An interpolative decoding approach for speech streaming services and Voice over IP

Tim Fingscheidt and Jes´us F. Guitarte P´erez
Siemens AG, Mobile Phones (ICM MP)

Abstract

Voice-over-IP services usually require a receiver buffer to overcome the effect of delay jitter and lost packets. In streaming services ARQ schemes (automatic repeat request) are used to retrieve the lost packets. Since streaming as a non-conversational service has relaxed delay constraints, a comparably large buffer can be implemented. Buffering in general is an important means for achieving robustness in these two application scenarios.

When frames of coded speech are buffered at the receiver (pre-decoder buffering), the speech coder parameters of lost frames can be computed by interpolation rather than extrapolation. Modern speech coders rely heavily on predictive quantization which makes interpolation a difficult task that was hardly addressed in literature. In [1] e.g. interpolative approaches have been proposed for non-predictively coded parameters, however it was stated: ”*Since it seems reasonable to expect improvements from the use of future in-formation even in predictive quantization, more work seems possible in this direction*”.

In this paper we present such work proposing a solution to interpolation of predictively coded parameters. Significant gains are achievable over GSM channels as well as random frame loss channels.