JOINT SPEECH CODEC PARAMETER AND CHANNEL DECODING OF PARAMETER INDIVIDUAL BLOCK CODES (PIBC)

Tim Fingscheidt, Stefan Heinen†, Peter Vary†*

*AT&T Labs
Shannon Laboratory
Florham Park, NJ 07932, USA
tim@research.att.com

†Institute of Communication Systems and Data Processing
Aachen University of Technology, Templergraben 55
D-52056 Aachen, Germany
{heinen, vary}@ind.rwth-aachen.de

ABSTRACT

In digital mobile speech transmission usually the most important (class 1a) bits provided by the speech coding scheme are protected by a CRC for error detection. As a consequence *all* parameters spanned by the class 1a bits have to be marked at the receiver either as reliable or as unreliable. In contrast to this somewhat coarse approach we propose the usage of what we call parameter individual block codes (PIBC) for the most important codec *parameters*. This allows joint speech codec parameter and PIBC decoding taking advantage of the error concealing properties of softbit speech decoding [1, 2].

1. INTRODUCTION

Digital cellular systems like GSM or IS-136 provide a CRC for the subjectively most important (class 1a) bits of the speech encoded bit stream. The class 1a bits usually belong to several speech codec parameters, so the CRC evaluation gives a measure of *frame* reliability rather than *parameter* reliability, called *bad frame indicator* (BFI). The BFI – besides other information – is used to initiate error concealment mechanisms, if necessary. In [1] it has been shown that the technique of softbit (or soft input) speech decoding as a means of error concealment is able to outperform conventional frame repetition methods that are driven by a BFI. So far softbit speech decoding has not used the BFI at all but instead relied on the soft output of the channel decoder, or equalizer, or demodulator - whatever precedes the speech decoder.

Now we show how softbit speech decoding is able not only to exploit the implicit redundancy of a speech codec parameter but also explicit redundancy in terms of parity bits to a speech codec *parameter* (rather than *all class 1a bits*). Applied to the most important parameters this allows individual parameters of a frame to be subject to error concealment. Assuming that a parameter has been protected by a parameter individual block code (PIBC) we propose the optimum joint parameter (= source) and channel decoding algorithm in an MMSE sense. Two examples of PIBCs are given with one code protecting the parameter value itself rather than its bits.

This work was done while first author was with †.

2. THE JOINT PARAMETER AND CHANNEL DECODING ALGORITHM

Let us consider a speech codec parameter $\tilde{\underline{v}}_0 \in \mathbb{R}^P$ being vector quantized $Q[\tilde{\underline{v}}_0] = \underline{v}_0$ by M bits and transmitted as depicted in Fig. 1. Dependent on the type of codec parameter, the time index 0 refers to a frame or a subframe instant. Accordingly, negative time indices denote previous (sub)frames. A bit mapping scheme (BM) assigns M bits $\underline{x}_0 = \{x_0(0), x_0(1), ..., x_0(m), ..., x_0(M-1)\} = \underline{x}_0{}^{(i)}$ with $i \in \{0, 1, ..., 2^M - 1\}$ being the quantization table index to the quantized parameter $\underline{v}_0$. Then a parameter individual block code (PIBC) of rate $r = M/K$ is applied to $\underline{x}_0$ yielding $\underline{y}_0 = \underline{y}_0{}^{(k)}$ with $k \in \{0, 1, ..., 2^K - 1\}$.

After transmission over a noisy equivalent channel that may contain channel (de-)coding, (de-)modulation, and equalization schemes a bit combination $\hat{\underline{y}}_0$ is received that probably differs from $\underline{y}_0$. In addition, estimated bit error probabilities $\underline{p_e}_0 = \{p_{e0}(0), p_{e0}(1), ..., p_{e0}(m), ..., p_{e0}(M-1)\}$ are assumed to be known. If the last algorithm of the equivalent channel is a Viterbi decoder, the $\underline{p_e}_0$ values can be derived e.g. from a soft output Viterbi algorithm (SOVA) [1, 3].

The parameter decoder [1, 2] uses $\hat{\underline{y}}_0$ and $\underline{p_e}_0$ to compute channel dependent transition probabilities $P(\hat{\underline{y}}_0 \mid \underline{y}_0{}^{(k)})$, $k \in \{0, 1, ..., 2^K - 1\}$. Modelling the quantized codec parameter vector as Markov process a parameter a priori knowledge in terms of $P(\underline{x}_0{}^{(i)} \mid \underline{x}_{-1}{}^{(j)})$ with $i, j \in \{0, 1, ..., 2^M - 1\}$ (1st order Markov) or $P(\underline{x}_0{}^{(i)})$ (0th order Markov) can be exploited. It can be measured by applying a large speech database to the speech encoder and by counting how often (pairs of) different levels of the quantizer output occur. Furthermore, *a priori* knowledge about the block code is required. It is used in terms of $P(\underline{y}_0{}^{(k)} \mid \underline{x}_0{}^{(i)})$ which is a very sparse matrix containing mostly zeros and a single one per row as well as per column if conventional[1] block codes are used. Both types of a priori knowledge are combined with the channel dependent transition probabilities to compute 2^M a posteriori probabilities $P(\underline{x}_0{}^{(i)} \mid \hat{\underline{y}}_0, \hat{\underline{y}}_{-1}, ...)$. Finally, a parameter estimation procedure is performed to achieve the parameter value itself. For most parameters the MMSE

[1]The term *conventional* for our purpose means that the number of valid codewords equals 2^M.

0-7803-5651-9/99/$10.00 ©1999 IEEE.

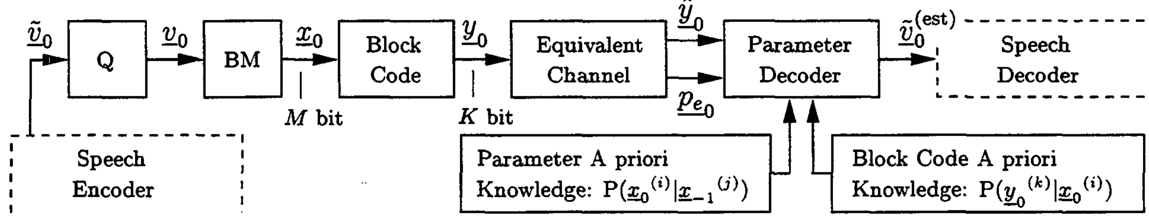

Figure 1: Block diagram of a speech codec parameter transmission: A PIBC scheme and its decoding

criterion is suitable in the sense that it reflects the decoded speech quality reasonably well. Therefore estimation can be done by

$$\tilde{\underline{v}}_0^{(\text{est})} = \sum_{i=0}^{2^M-1} \underline{v}^{(i)} \cdot P(\underline{x}_0^{(i)} \mid \hat{\underline{y}}_0, \hat{\underline{y}}_{-1}, \ldots) . \qquad (1)$$

Assuming the binary channel to be memoryless our proposal of joint speech codec parameter and channel decoding is performed by eq. (1) and the general recursion for the a posteriori probabilities

$$P(\underline{x}_0^{(i)} \mid \hat{\underline{Y}}_0) = \qquad (2)$$

$$C \cdot P(\underline{x}_0^{(i)} \mid \hat{\underline{Y}}_{-1}) \cdot \sum_{k=0}^{2^K-1} P(\hat{\underline{y}}_0 \mid \underline{y}_0^{(k)}) \cdot P(\underline{y}_0^{(k)} \mid \underline{x}_0^{(i)}) \quad \text{with}$$

$$P(\underline{x}_0^{(i)} \mid \hat{\underline{Y}}_{-1}) = \sum_{j=0}^{2^M-1} P(\underline{x}_0^{(i)} \mid \underline{x}_{-1}^{(j)}) \cdot P(\underline{x}_{-1}^{(j)} \mid \hat{\underline{Y}}_{-1})$$

with $\hat{\underline{Y}}_n = \hat{\underline{y}}_n, \hat{\underline{y}}_{n-1}, \ldots$ and the constant C normalizing the sum over the a posteriori probabilities to one. If a systematic block code is applied, i.e. $\underline{y}_0^{(i)} = (\underline{x}_0^{(i)}, \underline{z}_0^{(l)})$ with $\underline{z}_0^{(l)}$, $l \in \{0, 1, \ldots, 2^L - 1\}$ being L explicit parity bits, then recursion (2) becomes

$$P(\underline{x}_0^{(i)} \mid \hat{\underline{X}}_0, \hat{\underline{Z}}_0) = C \cdot P(\underline{x}_0^{(i)} \mid \hat{\underline{X}}_{-1}, \hat{\underline{Z}}_{-1}) \cdot \qquad (3)$$

$$\cdot P(\hat{\underline{x}}_0 \mid \underline{x}_0^{(i)}) \cdot \sum_{l=0}^{2^L-1} P(\hat{\underline{z}}_0 \mid \underline{z}_0^{(l)}) \cdot P(\underline{z}_0^{(l)} \mid \underline{x}_0^{(i)}) .$$

If conventional non-systematic block codes are used, recursion (2) simplifies to

$$P(\underline{x}_0^{(i)} \mid \hat{\underline{Y}}_0) = C \cdot P(\underline{x}_0^{(i)} \mid \hat{\underline{Y}}_{-1}) \cdot P(\hat{\underline{y}}_0 \mid \underline{y}_0^{(i)}) \qquad (4)$$

with $\underline{y}_0^{(i)}, i \in \{0, 1, \ldots, 2^M - 1\}$ being the valid codewords assigned to $\underline{x}_0^{(i)}$. Furthermore, if the conventional block code is systematic, then recursion (4) reads

$$P(\underline{x}_0^{(i)} \mid \hat{\underline{X}}_0, \hat{\underline{Z}}_0) =$$

$$C \cdot P(\underline{x}_0^{(i)} \mid \hat{\underline{X}}_{-1}, \hat{\underline{Z}}_{-1}) \cdot P(\hat{\underline{x}}_0 \mid \underline{x}_0^{(i)}) \cdot P(\hat{\underline{z}}_0 \mid \underline{z}_0^{(i)}) . \qquad (5)$$

This recursion was recently published by Görtz as an extension of the softbit speech decoding technique to include parameter individual, conventional, systematic codes [4].
If the quantized parameter is modelled as 0th order Markov process the terms $P(\underline{x}_0^{(i)} \mid \hat{\underline{Y}}_{-1})$ or $P(\underline{x}_0^{(i)} \mid \hat{\underline{X}}_{-1}, \hat{\underline{Z}}_{-1})$ in the above recursions are simply to be replaced by the 0th order parameter a priori knowledge $P(\underline{x}_0^{(i)})$.

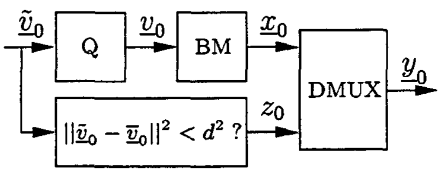

Figure 2: Signal driven PIBC generation

3. PARAMETER INDIVIDUAL BLOCK CODES

In this section we propose two simple parameter individual block codes (PIBCs).
The first one is a simple even parity code (PIBC1) of rate $r = M/K$ with $K = M + 1$. As it is a conventional systematic code, recursion (5) applies.
The second code (PIBC2) is also a systematic code of rate $r = M/K$, $K = M + 1$, but the parity bit is dependent on the unquantized parameter $\tilde{\underline{v}}_0$ rather than on the bit combination $\underline{x}_0$ of the quantized parameter $\underline{v}_0$. In contrast to the transmitter shown in Fig. 1 this code is generated as depicted in Fig. 2. The parity bit $y_0(M) = z_0$ is set to one if the Euclidean distance of $\tilde{\underline{v}}_0$ to its mean $\overline{\underline{v}}_0$ is below a certain threshold d, otherwise it is set to zero. The threshold d provides a degree of freedom in the design of this *signal driven* PIBC. The value of d can be optimally adjusted via simulation over a noisy channel model. The motivation for this kind of *parity* bit generation is that z_0 is expected to support the muting algorithm[2] at the decoder side. If decoding is performed by equations (1) and (3) then the muting effect of the minimum mean square estimation [2] is expected to be enhanced.
In general, the PIBC2 provides more than 2^M valid codewords because – e.g. in the case of a scalar quantizer – d must not necessarily fall on a quantizer decision level. Thus in the scalar case $2^M + 2$ valid codewords $\underline{y}_0^{(k)}$ exist what can be seen by the number of non-zero entries in $P(\underline{z}_0^{(l)} \mid \underline{x}_0^{(i)})$, $l \in \{0, 1\}$, (see e.g. eq. (7)).

4. SIMULATION RESULTS

To prove the capabilities of PIBC1 and PIBC2 when decoded using recursion (5) and (3), respectively, we performed simulations based on M bit scalar Lloyd-Max quantization of a unit variance, Gaussian, uncorrelated parameter. Although the bit mapping has influence on the de-

[2]Muting as a conventional means of error concealment pulls a (repeated) codec parameter toward its mean once several consecutive bad frames have been detected.

coding results (see e.g. [5]) we didn't focus on that in order to keep the PIBC design for a specific codec parameter as simple as possible. So we just assumed natural binary coding (NBC) of the parameter. The equivalent channel is an AWGN channel with BPSK modulation and coherent demodulation. The soft output of the channel is computed as given in [2]. The decoding quality is defined in terms of the parameter SNR

$$\mathrm{SNR} \ = 10 \log_{10} E\{\tilde{v}^2\}/E\{(\hat{v} - \tilde{v})^2\}$$

with $\tilde{v}$ and $\hat{v}$ being the unquantized and the decoded scalar parameters, respectively. The channel quality is given in terms of E_s/N_0 with E_s denoting the energy of a gross bit and $N_0/2$ being the power spectral density of the additive Gaussian noise.

The solid curves in Fig. 3 show simulation results to an $M =$ 3 bit quantization [6]. Simple hardbit decoding (HD, i.e. decoding via table lookup without any error concealment) yields the worst performance; the quality degrades rapidly with decreasing E_s/N_0. Employing softbit decoding (SD) the quality is enhanced. Due to the fact that we do not assume residual correlations of the codec parameter softbit decoding is performed by estimation using eq. (1) with the a posteriori probabilities

$$\mathrm{P}(\underline{x}_0{}^{(i)} \mid \hat{\underline{x}}_0) = C \cdot \mathrm{P}(\hat{\underline{x}}_0 \mid \underline{x}_0{}^{(i)}) \cdot \mathrm{P}(\underline{x}_0{}^{(i)}) \ . \qquad (6)$$

However, if an additional bit is spent to perform systematic block coding, the quality can be significantly enhanced. Recall that due to the uncorrelated codec parameter the $\mathrm{P}(\underline{x}_0{}^{(i)} \mid \underline{\hat{X}}_{-1}, \underline{\hat{Z}}_{-1})$ term in recursions (3) and (5) is replaced by the parameter a priori knowledge $\mathrm{P}(\underline{x}_0{}^{(i)})$.

If PIBC1 according to Fig. 1 is used and decoding is performed by eq. (5) E_s/N_0 gains are about 1 dB vs. softbit and at least 1.5 dB vs. hardbit decoding.

If PIBC2 according to Fig. 2 with a threshold value of $d =$ 1.1 is used then the measured code a priori knowledge is

$$\mathrm{P}(z_0^{(l)} \mid \underline{x}_0{}^{(i)}) = \begin{bmatrix} 1 & 0.8953 & 0 & 0 & 0 & 0 & 0.8953 & 1 \\ 0 & 0.1047 & 1 & 1 & 1 & 1 & 0.1047 & 0 \end{bmatrix} \quad (7)$$

with $l \in \{0, 1\}$. Every column of this matrix belongs to a quantized parameter, while in the 2nd and the 7th quantization interval the parity bit $z_0^{(l)}$ is only determined by $\tilde{v}_0$ but not by $\underline{x}_0{}^{(i)}$, because the threshold d lies within these intervals. Thus decoding has to be performed by eq. (3). The usage of such a signal driven parity bit leads to further gains of about 0.7 dB. Thus PIBC2 turns out to be preferable to the conventional parity code PIBC1.

Recall that in Fig. 3 the two solid curves at the left represent a bit rate of 4 bit/parameter while the two solid curves at the right represent 3 bit/parameter. If the Gaussian parameter is quantized with 4 bits (the two dashed curves) to yield the same gross bit rate as the PIBC schemes, the clear channel quality becomes better. However, if an SNR of about 14 dB is the required signal quality, 3 bit quantization with PIBC2 coding remains, even in that comparison, the best scheme.

If the codec parameter is correlated, recursions (2), (3), (4), or (5) can be applied leading to even more gain relative to the HD curve. The technique of parameter individual block coding has already successfully been used in a recent speech codec proposal [7].

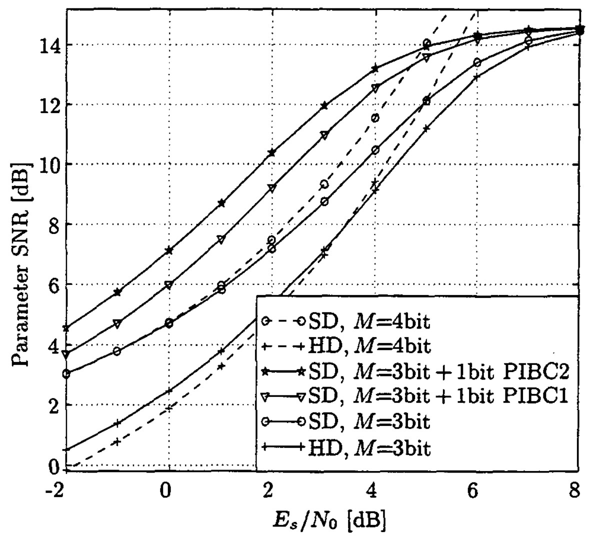

Figure 3: SNR performance of several schemes:
HD: harbit decoding, **SD**: softbit decoding, **PIBC1**: even parity code (as in [4]), **PIBC2**: signal driven code

5. CONCLUSIONS

Conventionally, a CRC over important bits of several speech codec parameters is used for the purpose of error detection. In contrast to that we showed that a simple additional bit attached to an important speech codec parameter in conjunction with an appropriate joint parameter and channel decoding scheme is able to provide a powerful means for combined error detection and concealment.

6. REFERENCES

[1] T. Fingscheidt and P. Vary, "Speech Decoding With Error Concealment Using Residual Source Redundancy," in *Proc. of IEEE Workshop on Speech Coding*, Pocono Manor, Pennsylvania, Sept. 1997, pp. 91–92.

[2] 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.

[3] J. Hagenauer and P. Hoeher, "A Viterbi Algorithm with Soft-Decision Outputs and its Applications," in *Proc. of GLOBECOM'89*, Dallas, Texas, 1989, pp. 1680–1686.

[4] N. Görtz, "How to Combine Forward Error Detection and Bit-Reliability Information in Source Decoding," in *Proc. of IEEE International Symposium on Information Theory*, MIT, Massachusetts, Aug. 1998.

[5] N. Farvardin and V. Vaishampayan, "Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source-Channel Coding," *IEEE Transactions on Inf. Theory*, vol. 33, no. 6, pp. 827–838, Nov. 1987.

[6] B. Dortschy, "Entwurf und Optimierung von Algorithmen zur Softbit-Sprachdecodierung im Mobilfunk," Diploma thesis, Institute of Communication Systems and Data Processing, Aachen University of Technology, Germany, Sept. 1998.

[7] S. Heinen, M. Adrat, O. Steil, P. Vary, and W. Xue, "A 6.1 to 13.3 kbit/s Variable Rate CELP Codec (VR-CELP) for AMR Speech Coding," in *Proc. of ICASSP'99*, Phoenix, Arizona, Mar. 1999.