GRACEFUL DEGRADATION IN ADPCM SPEECH TRANSMISSION

Tim Fingscheidt

Siemens AG, ICM Mobile Phones, Grillparzerstrasse 10-18, D – 81675 Munich, Germany
Email: Tim.Fingscheidt@siemens.com

ABSTRACT

Speech transmission according to the digital cordless telephony standards DECT, PHS, and PACS [1, 2, 3] is not protected by a forward error correction (FEC) scheme. In case of adverse channel conditions an error concealment technique must be applied to guarantee an acceptable speech quality or even to mute the speech signal. The standards do not specify a mandatory error concealment mechanism, accordingly there is room for new solutions to improve the decoding process of disturbed speech frames.

The mentioned cordless telephony standards employ the G.726 ADPCM speech codec [4]. Exploiting a frame reliability information from the demodulator in addition to some a priori knowledge about the ADPCM speech codec parameter, we apply the *softbit speech decoding* techniques proposed in [5, 6, 7, 8] to the ADPCM decoder. It turns out that the robustness of the ADPCM softbit speech decoder outperforms state of the art mechanisms to conceal bit errors in cordless telephony systems. The new algorithm encloses an inherent muting mechanism leading to a graceful degradation of speech quality in case of adverse transmission conditions. If the channel is error free, bit exactness as required by the standards can be preserved.

1. ADPCM DECODING

In the following we focus on the DECT standard, although the derived algorithms can be applied to any cordless telephony system employing the ADPCM codec. The 32 kbit/s G.726 ADPCM codec used in DECT speech transmission delivers 80 parameters each consisting of 4 bits $\underline{x}_n = (x_n(0), x_n(1), x_n(2), x_n(3))$ per 10 ms DECT frame. The first bit $x_0(0)$ at sample instant $n = 0$ denotes the sign, and the following 3 bits $\underline{z}_0 = (x_0(1), x_0(2), x_0(3))$ are an exponential representation of the magnitude of a residual signal sample dq_0.

Assuming the bits $x_n(m)$ and $\hat{x}_n(m)$ being bipolar, the conventional decoder performs

$$dq_0 = \hat{x}_0(0) \cdot 2^{\mathbf{BM}^{-1}[\underline{\hat{z}}_0] + y_0} \tag{1}$$

with $\mathbf{BM}^{-1}[\underline{\hat{z}}_0]$ denoting an inverse bit mapping scheme / lookup table with 8 values, and

$$y_0 = al_0 \cdot yu_{-1} + (1 - al_0) \cdot yl_{-1} \tag{2}$$

being a scale factor. The scale factor consists of a weighted sum of the *unlocked* scale factor yu_{-1} and the *locked* scale factor yl_{-1}, both depending on $\underline{\hat{z}}_{-1}$ in the previous sample instant $n = -1$. The adaptation speed parameter al_0 serves as weighting factor.

2. ADPCM SOFTBIT DECODING

An entropy measurement of $\underline{x}_n$ over a large data base yields $H = 3.62$ bit. Thus there is a redundancy of $\Delta R = 4$ bit$-H = 0.38$ bit that can be exploited by an appropriate error concealment if the discrete probability distribution $\mathrm{P}(\underline{x}_n{}^{(i)})$ with $i = 0, 1, ..., 15$ being the quantization table index is available at the decoder.

The ADPCM softbit speech decoder is controlled by a frame reliability information p_e denoting the mean bit error probability in the current frame of 320 bits. It is assumed to be perfectly estimated at the demodulator [8]. Nevertheless, a certain misadjustment of p_e can be tolerated. Now channel dependent transition probabilities $\mathrm{P}(\underline{\hat{x}}_0|\underline{x}_0{}^{(i)}), i = 0, 1, ..., 15$, from any possibly transmitted 4-bit combination $\underline{x}_0{}^{(i)}$ to the known received bit combination $\underline{\hat{x}}_0$ can be computed [8]. An a posteriori probability distribution is given by $\mathrm{P}(\underline{x}_0{}^{(i)}|\underline{\hat{x}}_0) = C \cdot \mathrm{P}(\underline{\hat{x}}_0|\underline{x}_0{}^{(i)}) \cdot \mathrm{P}(\underline{x}_0{}^{(i)})$ with C normalizing the sum of all $\mathrm{P}(\underline{x}_0{}^{(i)}|\underline{\hat{x}}_0), i = 0, 1, ..., 15$, to one.

In contrast to eq. (1), the ADPCM *softbit* speech decoder performs an estimation of the residual signal sample dq_0 according to

$$dq_0^{(\mathrm{MS})} = \sum_{i=0}^{15} x_0^{(i)}(0) \cdot 2^{\mathbf{BM}^{-1}[x_0^{(i)}(1), x_0^{(i)}(2), x_0^{(i)}(3)] + y_0^{(\mathrm{MS})}} \cdot \mathrm{P}(\underline{x}_0{}^{(i)}|\underline{\hat{x}}_0) \ . \tag{3}$$

Application of eq. (3) minimizes the mean squared error (MMSE) of the residual signal sample dq_0. The performance of the ADPCM decoder is very sensitive to a proper value of the scale factor y_0. Because the scale factor itself depends on previously received bit combinations $\underline{\hat{x}}_n, n = -1, -2, ...$, eq. (3) assumes it to be MS estimated as well, following

$$y_0^{(\mathrm{MS})} = \log_2 \left[\sum_{j=0}^{7} 2^{y_0^{(j)}} \cdot \mathrm{P}(\underline{z}_{-1}{}^{(j)}|\underline{\hat{z}}_{-1}) \right] \ . \tag{4}$$

The probability term related to $\underline{z}_{-1}{}^{(j)}$ can easily be derived by $\mathrm{P}(\underline{z}_{-1}{}^{(j)}|\underline{\hat{z}}_{-1}) = \mathrm{P}(\underline{x}_{-1}{}^{(j)}|\underline{\hat{x}}_{-1}) + \mathrm{P}(\underline{x}_{-1}{}^{(j+8)}|\underline{\hat{x}}_{-1})$ assuming that $x_{-1}^{(j)}(0)$ is the differing bit between $\underline{x}_{-1}{}^{(j)}$ and $\underline{x}_{-1}{}^{(j+8)}$. The scale factor $y_0^{(j)}$ requires 8 times computation of eq. (2), which is then

$$y_0^{(j)} = al_0 \cdot yu_{-1}^{(j)} + (1 - al_0) \cdot yl_{-1}^{(j)}, \ j = 0, 1, ..., 7. \tag{5}$$

Note that the adaptation speed al_0 in eq. (5) is not very sensitive against bit errors, therefore it is computed only once based on the received $\underline{\hat{z}}_{-1}$.

The 2×8 scale factors needed for eq. (5) are computed by

$$yu_{-1}^{(j)} = (1 - 2^{-5}) \cdot y_{-1}^{(\mathrm{MS})} + 2^{-5} \cdot W[j], \tag{6}$$

and

$$yl_{-1}^{(j)} = (1 - 2^{-6}) \cdot yl_{-2} + 2^{-6} \cdot yu_{-1}^{(j)}. \tag{7}$$

Here the estimated scale factor from the previous sample instant, a tabulated value $W[.]$ (see [4]), and, especially, a *single* value for yl_{-2} are needed.

We need to make available this single value of yl_{-1} for the computations of $y_1^{(j)}$ by eq. (5) in the next sample instant $n = 1$. Therefore, by putting eq. (7) into eq. (5) and omitting the index (j), we perform

$$yu_{-1} = \frac{y_0^{(\mathrm{MS})} - (1 - al_0) \cdot (1 - 2^{-6}) \cdot yl_{-2}}{2^{-6} \cdot al_0 \cdot (1 - al_0)} \tag{8}$$

using the estimated scale factor $y_0^{(\mathrm{MS})}$, and

$$yl_{-1} = (1 - 2^{-6}) \cdot yl_{-2} + 2^{-6} \cdot yu_{-1}. \tag{9}$$

Based on these *backward adapted* scale factors, again 8 versions of yu_0 and yl_0 can be computed using the recursive formulae eqs. (6) and (7) in the next sample instant.

We found that the LMS algorithm for updating the coefficients of the synthesis filter $H(z) = \frac{1+B(z)}{1-A(z)}$ prefers reliable signs rather than reliable values of the residual signal dq_0. Therefore these updates are based on the erroneously received bits $\underline{\hat{x}}_0$ in conjunction with the estimated scale factor $y_0^{(\mathrm{MS})}$. This requires the decoder to perform a second $B(z)$ filtering in addition to the $H(z)$ synthesis filtering of $dq_0^{(\mathrm{MS})}$.

3. SIMULATION RESULTS

To simulate a DECT-like channel behaviour, we used a frequency-nonselective Rayleigh fading channel model with perfect frame synchronisation and 2-path selection diversity. Given the E_b/N_0 ratio and a user speed, each 10 ms a fading factor α was obtained. Assuming it to be constant over the whole frame, the bit error probability [9]

$$p_e = \frac{1}{2} \cdot \mathrm{erfc}(\sqrt{\alpha^2 \frac{E_b}{2N_0}}) \tag{10}$$

of a binary FSK modulation was computed to control a binary symmetric channel model. The same p_e was used for ADPCM softbit speech decoding as well as for a reference error concealment technique employing simple frame repetition of the 320 bits, if $p_e > 2\%$ (i.e. a *bad* frame). The 2% threshold was found empirically to deliver best results for the frame repetition technique. In case of successive bad frames, earlier good frames are taken for repetition and after 4 bad frames muting takes place.

Fig. 1 shows the simulation results for a user speed of 30 cm/s (1.08 km/h). It turns out that there is a good correlation of speech SNR and speech quality in ADPCM speech transmission. Below $E_b/N_0 = 20$ dB the curve of the standard ADPCM decoder is characterized by a steep degradation of speech quality. Our reference frame repetition error concealment technique yields a better quality but still has severe clicks and artifacts. Finally, the proposed ADPCM softbit speech decoding technique gives a surprisingly good speech quality showing graceful degradation in the case of channel errors. Having the estimate of p_e available from the demodulator, the additional complexity compared to pure ADPCM decoding can roughly be estimated to be 3...4 MIPS.

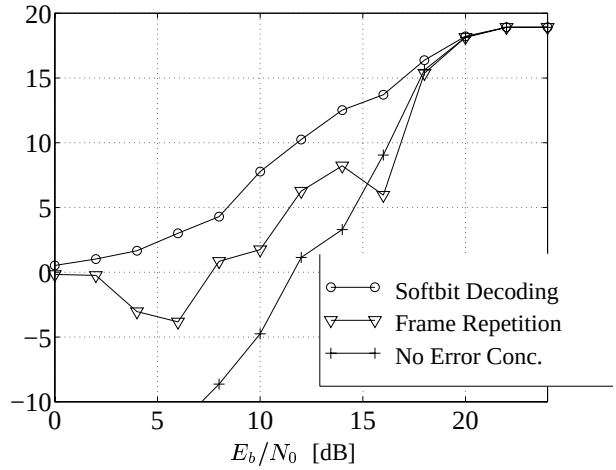

Fig. 1. Speech SNR in [dB]: DECT speech transmission using the ADPCM speech codec with 3 different decoding variants.

4. CONCLUSION

In this contribution we applied the softbit speech decoding technique to the highly recursive, sample-by-sample working ADPCM coder. By exploiting less than 10% redundancy in the transmitted 4 bit indices, simulations close to DECT-like channel behaviour show a graceful degradation and a significantly higher speech quality as achieved by conventional frame repetition algorithms.

5. REFERENCES

[1] "ETS 300 175, Digital Enhanced Cordless Telecommunications Common Interface," ETSI, 1992.

[2] "RCR STD-28, Personal Handy Phone System," ARIB, 1995.

[3] "ANSI J-STD-014, Personal Access Communications System Unlicensed," ANSI, 1995.

[4] "ITU-T Recommendation G.726, 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)," ITU, 1990.

[5] T. Fingscheidt and P. Vary, "Error Concealment by Softbit Speech Decoding," in *Proc. of ITG-Fachtagung "Sprachkommunikation"*, Frankfurt a.M., Germany, Sept. 1996, pp. 7–10, VDE–Verlag.

[6] T. Fingscheidt and P. Vary, "Robust Speech Decoding: A Universal Approach to Bit Error Concealment," in *Proc. of ICASSP'97*, Munich, Germany, Apr. 1997, vol. 3, pp. 1667–1670.

[7] P. Vary und T. Fingscheidt, "Von der Soft-Decision-Kanaldecodierung zur Softbit-Sprachdecodierung," *it+ti – Informationstechnik und Technische Informatik*, Schwerpunktthema: Neue Codierverfahren für die Telekommunikation, vol. 41, no. 1, Jan. 1999.

[8] T. Fingscheidt and P. Vary, "Softbit Speech Decoding: A New Approach to Error Concealment," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 9, no. 3, pp. 240–251, Mar. 2001.

[9] J.G. Proakis, *Digital Communications*, McGraw-Hill, New York, 1989.