来の無線デジタルスピーチ・システムでは、スピーチ・エンコーダ (Feature Extraction)

がピッチ、ボイシング、短期スペクトルなどの特徴を抽出して、一定のビット数（たとえば700ビット/秒）に量子化(Quantisation)します。

図 1: 左が従来の無線デジタルスピーチシステム、右がRADE。



前方誤り訂正 (FEC Encode) は、符号化された音声ビットをビット誤りから保護するために余分なビット(訂正符号ビット)を追加します。前方誤り訂正 (FEC Encode) でエンコードされたビットは変調器(Modulator)に渡され、無線周波数(Radio Channel)上に送信できるア

ナログ信号を生成します。復調器(Demodulator)は受信した信号をビットに変換します。ビットの中にはエラーがあり、FECデコード (FEC Decode) は、そのエラーを修正しようとしています。最後に、ビットはボコーダ(De-Quantised)で音声特徴(量)に戻され、スピーチシンセサイザー (Speech Synth) で音声に合成されます。

RADEは斬新な工夫を凝らしています。RADEエンコーダ (RADE Encoder) の機能は、ボコーダから直接 PSK (Phase Shift Keyed) シンボルに変換します。これは、量子化、FEC Encode、符号化、変調を効果的に組み合わせたものです。RADEデコーダ (RADE Decoder) は、受信したPSKシンボルを、高品質のFARGAN合成器を使って、音声合成します。

RADEエンコーダ、RADEデコーダ、FARGAN合成器は、最新の機械学習技術 (ML) を使用して構築しています。RADEは、HF無線周波数に歪みがあっても高品質の音声を生成するように訓練されています。図 1には示していませんが、PSKシンボルをOFDM 信号に変換したり、OFDM信号からPSKシンボルに変換したりする従来の DSPと、同期のようなハウスキーピング・タスクがあります。PSKシンボルは2000シンボル/秒で無線周波数上に送信されます。

図 2: RADEエンコーダ PSKシンボルのコンステレーション・プロット。従来のQPSKと比較すると、コンステレーションはノイズのように見えます。

図 2: Scatter

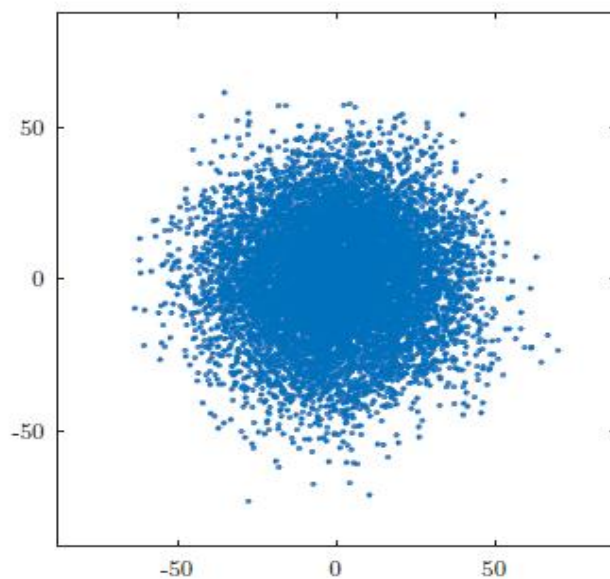
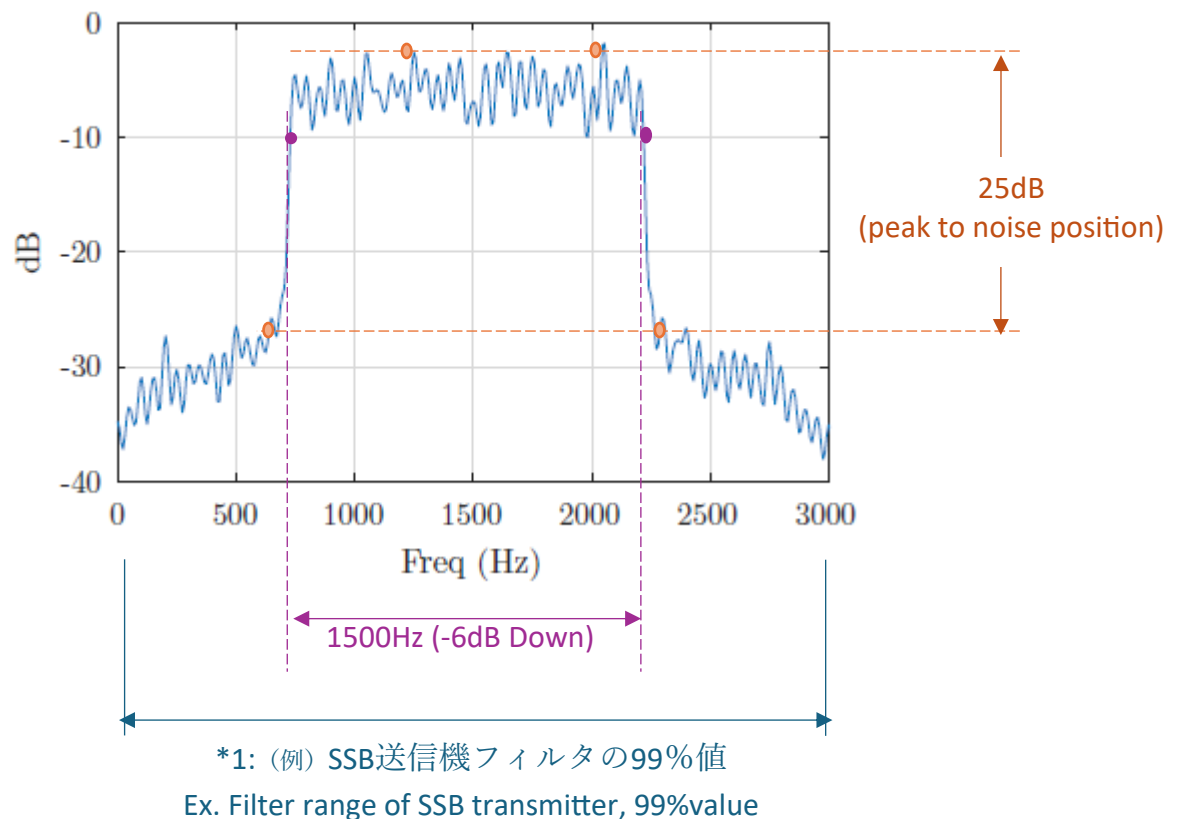


図2に示すようにRADEエンコーダ(RADE Encoder)のPSKシンボルは、従来のデジタルモデムのように離散的なコンステレーションポイントではなく、ランダムに配置されてい

るように見えます。このコンステレーションは、HF無線周波数での多くの音声信号例を用い、オートエンコーダを訓練することで、つくられたものです。特徴的なのは、RADEシステムはビットで処理をしません。特徴抽出器 (Feature Extraction) を通りPSKシンボル合成まで浮動小数点で処理します。つまり、RADE信号は、機械学習と古典的なDSP技術を組み合わせて作った、一種のアナログPSKと見ることができます。

図 3 はRADE V1信号のスペクトラムです。波形は他のOFDMと同様で、図のRF帯域幅は1500Hzです。スペクトルの高・低エッジの「ノイズ (grass)」までは、ピークから約-25dBと比較的高く、V1リリースではこの抑制が最適化されていません。

図 3: Spectrum of RADE V1 Signal.



図の補記は訳者の追記です。注意参照 (See Note)
Figures are supplemented by the translator's additions.

3 波形概要(Waveform Description)

パラメータ	値	コメント
音声帯域	100-7900 Hz	
RF帯域幅 (-6dB値)	1500 Hz	ピークから-6dBの位置で1500Hz
占有帯域幅 (99%値) 注意参照 (See Note)	TBD (1.5k - 3kHz)	OBW: 図 3 の条件では 3 KHz, 最小 1.5KHz. OBW:3KHz in Figure 3, Min 1.5KHz.
変調方式	OFDM	離散時間 , continuously valued PSK symbols
フレームサイズ	120ms	アルゴリズムによる遅延
ボコーダ	FARGAN	低CPU機械学習ボコーダ
パイロード シンボルレート	2000 Hz	全サブキャリアの合計
サブキャリア数	30	
サブキャリアシンボルレート	50 Hz	
サブキャリア間隔	50 Hz	
サイクリックプリフィックス	4ms	
ピーク対平均電力比 PAPR	< 1dB	
SNR 閾値 (AWGN)	-3dB	AWGN環境, ノイズ帯域幅 3000 Hz
C/No 閾値 (AWGN)	32 dBHz	AWGN環境
SNR 閾値 (MPP)	0 dB	Multipath Poor (MPP) 環境(ドップラ周波数1Hz, 伝搬遅延 2ms), ノイズ帯域幅 3000 Hz
C/No 閾値 (MPP)	35 dBHz	
ワーストケースチャネル	MPD	Multipath Disturbed (MPD)環境 (ドップラ拡散2Hz, 伝搬遅延 4ms, 2パスWattersonモデル)
アキュイジション時間	< 1.5s	0dB SNR MPP環境において
アキュイジション周波数	+/- 50 Hz	
アキュイジション隣接妨害耐性	-3dBC	平均アキュイジション時間が2秒以下となる正弦波干渉レベルを受けた状態において
補助テキストチャネル	25 bits/s	RADE V1は全ての補助ビットが同期に使用され、テキストに使用できるビットはない
SNR 測定	不可	
スケルチ	-2.5dB	早期リリースは固定値 squelch are fixed value.

表 1: RADE V1の波形とパフォーマンス・パラメーター

青文字は訳者追記. Blue letters indicate translator's additions

訳者追記 (translator's additions.)

参考資料(Additional Reference)

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

https://github.com/drowe67/radae/blob/dcd33aa142f5e418e82f1d3d1873d081c2f5c39d/doc/rade_intro_waveform.pdf

Watterson Channel Model for gnuradio

<https://github.com/noahthurston/gr-watterson>

FARGAN Vocoder

<https://arxiv.org/html/2405.21069v1#S2>

注意 (Note.)

*1: 占有帯域幅 (OBW) は変調後の電力波形の99%が占める帯域幅値であり、実送信機のフィルタ特性に依存するため、送信機に対する出力レベルをソフトウェアで調整し占有帯域幅を確認する必要があります。マイク入力端子を使う場合は特に注意をして下さい。

なぜなら、占有帯域幅は送信機のフィルタによっては予想以上に広がる場合があるためです。なお、送信機フィルタによっては最小占有帯域を1.5KHzまで制限できます。

図3 はでSSB変調でのシュミレーションのため、実際の送信波形とは違いがでます。

*1: In Japan, the occupied bandwidth (OBW) after modulation is defined as the 99% value occupied by the power waveform. Since it depends on the filter characteristics of the actual transmitter, adjust the input level of the transmitter and check the occupied bandwidth. Special care should be taken when using a microphone input jack.

This is because although this occupied bandwidth may be wider than expected for some transmitter filters, occupied bandwidth can be limited to a minimum of 1.5 KHz by the transmitter filter.

Since this occupied bandwidth depends on the transmitter filter characteristics, it should be checked before operation after the transmitter input level is adjusted. Figure 3 is a simulation using SSB modulation, so there should be differences from the actual waveform..

「このバージョンは技術検証のための早期リリースです。利用者は自身の置かれた環境に十分留意して自己責任で使ってください。」

This version is an early release for technical verification. Users should be aware of their environment and use at their own risk