

Platforms for simulating the impact of mobile phone networks on speech

Esam A. S. Alzghoul¹, Balamurali B. T. Nair¹, Bernard J. Guillemin^{1*}

¹ Department of Electrical and Computer Engineering, The University of Auckland, New Zealand

*bj.guillemin@auckland.ac.nz

In the mobile phone arena, there are a number of network providers utilizing different technologies such as the GSM and CDMA networks. These networks are different in their internal operation which makes their impact on the speech signal also different. The medium of transmission in these networks is wireless and it is constantly changing, which required the development of highly innovative speech codecs to control a variable data rate capacity while maintaining a good quality of speech. The most common codecs in these networks are the Adaptive multi-rate codec (AMR) found in the GSM network [1] and the Enhanced Variable Rate Codec (EVRC) found in the CDMA network [2].

At The University of Auckland speech-codec platforms for these mobile phone networks are currently in the final stages of development. The purpose of these is to simulate their impact on the speech signal. One of these platforms, namely the EVRC-codec platform, is shown in Figure 1. These facilities will enable forensics scientists to understand the impact of mobile phone networks under different transmission conditions on the speech signal and to what extent these might impact on the outcome of a forensic voice comparison (FVC).

The new facilities incorporate several features. Firstly, they can be used to simulate the dynamic-rate coding aspect. For both codecs, bit rates can be different for every 20ms frame [1, 2]. Another aspect which can be investigated with these platforms is the misclassification of active speech frames as silence, especially when the speech is of low amplitude. The third is in respect to frames which are either lost or corrupted.

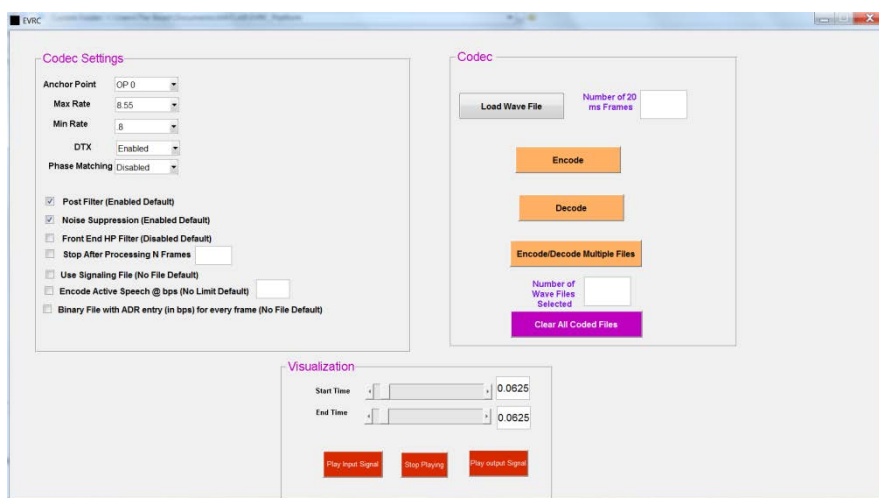


Figure 1. The EVRC-codec platform.

In the case of the AMR codec, its strategy is to replace a lost frame with a best guess based upon

previous frames. In the case of the EVRC codec the strategy is much simpler. Lost frames are simply replaced by the last good frame. One final important aspect is Noise Suppression (NS). This only applies to the EVRC codec as the GSM codec is not equipped with a similar capability. The concern with NS is that it can potentially remove part of the original speech in the presence of high background noise levels [3].

These two platforms will allow forensic speech scientists to investigate all aspects of the operation of these codecs on parameters important to FVC. This is particularly important because amongst a number of forensic speech scientists is the belief that the landline and mobile phone networks impact similarly on the speech signal. Further, there is the view that amongst mobile phone networks, different mobile phone technologies act similarly on the speech signal. Neither of these is true.

References

- [1] 3GPP (2002a) *TS 26.071 Adaptive Multi-Rate (AMR) speech codec (general description)*. Retrieved on 25 of March 2012 from http://www.3gpp.org/ftp/Specs/-archive/26_series/.
- [2] 3GPP2, C. S0018-D, Minimum Performance Specification for the Enhanced Variable Rate Codec, Speech Service Options 3, 68, 70, and 73 for Wideband Spread Spectrum Digital Systems. 2010. Retrieved on 15 of March 2012 from http://www.3gpp2.org/public_html/-specs/C.S0018-D_v1.0_MPS_for_EVCR.pdf.
- [3] Macdonald, M., EVRC: Best of Both Worlds. Retrieved on 5 of April 2012 from http://connectedplanetonline.com/wireless/mag/wireless_evrc_best_worlds/. March, 1999.