

AN IMPROVED ADPCM DECODER BY ADAPTIVELY CONTROLLED QUANTIZATION INTERVAL CENTROIDS

Sai Han and Tim Fingscheidt

Institute for Communications Technology, Technische Universität Braunschweig
Schleinitzstr. 22, 38106 Braunschweig, Germany
{s.han,t.fingscheidt}@tu-bs.de

ABSTRACT

Adaptive differential pulse code modulation (ADPCM) has been standardized in ITU-T Recommendations G.726 and G.722 and is widely used in IP and cordless telephony. Although adaptive quantization and adaptive prediction is employed in ADPCM using a fixed scalar quantization codebook/lookup table, residual correlation of the quantizer input samples is yet observed. Exploiting source correlation, it has been shown that scalar quantization performance can be improved by a time-variant quantization interval *centroid* leading to an adaptive codebook in the *decoder*. Using an ADPCM encoder and applying this principle to the ADPCM decoder with its own adaptive quantization and prediction, the mean opinion score (MOS) of perceptual evaluation of speech quality (PESQ) is shown to improve by about 0.15 points for low bit rate ADPCM in error-free transmission conditions.

Index Terms— ADPCM, probability density function, centroid condition, prediction

1. INTRODUCTION

Quantization plays a central role in digital communications. As outlined in [1], a quantizer can be memoryless with a fixed codebook [2–4], or have memory reflecting the statistical properties of the source process with either a fixed or a time-variant codebook [5–7].

Memoryless quantizers [2–4] are designed in the same way for correlated as for uncorrelated processes. Utilizing source correlation asks for vector quantization (VQ) [8] or scalar quantization (SQ) with memory [5–7], the latter having predictive quantization [9, 10] and transform coding [11] as two representative approaches. In predictive quantization, predictors are required both at the encoder and decoder side; the difference between the original signal and its predicted signal is quantized [6, 7, 12–14].

Different from predictive quantization, scalar quantization performance for correlated processes can also be improved by our previous work [15–17] which leaves the encoder unchanged and employs a predictor only at the decoder side. In this approach, the signal is predicted depending on the previously reconstructed samples in the decoder. At each time instant, according to the fixed encoder decision levels and an instantaneously shifted *decoder-sided* prediction error probability density function (PDF), a *time-variant* quantization codebook can be computed based on the centroid condition. This approach leads to significant improvements especially for low rate quantization and highly correlated processes. Moreover, it can advantageously be applied both in error-free and error-prone transmission conditions, either using hard decisions or soft decisions [15].

Adaptive differential pulse code modulation (ADPCM) [13, 14], including adaptive prediction backwards (APB) and adaptive quan-

tization backwards (AQB) methods, offers speech quality at different levels and is widely used in digital enhanced cordless telephony (DECT) [18], New Generation DECT [19], and voice over IP (VoIP). When the input process is nonstationary (e.g., speech), the variance changes over time; then adaptive quantization can provide a better performance, especially for low bit rate quantization [5]. In adaptive quantization with backward estimation (AQB), the step size adaptation is performed on the basis of the reconstructed/quantized signal. Moreover, the nonstationarity of speech also requires adapting the predictor coefficients over time leading to a higher prediction gain than using fixed predictor coefficients [5]. In ADPCM, the adaptation information is recursively computed according to previous quantized signals (APB) [13, 14]. For speech with a sampling rate of 8 kHz, the predictor order of 10 is normally adequate. However, in G.726 and G.722 ADPCM [13, 14], only a second-order all-pole predictor and a sixth-order all-zero predictor are employed (autoregressive moving-average (ARMA) predictor).

Two frequency-domain approaches have been proposed to enhance the quality of coded speech: noise spectral shaping at the speech encoder [20] and adaptive post-filtering at the decoder output [21]. The combination with both encoder and decoder being modified has been investigated in [22, 23]. The soft-decision decoding technique from [24] has been applied to the G.726 ADPCM decoder in [25]; however, ADPCM performance was only improved in erroneous transmission conditions.

In ADPCM schemes, although the difference between the original signal and the ARMA-predicted signal is quantized by a scalar quantizer, residual correlation of the difference signal is yet observed. Therefore, we apply the improved scalar quantization/lookup approach from [15–17] to the ADPCM decoder to obtain improved speech quality even in error-free transmission conditions. Due to the nonstationarity of speech, and different from using a fixed *decoder-sided* predictor coefficient for a Gaussian autoregressive process as in our previous work, the normalized least mean-squares (NLMS) algorithm [26, 27] is utilized here to update the *decoder-sided* lookup table predictor (i.e., not the ARMA predictor) coefficients adaptively. These predictor coefficients are further used for computing the new quantizer centroids of the AQB scheme. The proposed system with a standard ADPCM encoder would be perfectly suited for system-compatible use in all ADPCM applications.

The paper is structured as follows: Section 2 briefly revisits a standard ADPCM encoder and decoder for later reference and notations. Section 3 presents our improved ADPCM decoder. Simulation results are discussed in Section 4. Finally, conclusions are drawn in Section 5.

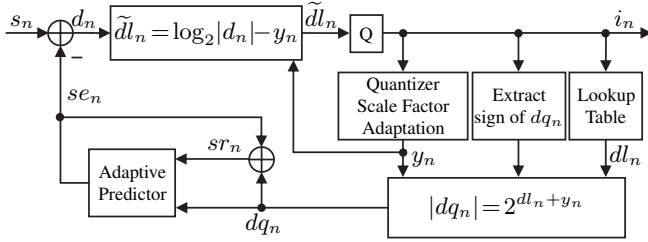


Fig. 1: Block diagram of a standard ADPCM encoder.

2. ADPCM CODING

2.1. Encoding

The ITU-T Recommendations G.726 narrowband codec and G.722 wideband codec are mandatory for New Generation DECT devices [19]. In G.722 subband-ADPCM [14], speech signals with 16 kHz sampling rate are split into a higher band and a lower band, with the signals in each subband (sampled at 8 kHz) being ADPCM-encoded. The higher subband can produce 16 kbit/s signals using a 2 bit quantizer, while the lower subband is operating at 48 kbit/s with a 6 bit quantizer. Due to the similarity of G.726 ADPCM [13] and the subband structure in G.722 ADPCM, we focus on the description and improvement of (G.726) ADPCM in this paper.

The simplified block diagram of a standard ADPCM [13] encoder is depicted in Figure 1. In the ADPCM encoder, the difference between the input signal s_n (uniform PCM) and the estimation of the input signal s_{e_n} is denoted as d_n , with n being the time index. Before quantization, d_n is first transformed to a base-2 logarithmic representation and thereafter scaled by a scaling factor y_n which is calculated by the quantizer scale factor adaptation block. This normalized quantizer input signal \tilde{d}_n is further scalar quantized and expressed by a corresponding quantization index i_n with five, four, three or two bits representation for operating at 40, 32, 24 or 16 kbit/s, respectively, with the first bit representing the sign of the difference signal. The quantization index i_n is assigned to a quantized normalized value dl_n by a lookup table. Extracting the sign from i_n and performing $|dq_n| = 2^{dl_n + y_n}$, the quantized difference signal dq_n is obtained. The reconstructed signal sr_n , which is produced by adding the quantized difference signal dq_n and the estimated signal s_{e_n} , along with the quantized difference signal dq_n are used as the input to the adaptive ARMA predictor generating the estimated signal s_{e_n} in a feedback loop.

2.2. Decoding

Figure 2 shows the simplified block diagram of an ADPCM decoder [13] including the same feedback loop structure as the encoder. The quantized difference signal dq_n , estimated signal s_{e_n} , and reconstructed signal sr_n remain the same as in the encoder. In the further description, we focus on error-free transmission conditions, which means the ADPCM encoder output i_n is received unchanged at the decoder.

3. IMPROVED ADPCM DECODING

3.1. An Outline

An improved ADPCM performance can be achieved by getting a better reconstructed signal sr_n . Since $sr_n = dq_n + s_{e_n}$, the signal can

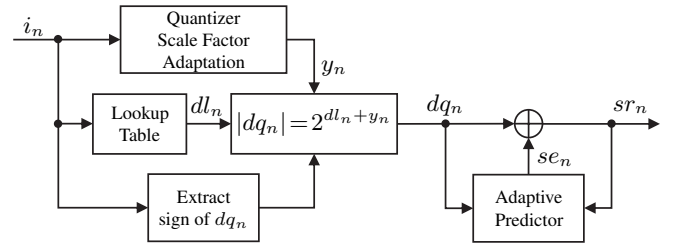


Fig. 2: Block diagram of a standard ADPCM decoder.

be reconstructed better if either the difference signal is better quantized or the signal is better estimated. However, as mentioned in Section 2.1, the estimated signal is produced by an adaptive ARMA predictor with dq_n and sr_n being the input in a feedback loop. Therefore, in order to keep the ARMA predictor coefficients estimation in encoder and decoder synchronized, a complete standard decoder should be employed, providing signal s_{e_n} (and y_n). The idea now is to compute a (new) better quantized difference signal $\hat{d}q_n$ and use it for a better final reconstructed signal $\hat{s}r_n = s_{e_n} + \hat{d}q_n$.

Furthermore, as shown in Figure 1, since the encoder quantizer input is the normalized difference signal \tilde{d}_n , comparing $|dq_n| = 2^{dl_n + y_n}$ in the decoder and $|d_n| = 2^{\tilde{d}_n + y_n}$ in the encoder, we can state that if \tilde{d}_n is quantized with smaller quantization error (i.e., dl_n is closer to \tilde{d}_n), the new quantized difference signal $\hat{d}q_n$ will accordingly be closer to the difference signal d_n , on the basis of the same given scale factor y_n . The derivation of a better quantized normalized value \tilde{d}_n is presented in the following sections.

3.2. Improved Scalar Quantization With an Adaptive Codebook

Due to the scalar quantization in ADPCM, the quantizer input \tilde{d}_n and the quantizer output dl_n are used to explain the *decoder-sided* improved scalar quantization approach, which has been presented in [15–17] and is summarized as follows:

The scalar non-uniform Lloyd-Max quantization (LMQ) [2, 3] is optimized for the minimum mean squared error (MMSE), with a fixed quantization codebook and the reconstruction levels being the centroid of the region of the (unquantized) signal PDF in the corresponding quantization interval, which is known as the centroid condition [28]:

$$dl^{(i)} = \frac{\int_{I^{(i)}} \tilde{d} \cdot p_{\tilde{D}}(\tilde{d}) d\tilde{d}}{\int_{I^{(i)}} p_{\tilde{D}}(\tilde{d}) d\tilde{d}}, \quad (1)$$

with $p_{\tilde{D}}(\tilde{d})$ being the PDF of the unquantized signal \tilde{d} and $I^{(i)}$ being the corresponding quantization interval for quantization index i .

For correlated processes, at each time index n , the current signal can be estimated according to the previously reconstructed signal samples generating the predicted signal $\hat{d}l_n^+$ at the decoder side and yielding the prediction error $\hat{e}_n = \tilde{d}_n - \hat{d}l_n^+$, which is the difference between the original signal and its *decoder-sided* predicted signal. The PDF of \tilde{d}_n conditioned on a (*decoder-sided*) known deterministic value $\hat{d}l_n^+$ turns out to be the prediction error PDF $p_{\hat{E}}(\cdot)$ of \hat{e}_n shifted by $\hat{d}l_n^+$:

$$p_{\tilde{D}}(\tilde{d}_n | \hat{d}l_n^+) = p_{\hat{E}}(\hat{e}_n = \tilde{d}_n - \hat{d}l_n^+) = f(\tilde{d}_n), \quad (2)$$

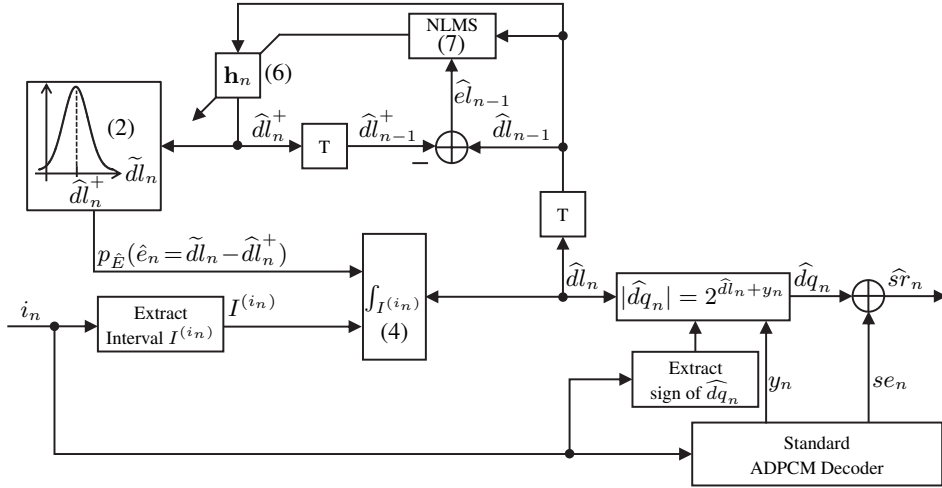


Fig. 3: Block diagram of the newly proposed bit-stream compatible ADPCM decoder.

with the shifted prediction error PDF being a function of \tilde{d}_n .

Furthermore, the centroid condition from (1) given a known predicted signal becomes a conditional one:

$$d_l^{(i_n)} = \frac{\int_{I^{(i_n)}} \tilde{d}_n \cdot p_{\text{DL}}(\tilde{d}_n | \hat{d}_n^+) d\tilde{d}_n}{\int_{I^{(i_n)}} p_{\text{DL}}(\tilde{d}_n | \hat{d}_n^+) d\tilde{d}_n}, \quad (3)$$

with $I^{(i_n)}$ being the quantization interval at time index n .

Comparing the prediction error PDF $p_{\hat{E}}()$ and the unquantized normalized signal PDF $p_{\text{DL}}(\tilde{d}_l)$, the prediction error has a smaller variance in general. Additionally, the two PDF shapes in the fixed quantization interval $I^{(i_n)}$ are different. As a result, applying the conditional PDF $p_{\text{DL}}(\tilde{d}_n | \hat{d}_n^+)$ from (2) to (3), a different centroid resulting in a new reconstruction level can be obtained by

$$d_l^{(i_n)} = \frac{\int_{I^{(i_n)}} \tilde{d}_n \cdot p_{\hat{E}}(\hat{e}_n = \tilde{d}_n - \hat{d}_n^+) d\tilde{d}_n}{\int_{I^{(i_n)}} p_{\hat{E}}(\hat{e}_n = \tilde{d}_n - \hat{d}_n^+) d\tilde{d}_n}, \quad (4)$$

basically constituting an adaptive lookup table. The training of the prediction error PDF $p_{\hat{E}}()$ will be explained in Section 4.1.

3.3. Application to ADPCM

In the following we describe the application of the technique above to any standard ADPCM decoder. For ease of description we focus on the 16 kbit/s mode with a 4-level quantizer (i.e., $i_n \in \{0, 1, 2, 3\}$) as being employed in the upperband of G.722 [14], or in G.726 [13], but application to the other bit rates is straightforward. The active speech part of 20 languages excluding American English from NTT monaural speech database [29] has been used for training to acquire the PDF $p_{\text{DL}}(\tilde{d}_l)$. Thereafter, a new LMQ quantization codebook is obtained using the centroid condition (1). As expected, the same mean opinion score (MOS) 2.48 of perceptual evaluation of speech quality (PESQ) (P.862) [30] is obtained compared to the MOS value using the standard quantization codebook/lookup table of the G.726 Recommendation. The above result is the mean MOS value from 96 American English speech files including 4 female and 4 male speakers. In consequence, we can state that the conditional centroid condition (4) can be further used for improving ADPCM quantization

performance.

The block diagram of our proposed decoder is depicted in Figure 3. The quantization index i_n having a two-bit representation is transmitted to the decoder, with the first bit representing the sign and the second bit representing the magnitude. In the ADPCM quantization domain (i.e., \tilde{d}_n), in fact one bit amplitude quantization is performed according to the value of $|d_n|$ and resulting in $i_n = 0$ or $i_n = 1$ for $d_n > 0$, while $i_n = 3$ or $i_n = 2$ for $d_n < 0$, respectively. Therefore, the actual 1-bit quantization interval I_n needs to be extracted by

$$I^{(i_n)} = \begin{cases} I_0, & \text{if } i_n = 0 \text{ or } i_n = 3, \\ I_1, & \text{else.} \end{cases} \quad (5)$$

The *decoder-sided* predicted signal \hat{d}_n^+ can be obtained by a linear prediction with predictor order of N_p :

$$\hat{d}_n^+ = \mathbf{h}_n^T \cdot \hat{\mathbf{d}}_{n-N_p}^{n-1}, \quad (6)$$

with the predictor coefficients $\mathbf{h}_n = (h_n(1), h_n(2), \dots, h_n(N_p))^T$ in a transposed vector manner and the previously received signal $\hat{\mathbf{d}}_{n-N_p}^{n-1} = (\hat{d}_{n-1}, \hat{d}_{n-2}, \dots, \hat{d}_{n-N_p})^T$. Since speech is non-stationary, at each time index n , the predictor coefficients \mathbf{h}_n should be adaptively updated to minimize the mean square error $E\{(\hat{e}_n)^2\}$ with $\hat{e}_n = \hat{d}_n - \hat{d}_n^+$. This can be achieved by the normalized least-mean-squares (NLMS) algorithm which will be further explained in Section 3.4. Thereafter, the prediction error PDF is shifted by \hat{d}_n^+ . Since the standard ADPCM encoder is used, the correct quantization interval $I^{(i_n)}$ where the original signal occurs is known by the decoder in error-free transmission conditions. In consequence, the decision levels are taken from the standard ADPCM quantizer as they are used in the G.726 Recommendation [13].

To sum up, based on the standard decision levels and using the shifted prediction error PDF $p_{\hat{E}}(\hat{e}_n = \tilde{d}_n - \hat{d}_n^+)$, \hat{d}_n can be obtained by the centroid condition from (4) at each time index with $\hat{d}_n = d_l^{(i_n)}$. Initially, for time index $n = 0$, $\hat{d}_0^+ = 0$, the original standard quantized value is used as the received signal $\hat{d}_0 = d_{l_0}$.

Taking the scale factor y_n from the standard ADPCM decoder and extracting the sign from i_n , the new quantized difference signal

can be calculated according to $|\widehat{d}q_n| = 2^{\widehat{d}l_n + y_n}$. As mentioned in Section 2.2, the estimated signal se_n should be taken from the standard decoder, in consequence, the new reconstructed signal $\widehat{s}r_n$ can be obtained by $\widehat{s}r_n = \widehat{d}q_n + se_n$.

3.4. NLMS Algorithm

Using the NLMS algorithm, which is widely used for adaptive filtering [26], the *decoder-sided* predictor coefficients can be adaptively estimated by:

$$\mathbf{h}_n = \mathbf{h}_{n-1} + \frac{\widehat{e}l_{n-1}}{1 + \lambda \cdot \|\widehat{\mathbf{d}}\mathbf{l}_{n-N_p}^{n-1}\|^2} \cdot \widehat{\mathbf{d}}\mathbf{l}_{n-N_p}^{n-1}, \quad (7)$$

with $\|\widehat{\mathbf{d}}\mathbf{l}_{n-N_p}^{n-1}\|$ being the Euclidean norm. Instead of using both step-size and regularization parameters [26], only one tuning parameter λ is used [27, 31]. The initialization of \mathbf{h} and $\widehat{\mathbf{d}}\mathbf{l}$ for time index $n = 0$ are $\mathbf{h}_{-1} = (1/N_p, \dots, 1/N_p)^T$ and $\widehat{\mathbf{d}}\mathbf{l}_{-N_p}^{-1} = (0, \dots, 0)^T$.

4. SIMULATION RESULTS

4.1. Simulation Setup

In this paper, we investigate ADPCM with our new decoder operating at its 16 kbit/s mode. The NTT monaural speech database [29] with 8 kHz sampling rate is used in our simulations. All signals are normalized to -26 dBov. The perceptual evaluation of speech quality (PESQ) (P.862) [30] is used as an instrumental measure for the MOS. A better speech quality can be identified by a higher MOS value.

The core of the proposed scalar quantization decoding approach is to use a shifted prediction error PDF $p_{\widehat{E}}(\widehat{e}_n = \widetilde{d}l_n - \widehat{d}l_n^+)$ to generate a time-variant quantization codebook. In order to identify the optimal prediction error PDF $p_{\widehat{E}}(\cdot)$, we did experiments with both Laplacian and Gaussian PDFs by varying the prediction error variance $\sigma_{\widehat{e}}^2$. For each PDF, a full numerical search over the number range $0.5 < \sigma_{\widehat{e}} < 1.1$ in steps of 0.1 is performed in advance. A number of 96 English speech files spoken by American native speakers, which includes 4 male and 4 female speakers, each with length 8 s is exclusively used for the optimization. The optimal values $\widehat{\sigma}_{\widehat{e}}^{\text{opt}}$ and $\widehat{\mu}_{\widehat{e}}^{\text{opt}}$ are determined by the maximum MOS score of PESQ and are used in the decoder. $N_p = 10$ and $\lambda = 10000$ are adopted in the NLMS algorithm. It is found that a better improvement can be achieved by using a Gaussian PDF as the shifted prediction error PDF:

$$p_{\widehat{E}}(\widehat{e}_n = \widetilde{d}l_n - \widehat{d}l_n^+) = \frac{1}{\sqrt{2\pi}\widehat{\sigma}_{\widehat{e}}} \exp\left(-\frac{(\widetilde{d}l_n - \widehat{d}l_n^+ - \widehat{\mu}_{\widehat{e}})^2}{2\widehat{\sigma}_{\widehat{e}}^2}\right), \quad (8)$$

with $\widehat{\sigma}_{\widehat{e}}^{\text{opt}} = 0.9$ and the mean of the prediction error $\widehat{\mu}_{\widehat{e}} = 0$. Interestingly, it is found that the performance can be further slightly improved by varying the mean value $\widehat{\mu}_{\widehat{e}}$. Using $\widehat{\sigma}_{\widehat{e}}^{\text{opt}} = 0.9$, a full numerical search over the number range $-0.3 < \widehat{\mu}_{\widehat{e}} < 0.1$ in steps of 0.01 is performed to identify the optimal value $\widehat{\mu}_{\widehat{e}}^{\text{opt}} = -0.23$.

For testing, we use 20 different languages (excluding American English), each language with 96 speech files including 4 male and 4 female speakers.

In order to observe how well the NLMS-based prediction

Test Language	Standard Decoder	Proposed Decoder
English (British)	2.51	2.67
Chinese	2.44	2.61
French	2.46	2.62
German	2.48	2.61
Spanish	2.58	2.73
Average	2.56	2.71

Table 1: MOS results of PESQ using the standard ADPCM decoder and the proposed ADPCM decoder (16 kbit/s mode).

Test Language	G_p (dB)
English (British)	21.97
Chinese	22.80
French	21.82
German	20.89
Spanish	22.26
Average	22.65

Table 2: Prediction gain (in dB) using the NLMS algorithm to predict the normalized difference signal $\widetilde{d}l_n$ in the 16 kbit/s ADPCM decoder.

worked, the prediction gain is measured by

$$G_p = 10 \cdot \log_{10}\left(\frac{E\{\widehat{d}l_n^2\}}{E\{\widehat{e}l_n^2\}}\right), \quad (9)$$

with $E\{\cdot\}$ being the expectation value. For each language, $\frac{E\{\widehat{d}l_n^2\}}{E\{\widehat{e}l_n^2\}}$ is calculated upon each speech file considering all samples: the final result is obtained by the mean of the 96 (linear) G_p values and thereafter transformed into the dB domain.

4.2. Discussion

The PESQ MOS results of the standard ADPCM and our proposed decoding approach are shown in Table 1, with the explicit results of British English, Chinese, German, French and Spanish, and with the average value of all 20 languages. MOS gains in the range of 0.12 to 0.19 can be observed by using our proposed decoder. The average MOS of the 20 languages has been improved by 0.15. Moreover, as shown in Table 2, more than 20 dB prediction gain can be achieved, which implies the 10th-order NLMS-based predictor works reasonably well in our system. These high values are particularly remarkable, since in ADPCM already an ARMA predictor is employed. Therefore, we can conclude that our approach significantly and consistently outperforms the standard ADPCM decoding (even) in error-free transmission conditions.

It is noted that the optimization/training of the prediction error PDF is carried out with the American English database, but the optimal values can also be applied to other languages. We can state that the optimization is independent of language.

5. CONCLUSIONS

In this paper, we present an improved adaptive differential pulse code modulation (ADPCM) system using an improved scalar quantization

approach, which adopts an unmodified encoder and employs an additional predictor on the basis of the normalized least-mean-squares (NLMS) algorithm in the decoder. According to the standard decision levels and using the centroid condition, a new adaptive quantization codebook can be generated to be used only in the decoder. We apply this approach to the ADPCM decoder operating at 16 kbit/s and show that the mean opinion score (MOS) can be increased by about 0.15 points using our proposed approach in error-free transmission conditions for low bit rate. The proposed approach is applicable in a standard-compatible fashion and can straightforwardly be employed to improve New Generation DECT devices.

REFERENCES

- [1] R.M. Gray and D.L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.
- [2] S.P. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–136, Mar. 1982.
- [3] J. Max, "Quantizing for Minimum Distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.
- [4] W.B. Kleijn and R. Hagen, "On Memoryless Quantization in Speech Coding," *IEEE Signal Processing Letters*, vol. 3, no. 8, pp. 228–230, Aug. 1996.
- [5] N.S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1984.
- [6] N.S. Jayant, "Digital Coding of Speech Waveforms: PCM, DPCM, and DM Quantizers," *Proc. IEEE*, vol. 62, no. 5, pp. 611–632, May. 1974.
- [7] R. A. McDonald, "Signal-to-Noise and Idle Channel Performance of Differential Pulse Code Modulation Systems — Particular Applications to Voice Signals," *Bell System Technical Journal*, vol. 45, no. 7, pp. 1123–1151, Sep. 1966.
- [8] Y. Linde, A. Buzo, and R.M. Gray, "An Algorithm for Vector Quantizer Design," *IEEE Transactions on Communications*, vol. 28, no. 1, pp. 84–95, Jan. 1980.
- [9] P. Elias, "Predictive Coding—I," *IRE Transactions on Information Theory*, vol. 1, no. 1, pp. 16–24, Mar. 1955.
- [10] D.S. Arnstein, "Quantization Error in Predictive Coders," *IEEE Transactions on Communications*, vol. 23, no. 4, pp. 423–429, Apr. 1975.
- [11] J. Huang and P.M. Schultheiss, "Block Quantization of Correlated Gaussian Random Variables," *IEEE Transactions on Communications Systems*, vol. 11, no. 3, pp. 289–296, Sep. 1963.
- [12] P. Cummiskey, N. S. Jayant, and J. L. Flanagan, "Adaptive Quantization in Differential PCM Coding of Speech," *Bell System Technical Journal*, vol. 52, no. 7, pp. 1105–1118, Sep. 1973.
- [13] "ITU-T Recommendation G.726, 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)," ITU-T, Aug. 1990.
- [14] "ITU-T Recommendation G.722, 7 kHz Audio-Coding Within 64 kbit/s," ITU-T, Nov. 1988.
- [15] S. Han and T. Fingscheidt, "Improving Scalar Quantization for Correlated Processes Using Adaptive Codebooks Only At the Receiver," in *Proc. of EUSIPCO*, Lisbon, Portugal, Sep. 2014, pp. 386–390.
- [16] S. Han and T. Fingscheidt, "Scalar Quantization With Optimized Receiver-Sided Adaptive Codebook Reconstruction Levels Controlled by a Predictor," in *Proc. of 11th ITG Conference on Speech Communication*, Erlangen, Germany, Sep. 2014, pp. 47–50.
- [17] S. Han and T. Fingscheidt, "Lloyd-Max Quantization of Correlated Processes: How to Obtain Gains by Receiver-Sided Time-Variant Codebooks," in *Proc. of 10th International ITG Conference on Systems, Communications and Coding*, Hamburg, Germany, Feb. 2015, pp. 1–5.
- [18] "ETSI EN 300 175, Digital Enhanced Cordless Telecommunications Common Interface," ETSI, 1992.
- [19] "ETSI TS 102 527, Digital Enhanced Cordless Telecommunications New Generation DECT," ETSI, 2007.
- [20] J. Makhoul and M. Berouti, "Adaptive Noise Spectral Shaping and Entropy Coding in Predictive Coding of Speech," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, no. 1, pp. 63–73, Feb. 1979.
- [21] J. Chen and A. Gersho, "Adaptive Postfiltering for Quality Enhancement of Coded Speech," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 1, pp. 59–71, Jan. 1995.
- [22] V. Ramamoorthy, N.S. Jayant, R.V. Cox, and M.M. Sondhi, "Enhancement of ADPCM Speech Coding with Backward-Adaptive Algorithms for Postfiltering and Noise Feedback," *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 2, pp. 364–382, Feb. 1988.
- [23] C. Lee, "An Enhanced ADPCM Coder for Voice Over Packet Networks," *International Journal of Speech Technology*, vol. 2, no. 4, pp. 343–357, May. 1999.
- [24] T. Fingscheidt and P. Vary, "Softbit Speech Decoding: A New Approach to Error Concealment," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 3, pp. 240–251, Mar. 2001.
- [25] T. Fingscheidt, "Graceful Degradation in ADPCM Speech Transmission," in *Proc. of DAGA*, Aachen, Germany, Mar. 2003, pp. 748–749.
- [26] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, New Jersey, fourth edition, 2002.
- [27] F. Pflug and T. Fingscheidt, "Delayless Soft-Decision Decoding of High-Quality Audio Transmitted Over AWGN Channels," in *Proc. of ICASSP 2011*, Prague, Czech Republic, May. 2011, pp. 489–492.
- [28] A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, Dordrecht, London, 1992.
- [29] "Multi-Lingual Speech Database for Telephonometry," NTT-AT, 1994.
- [30] "ITU-T Recommendation P.862, Perceptual Evaluation of Speech Quality (PESQ)," ITU-T, Feb. 2001.
- [31] G.D.T. Schuller, B. Yu, D. Huang, and B. Edler, "Perceptual Audio Coding Using Adaptive Pre- and Post-filters and Lossless Compression," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 6, pp. 379–390, Sep. 2002.