Deep Redundancy for the Opus Codec

`draft-ietf-mlcodec-opus-dred-02`

Presenter: Jean-Marc Valin
jmvalin@jmvalin.ca

IETF 122

Changes Since IETF 120

- No change to code or spec

Training Procedure

- Documented in dnn/torch/rdovae/README.md
- New scripts to download and process data
  - download_datasets.sh
  - process_speech.sh
- Training only partially documented
  - Will work, but some fine-tuning steps missing
- **Will not produce the exact weights that will be in the standard**

Testvectors Proposal

- Provide a set of DRED payloads in simple *ad hoc* binary format (similar to opus_demo format)
- Provide corresponding decoded float features
- Create tool that decodes DRED payloads and compares them to reference decoded features
  - How loose should the test be?

Open Questions

- Where to publish weights?
- Options for defining the features/vocoder:
  - Provide reference vocoder, allow deviation
  - Define minimal objective metrics on vocoder synthesis
  - “Feature analysis on vocoder output must match input”
  - Encoder-side cepstrum + neural pitch calculation
- What else do we need to finish?