A 2.1 kBd Speech Transmission System for Rayleigh-Fading Channels

L. Hanzo, R. Salami, R. Steele

1 Introduction

In the global microcellular mobile network of the near future extremely high traffic densities have to be carried. The well-known GSM system is capable of handling a traffic density of typically 10 erlang/MHz/km^2 among adverse channel conditions in the large, maximum 35 km diameter traffic cells. In the future global microcellular mobile network the initiated traffic densities can be as much as three orders of magnitude higher, whence extreme spectral efficiency is required. Also, the frequency re-use distances have to be reduced to enable handling of the initiated traffic. This results into a microcellular structure with favourable propagation characteristics. In this situation the channel can generally be described by a Rician model, although in the worst-case scenario the flat Rayleigh-fading model applies. As a consequence, low bitrate speech codecs with moderate error correction coding and multi-level modulation schemes without channel equalisers can be deployed to achieve the required bandwidth efficiency.

2 The Speech Codec

Constrained by the extreme spectral efficiency demand we propose to deploy a low-complexity 4.8 kbit/s transformed code excited linear predictive (**CELP**) codec [1]. A strong justification to advocate the CELP codec is that at 4.8 kbit/s transmission rate it proved to be most promising candidate for the American Department of Defence (DOD) standard [2]. It is also the most likely solution to the half-rate GSM codec [3] and to the 16 kbit/s low-delay CCITT speech coding standard.

In our CELP codec 36 bits are allocated to the scalar quantisation of ten line spectrum frequencies (**LSFs**), which represent the speech spectral envelope for a 30 ms frame. This frame is divided into four subsegments of 60 samples. Sparse transformed duo-binary excitation samples are used with a decimation factor of D=5, which helps to keep the complexity low. The long-term predictor (**LTP**) delays are encoded using seven or five bits per subsegment, while the LTP gain is quantised with three bits. The gridposition of the regularly spaced candidate excitation sequence is represented by two bits, the codebook gain by four bits and the excitation samples by 12 bits. Therefore in a subsegment on the average 27 bits are utilised for quantisation and the total number of bits to be transmitted in a 30 ms interval is 36+4x27=144, which results into a transmission rate of 4.8 kbit/s.

The bits of the CELP codec have varying bit error sensitivities both in objective terms expressed in segmental signal to noise ratio (**SEG-SNR**) or cepstral distance (**CD**) as well as perceptually. The CD measure is more appropriate to evaluate the sensitivity of the LSF parameters, while the SEG-SNR to assess that of the excitation parameters. On the basis of their sensitivities the bits are assigned into two classes (**CI&CII**) of different subjective and objective importance. Then appropriate embedded forward error correction (**FEC**) coding is applied to both classes of the encoded speech bits, as we highlight in the following section.

3 The Channel Codec

In the mobile environment both convolutional and block FEC codes can successfully be deployed, if they are carefully matched to the surrounding system environment [4].

L. Hanzo, R. Salami and R. Steele are with the Dept. of Electr., Univ. of Southampton

The power of convolutional codes comes true, if they are Viterbi-decoded using soft-decision channel measurement information and the interleaving depth is sufficiently high to render the originally bursty channel's error statistics to become random. Then the higher is the code constraint-length, viz. the number of states in the Viterbi-decoder, the better is the performance, while the complexity increases exponentially. This is one factor, limiting the correcting power of convolutional codes.The other probelm is that in speech trasnmission systems the interleaving depth is rather limited by the low delay requirements imposed. Finally, there is no confident indication, whether the convolutional codec succeeded to correct the received sequence.

Block codes lend themselves to deployment via both Gaussian and Rayleigh-fading channels. Due to their higher blocklength they inherently allow for some degree of freedom in the distribution of channel errors. The higher the blocklength, the larger is this freedom in the error-distribution. By increasing the codeword length the ideal situation can be reached when the signal envelope fades can be overbridged by the codewords and thence we are faced with a system, where the FEC codec's memory-length matches that of the channel. The price we have to pay is the increased FEC codec complexity and delay. Nearly equivalent performance can be achieved when deploying interleaving and using shorter FEC block codes, having lower decoding complexities.

The encoders and decoders of cyclic block codes can be easily implemented using simple linear shift-register circuitries, whence they are important in the FEC coding theory. A prominent class of cyclic block codes over tha mathematical construction of finite Galois Fields (**GFs**) is constituted by the Bose-Chaudhuri-Hocquenghem (**BCH**) codes. An important non-binary subclass of BCH codes with optimum distance properties is that of the **Reed-Solomon** (**RS**) codes, having identical symbol field and error locator field. Although RS codes possess maximum minimum distance properties, their performance is often matched by some binary BCH codes, having considerably lower decoding complexities [5]. The error correcting power at a given coding rate and codeword length is increased by deploying trellis decoding with soft-decisions [5], but the decoding complexity is prohibitive at state-of-art technology in the codeword length domain of sufficient correcting power.

In summary of the FEC coding aspects, when low decoding complexity associated with high error correcting power is required and no soft-decision information is used, binary BCH codes with moderate blocklengths have favourable properties, whence in this system we propose to utilize BCH codes. The bit error rate (**BER**) performance versus channel **signal** to **noise** ratio (**SNR**) is displayed for a set of BCH codes in Figure 1, for a Rayleigh-fading channel, where Minimum Shift Keying (**MSK**) modulation is used, which is as robust as **binary phase shift** keying (BPSK) at double spectral efficiency. The propagation frequency is 900 MHz and the vehicular speed is 30 mph. We use the BER curves in Figure 1 to select the appropriate BCH code for our speech transmission system after considering aspects of the modulation scheme preferred. Finally, we emphasize that with the proviso of reasonable error correcting power BCH codes also boast of reliable error detection capability. This can be advantageously exploited to invoke speech post-processing algorithms [6] or to control handovers [7].

4 The Modulation Scheme

The high traffic density demand requires not only a low bitrate speech codec but also a highly bandwidth efficient modulation scheme. Although efficient multi-level modulation methods require usually comparatively high channel SNRs, this requirement is readily met in the microcellular environment. Encouraged by our experiments [8] we propose to utilise **16**-level quadrature amplitude **modulation** (**16-QAM**).

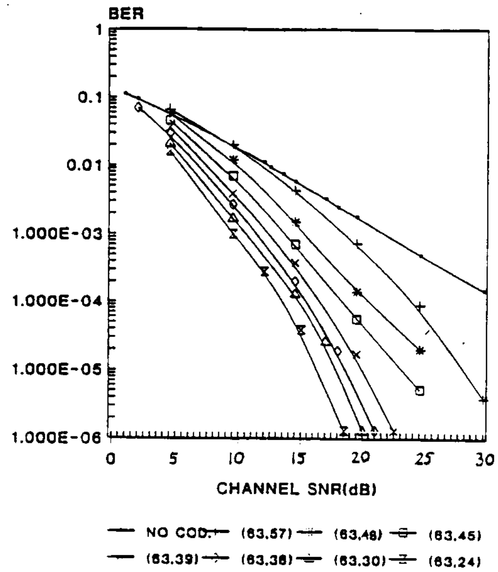

Figure 1. BER performance of various BCH codes

In the Gray-coded 16-QAM signal constellation each phasor is represented by a four-bit symbol and the position of a bit is crucial, as far as its error probability is concerned [8]. To be specific, the first two bits, which we refer to as class I (CI) bits, are less likely to be erroneous than the second two, referred to as class II (CII) bits. This holds for both additive white Gaussian noise(AWGN) and Rayleigh-fading channels, but the BER differences are more dramatic for the Rayleigh-fading channel. This is due to the fading nature of the channel, which effect can be mitigated by deploying a fade-tracking automatic gain control (AGC). Interestingly, the fade-tracking AGC does not effect the BER performance of the better CI subchannel but improves that of the poorer CII subchannel. Then we have two modulation subchannels with different integrities that can be exploited in our combined system.

5 The 2.1 kBd System

The proposed combined system's block diagram is shown in Figure 2. The 4.8 kbit/s CELP codec's output sequence is mapped, using the 'CELP MAP' block, into the CI and CII subgroups according to their subjective and objective importance. Both the CI and CII subchannels are protected by BCH error correction coding, the correcting power of which is carefully matched to the required integrity in both subchannels and to that of the 16-QAM subchannels. In other words, the BCH codecs with their associated CI and CII 16-QAM subchannels constitute a pair of superchannels for transmitting the CI and CII CELP bits. After careful considerations the BCH(63,48,2) code has been chosen to protect the CI CELP bits when transmitting through the CI 16-QAM subchannel. In the more hostile CII 16-QAM subchannel the stronger BCH(63,24,7) code has been utilised to provide a sufficiently robust superchannel for the transmission of the CII CELP bits.

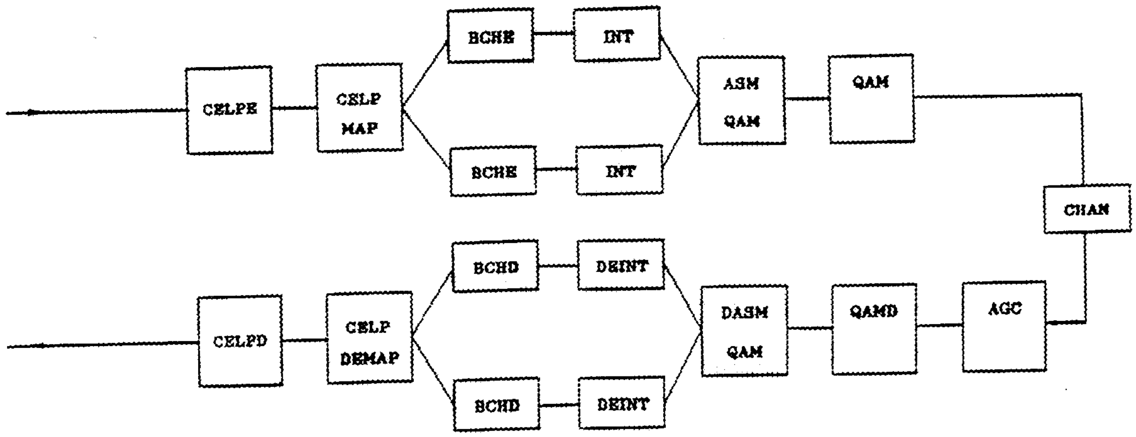

Figure2. The 2.1 KBd system

The selection of the BCH codes is also matched with the channel capacity of both 16-QAM subchannels. Namely, both subchannels can carry the same number of bits, therefore we have to ensure that after BCH encoding this condition is met. Furthermore, the BCH coded words should preferably protect an integer number of CELP frames, so that in case of incidental BCH decoding errors these are confined to as low a number of CELP frames as possible. This also helps to harmonize the choice of the proper interleaver.

Specifically, the CELP codec delivers 144 bits in each 30 ms interval, which are then split into 96 CI bits and 48 CII bits. Both of these constitute then a pair of BCH(63,48,2) and BCH(63,24,7) codewords, respectively. Then rectangular interleaving over four such codewords, viz. two consecutive CELP frames is carried out in the block 'INT' in Figure 2, which results into a total of 60 ms speech delay.

The BCH encoded and interleaved sequences are assembled in the block 'ASM QAM' for modulation in the block 'QAM'. This signal is then transmitted over the channel, which is in best case Gaussian, while in the worst situation of Rayleigh-fading type. After AGC and 16-QAM demodulation the bitstream is disassembled, deinterleaved and BCH decoded. Finally, the bits are demapped into their original positions and CELP decoded to recover the original speech signal.

The transmission rate is computed as follows: The 4.8 kbit/s CELP bit sequence is split into a 1.6 kbit/s CII and a 3.2 kbit/s CI subgroup. After BCH(63,48,2) and BCH(63,24,7) encoding the transmission rate in both subchannels is 4.2 kbit/s. Whence the total transmission rate is 8.4 kbit/s, viz. 2.1 kBd.

6 Results and Conclusion

As overall speech quality measure the speech SEG-SNR between the original and processed speech is used. The best and worst possible channel conditions are represented by the Gaussian and Rayleigh-fading channel, respectively. The performance for the Gaussian channel actually represents the stationary vehicle scenario, where there is no fading and the physical position of the receiver determines the actual SNR value.

The speech SEG-SNR versus channel SNR performance has been depicted in Figure 3 for the AWGN channel, as well as for the Rayleigh-fading channel in case of 3.75 mph. Observe the speech quality degradation as the mobile speed is reduced to 3.75 mph. This is due to the fact that at lower speeds the mobile is idleing too long in a fade. In this situation the limited-depth interleaving cannot randomise the error statistics and the BCH decoder is often overloaded. Best performance is achieved, however, in the sationary scenario, when there are no fades and the speech quality is determined solely by the SNR value the mobile experiences.

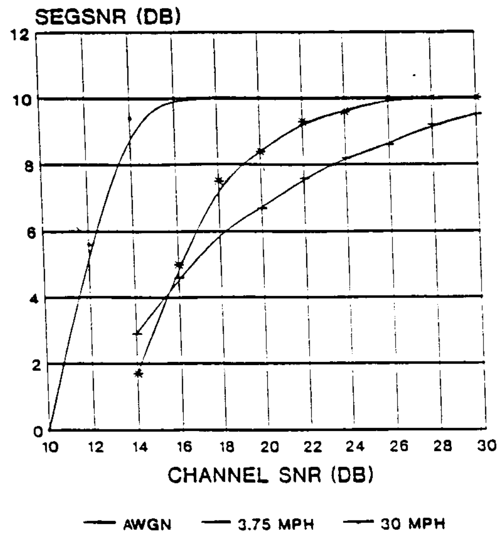

Figure 3. SEGSNR of the 2.1 KBd system via AWGN and Rayleigh channels

In summary, the proposed 2.1 kBd system guarantees extrem bandwidth efficiency in return for the favourable channel conditions experienced in the microcellular environment.

References

[1] R. Salami, K.H.H. Wong, D. Appleby, R. Steele: A robust transformed binary vector excited coder with embedded error correction coding, IEE Colloquium, 9. Oct. 1989, Savoy Place, London

[2] D.P. Kemp, R.A. Sueda, T.E. Tremain: An evaluation of 4800 bps voice coders, Proc. of ICASSP'89, pp 200-203

[3] Proc. of the Second EURASIP Workshop on Medium to Low-Rate Speech Coding, 20-22. Sept. 1989, Hersbruck, West-Germany

[4] L. Hanzo, K.H.H. Wong, R. Steele: Efficient channel coding and interleaving schemes for mobile radio communications, Proc. of IEE Colloq., 22 Febr. 1989, Savoy Place, London, U.K.

[5] K.H.H. Wong, L. Hanzo, R. Steele: Channel coding for satellite mobile channels, Int. Journal of Satellite Communications, Vol. 7, pp 143-163, 1989, John Wiley LTD

[6] K.H.J. Wong, L. Hanzo, R. Steele: A subband codec with embedded Reed-Solomon coding for mobile radio speech communication, Proc. of IEEE ICCS'88, pp 709-713

[7] R. Steele, D. Twelves, L. Hanzo: Effect of cochannel interference on handover in microcellular mobile radio, accepted by Electronics Letters, 1989

[8] L. Hanzo, R. Steele, P.M. Fortune: A subband coding, BCH coding and 16-QAM system for mobile radio speech communications, submitted to IEEE Vehicular Technology, 1989