

# A Low-Rate Multi-Level Voice/Video Transceiver for Personal Communications\*

L. HANZO, J. STREIT, R. A. SALAMI<sup>1</sup> and W. T. WEBB<sup>2</sup>

*Department of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK  
(Tel.: +44 1703 593 125; Fax: +44 1703 593 045; E-mail: lh@ecs.soton.ac.uk, http://www.ecs.soton.ac.uk)*

<sup>1</sup>*Current address: University of Sherbrook, Department of Electrical Engineering, Quebec, Canada;* <sup>2</sup>*Current address: Smith Associates, UK*

**Abstract.** A personal communication system (PCS) transceiver is proposed and investigated. A 4.8 kbit/s transformed binary pulse excited (TBPE) linear predictive speech codec, embedded source sensitivity-matched binary Bose–Chaudhuri–Hocquenghem (BCH) block error correction codecs, non-coherent differentially coded 16-level quadrature amplitude modulation (16-QAM) modem and packet reservation multiple access (PRMA) are deployed. The 2.15 kBd transceiver requires a signal-to-noise ratio (SNR) and signal-to-interference ratio (SIR) in excess of about 24 dB over Rayleigh-fading channels in order to support 10–11 nearly un-impaired voice conversations within a bandwidth of 30 kHz. Additionally, by reserving two PRMA time slots for video telephony, an 8.52 kbps videophone user can also be supported.

**Key words:** Wireless multimedia communications, mobile voice and video telephony, transformed binary pulse excited speech coding, multilevel modulation, packet reservation multiple access, the Pan-American IS-54 mobile system

## 1. Introduction

While the second generation digital mobile radio systems are being deployed in Europe, throughout the Pacific Rim and the United States, researchers turned their attention towards the true personal communication system (PCS) of the near future. Quite recently a special issue of the *Communications Magazine* was dedicated to general aspects of PCSs [1], covering a variety of existing standardised schemes, networking problems, future modulation arrangements as well as multimedia schemes. Combined modulation and coding was the feature topic of references [2, 3]. A plethora of subsystem details on speech codecs, channel codecs, modems and multiple access can be found in recent issues of the *IEEE Transactions on Communications and Vehicular Technology*, in particular for example in the August 1994 Special Issue on Personal Communications. The November, 1993 issue of the *Communications Magazine* was focused on general aspects of speech processing with three contributions considering speech codecs for various applications, while the paper by R. Steele [4] analysed the pertinent PCS trade-offs and gave a rudimentary comparison of existing wireless systems.

In this contribution we turned our attention to specific algorithmic and system aspects of a complete voice/video phone transceiver suitable for the future third generation PCS and evaluated its expected performance. The system bandwidth was assumed to be 30 kHz, as in the American IS-54 standard system, which allowed us to assess the potential of the proposed scheme in comparison to a well known benchmarker. The paper is organized as follows. Section 2 describes the proposed 4.8 kbit/s transformed binary pulse excited (TBPE) speech

\* The constructive criticism of the anonymous Reviewers is gratefully acknowledged.

codec, which is followed by a discussion on the proposed video codec in Section 3, while Section 4 is focused on the choice of modulation, in particular on 16-level quadrature amplitude modulation (16-QAM). Forward error correction coding (FEC) is considered in Section 5, while Section 6 highlights, how packet reservation multiple access (PRMA) improves the efficiency of the radio link by surrendering passive time slots for active users contending for an available slot. The proposed transceiver scheme is described in Section 7 and the system performance is analysed in Section 8, before some conclusions are drawn in Section 9.

## 2. The 4.8 kbit/s Speech Codec

In code excited linear predictive (CELP) codecs a Gaussian process with slowly varying power spectrum is used to represent the residual signal after short-term and long-term prediction, and the speech waveform is generated by filtering Gaussian excitation vectors through the time-varying linear pitch and LPC synthesis filters [5]. The Gaussian excitation vectors of dimension  $N$  are stored in a codebook of typically 1024 entries and the optimum excitation sequence is determined by the exhaustive search of the excitation codebook. The codebook entries  $c_k(n)$ ,  $k = 1 \dots L$ ,  $n = 0 \dots N - 1$ , after scaling by a gain factor  $G_k$ , are filtered through the pitch synthesis filter  $1/P(z)$  and the error weighting filter  $W(z) = 1/A(z/\gamma)$  to produce the weighted synthetic speech  $\tilde{s}_w(n)$ , which is compared to the weighted original speech  $s_w(n)$ .

Let  $x(n)$  be the weighted original speech after removing the memory contribution of the concatenated pitch synthesis and error weighting filters  $W(z) \cdot 1/P(z)$  from previous frames and  $h(n)$  be the impulse response of the weighting filter  $W(z)$ . Then the mean squared weighted error (mswe) between the original and synthesized speech is given by:

$$E = \sum_{n=0}^{N-1} [x(n) - G_k c_k(n) * h(n)]^2. \quad (1)$$

Setting  $\partial E / \partial G_k = 0$  leads to the minimum mean square weighted error expression [6]:

$$E_{min} = \sum_{n=0}^{N-1} x^2(n) - \frac{\left[ \sum_{i=0}^{N-1} \psi(i) c_k(i) \right]^2}{\sum_{i=0}^{N-1} c_k^2(i) \phi(i, i) + 2 \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} c_k(i) c_k(j) \phi(i, j)}, \quad (2)$$

where  $\psi(i)$  represents the correlation between the weighting filter's impulse response  $h(n)$  and the signal  $x(n)$ , given by  $\psi(n) = x(n) * h(-n)$ , while  $\phi(i, j)$  represents the covariances of  $h(n)$ :

$$\phi(i, j) = \sum_{n=0}^{N-1} h(n-i) h(n-j). \quad (3)$$

The best innovation sequence is constituted by that codebook entry  $c_k$  with index  $k$  ( $k = 1 \dots L$ ), which minimises the mean squared weighted error in Equation 2.

Since  $\psi(n)$  and  $\phi(n)$  are computed outside the error minimisation loop, the computational complexity is predetermined by the number of operations needed to evaluate the second term of Equation 2 for all the codebook entries. For a typical excitation frame length of  $N = 40$  and codebook size  $L = 1024$ , the CELP complexity becomes excessively high for real-time implementation. In recent years a plethora of efficient solutions have been suggested [8–10]

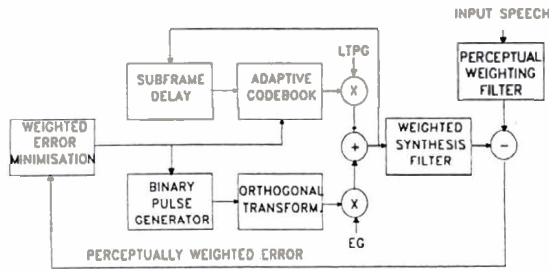


Fig. 1. Block diagram of the 4.8 kbit/s TBPE codec

in order to ease the computational load encountered, while still maintaining perceptually high speech quality.

In our PCS transceiver we favour a transformed binary pulse excited (TBPE) speech codec [8]. The great attraction of TBPE codecs when compared to CELP codecs accrues from the fact that the excitation optimisation can be achieved in a direct computation step [6, 8]. The sparse Gaussian excitation vector is assumed to take the form of

$$c = Ab, \quad (4)$$

where the binary vector  $b$  has  $M$  elements of  $\pm 1$ , while the  $M \cdot M$  matrix  $A$  represents an orthogonal transformation. Due to the orthogonality of  $A$  the binary excitation pulses of  $b$  are transformed into independent, unit variance Gaussian components of  $c$ . The set of  $2^M$  binary excitation vectors gives rise to  $2^M$  Gaussian vectors of the original CELP codec.

The block diagram of the TBPE codec is shown in Fig. 1. The weighted synthetic speech is generated for all  $2^M = 1024$  codebook vectors and subtracted from the weighted input speech in order to find the one resulting in the best synthesized speech quality. The weighted synthesis filter is excited by the vectors given by the superposition of the adaptive codebook vector scaled by the long term predictor gain (LTPG) and the orthogonally transformed binary pulse generator scaled by the excitation gain (EG).

The direct excitation computation of the TBPE codec accrues from the matrix representation of Equation 2 using Equation 4, viz.:

$$E = x^T x - \frac{(\Psi^T A b)^2}{b^T A^T \Phi A b}. \quad (5)$$

The denominator in Equation 5 is nearly constant over the entire codebook and hence plays practically no role in the excitation optimisation. This is due to the fact that the autocorrelation matrix  $\Phi$  is strongly diagonal, since the impulse response  $h(n)$  decays sharply. Due to the orthogonality of  $A$  we have  $A^T A = I$ , where  $I$  is the identity matrix, causing the denominator to be constant.

Closer scrutiny of Equation 5 reveals that its second term reaches its maximum if the binary vector element is given by  $b(i) = -1$ , whenever the vector element  $\Psi^T A$  is negative, and vica-versa, i.e.,  $b(i) = +1$  if  $\Psi^T A$  is positive. The numerator of Equation 5 is then constituted by exclusively positive terms, i.e., it is maximum, and the weighted mean squared error is minimum. The optimum Gaussian excitation is computed from the binary vector  $b$  using Equation 4 in both the encoder and decoder. Only the  $M$ -bit index representing the optimum binary excitation vector  $b$  has to be transmitted. The evaluation of the vectors  $\Psi^T A$  and  $c = Ab$  requires  $2M^2$  number of multiplications/additions, which gives typically 5

*Table 1.* Bit Allocation of 4.8 Kbit/s TBPE Codec

| Parameter  | Bitnumber               |
|------------|-------------------------|
| 10 LSFs    | 36                      |
| LTPD       | $2 \cdot 7 + 2 \cdot 5$ |
| LTPG       | $4 \cdot 3$             |
| GP         | $4 \cdot 2$             |
| EG         | $4 \cdot 4$             |
| Excitation | $4 \cdot 12$            |
| Total      | 144/20 ms               |

combined operations per output speech sample, a value 400 times lower than the complexity of the equivalent quality CELP codec.

The bit allocation of our TBPE codec is summarised in Table 1. The spectral envelope is represented by ten line spectrum frequencies (LSFs) which are scalar quantised using 36 bits. The 30 ms long speech frames having 240 samples are divided into four 7.5 ms subsegments having 60 samples. The subsegment excitation vectors  $b$  have 12 transformed duo-binary samples with a pulse-spacing of  $D = 5$ . The long term predictor (LTP) delays (LTPD) are quantised with seven bits in odd and five bits in even indexed subsegments, while the LTP gain (LTPG) is quantised with three bits. The excitation gain (EG) factor is encoded with four bits, while the grid position (GP) of candidate excitation sequences by two bits. A total of 28 or 26 bits per subsegment is used for quantisation, which yields  $36 + 2 \cdot 28 + 2 \cdot 26 = 144$  bits/30 ms, i.e., a bitrate of 4.8 kbit/s.

The TBPE codec was subjected to rigorous bit-sensitivity analysis [11] and the bits were assigned in three sensitivity classes for embedded source-matched forward error correction to be detailed in Section 5. Having described the proposed 4.8 kbit/s TBPE speech codec we now focus our attention on the design of the 8.52 kbps videophone codec proposed.

### 3. Video Codec

The video codec's outline is depicted in Fig. 2. It was designed to achieve a time-invariant compression ratio associated with an encoded video rate of 8.52 kbps. The codec's operation is initialised in the intra-frame mode, but once it switched to the inter-frame mode, any further mode switches are optional and only required if a drastic scene change occurs.

In the intra-frame mode the encoder transmits the coarsely quantised block averages for the current frame, which provides a low-resolution initial frame required for the operation of the inter-frame codec at both the commencement and during later stages of communications in order to prevent encoder/decoder misalignment. For  $176 \times 144$  pixel CCITT standard Quarter Common Intermediate Format (QCIF) images we limited the number of video encoding bits per frame to 852. In order to transmit all block averages with a 4-bit resolution while not exceeding the 852 bits/frame rate the forced-update block size is fixed to  $11 \times 11$  pixels.

In the motion-compensation  $8 \times 8$  blocks are used. At the commencement of the encoding procedure the motion compensation (MC) scheme determines a motion vector (MV) for each of the  $8 \times 8$  blocks using full-search [26]. The MC search window is fixed to  $4 \times 4$  pels around the center of each block and hence a total of 4 bits are required for the encoding of 16 possible positions for each MV. Before the actual motion compensation takes place, the

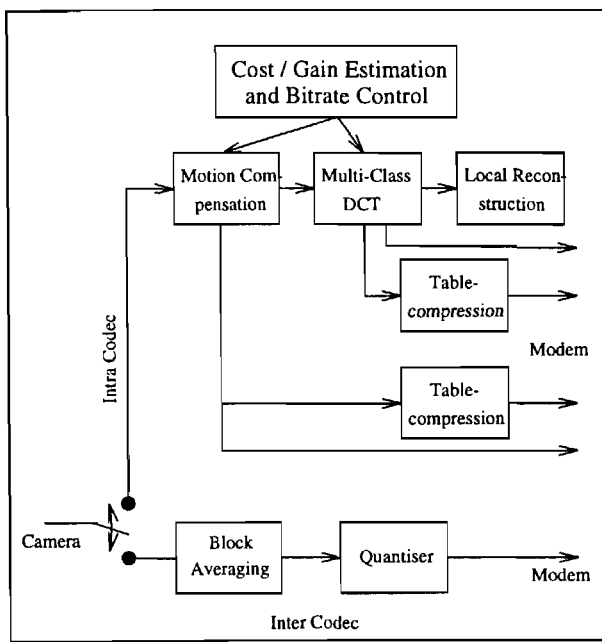


Fig. 2. The intra frame codec schematic.

codec tentatively determines the potential benefit of the compensation in terms of motion compensated error energy reduction. Then the codec selects those blocks as 'motion-active' whose gain exceeds a certain threshold. This method of classifying the blocks as motion-active and motion-passive results in an active/passive table, which consists of a one bit flag for each block, marking it as passive or active. In case of  $8 \times 8$  blocks and  $176 \times 144$  pel images this table consists of 396 entries which is compressed using elements of a two stage quad tree as follows.

First the whole table is grouped in  $2 \times 2$  blocks and a four bit symbol is allocated to those blocks which contain at least one active flag. These symbols are then run length encoded and transmitted to the decoder. This concept requires a second active table containing  $396 / 4 = 99$  flags in order to determine which of the two by two blocks contain active vectors. Three consecutive flags in this table are packetised to a symbol and then run length encoded. As a result, a typical 396-bit active/passive table containing 30 active flags can be compressed to less than 150 bits. Due to their low correlation the motion vectors themselves are not run length encoded. If at this stage of the encoding process the number of bits allocated to the compressed tables and active motion vectors exceeds half of the total number of available bits/frame, a number of blocks satisfying the motion-active criterion will be relegated to the motion-passive class. This process takes account of the subjective importance of various blocks and does not ignore motion-active blocks in the central eye and lip regions of the image, while relegating those, which are closer to the fringes of the image.

Pursuing a similar approach, gain control is also applied to the Discrete Cosine Transform (DCT) based compression [26]. Every block is DCT transformed and quantised. In order to take account of the non-stationary nature of the motion compensated error residual (MCER) and its time-variant frequency-domain distribution, four different sets of DCT quantisers were designed. The quantisation distortion associated with each quantiser is computed in order to

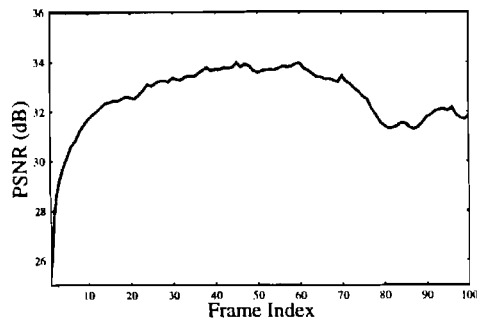


Fig. 3. PSNR versus frame index performance of the 8.52 kbps video codec for the 'Miss America' sequence.<sup>1</sup>

be able to choose the best one. Ten bits are allocated for each quantiser, each of which are trained Max-Lloyd quantisers catering for a specific frequency-domain energy distribution class. All DCT blocks whose coding gain exceeds a certain threshold are marked as DCT-active resulting in a similar active/passive table as for the motion vectors. For this second table we apply the same run length compression technique, as above. Again, if the number of bits required for the encoding of the DCT-active blocks exceeds half of the maximum allowable number, blocks around the fringes of the image are considered DCT-passive, rather than those in the central eye and lip sections. If, however, the active DCT coefficient and activity-table do not fill up the fixed-length transmission burst, the thresholds for active DCT blocks is lowered and all tables are recomputed.

The encoded parameters are transmitted to the decoder and also locally decoded in order to be used in future motion predictions. The video codec's Peak Signal to Noise Ratio (PSNR) versus frame index performance is shown in Fig. 3, where an average PSNR of about 33.3 dB was achieved for the MA sequence.

## 4. Modulation Issues

### 4.1. CHOICE OF MODULATION

In conventional mobile systems, such as the Pan-European GSM system [7] or the Digital European Cordless Telecommunications (DECT) scheme [6] constant envelope partial response Gaussian Minimum Shift Keying (GMSK) [6] is employed. Its main advantage is that it ignores any fading-induced amplitude fluctuation present in the received signal and hence facilitates the utilisation of power-efficient non-linear class-C amplification. In third generation personal communication systems however benign pico- and micro-cells will be employed, where low transmitted power and low signal dispersion are characteristic. Hence the employment of more bandwidth efficient multilevel modulation schemes becomes realistic. In fact the American and Japanese second generation digital systems have already opted for 2 bits/symbol modulation.

Multi-level modulation schemes have been considered in depth in reference [45] and in Chapters 17 and 18 we have shown that the bandwidth efficiency and minimum required signal-to-noise ratio (SNR) and Signal-to-Interference Ratio (SIR) of a modulation scheme in a given frequency re-use structure is dependent on the bit error ratio (BER) targeted. The required BER in turn is dependent on the robustness of the source codecs used. Furthermore,

<sup>1</sup> This video sequence can be viewed under <http://rice.ecs.soton.ac.uk>.

in indoors scenarios the partitioning walls and floors mitigate the co-channel interference and this facilitates the employment of 16-level Quadrature Amplitude Modulation (16QAM).

The 'maximum minimum distance' square-shaped QAM constellation [12] is optimum for transmissions over Additive White Gaussian (AWGN) channels, since it has the highest average distance amongst its constellation points, yielding the highest noise protection distances. Until quite recently QAM developments were focused at the benign AWGN telephone line and point-to-point radio applications [13], which led to the definition of the CCITT telephone circuit modem standards V.29-V.33 based on various QAM constellations ranging from uncoded 16-QAM to trellis coded (TC) 128-QAM. In recent years QAM research for hostile fading mobile channels has been motivated by the ever-increasing bandwidth efficiency demand for mobile telephony [14–25], although it requires power-inefficient class A or AB linear amplification [27–30]. However, the power consumption of the low-efficiency class-A amplifier [29, 30], is less critical than that of the digital speech, image and channel codecs. Out-of-band emissions due to class AB amplifier non-linearities generating adjacent channel interferences can be reduced by some 15–20 dB using the adaptive predistorter proposed by Stapleton et al. [32, 33].

When using the square-shaped 16-QAM constellation it is essential to be able to separate the information modulated on to the in-phase (I) and quadrature-phase (Q) carriers by the help of coherent demodulation, invoking the Transparent-tone-in-band (TTIB) principle invented by McGeehan and Bateman [35–37] or invoking Pilot Symbol Assisted Modulation (PSAM) [34]. Although these methods can eliminate the residual BER, they are significantly more complex to implement than their non-coherently detected differentially coded counterparts. Based on the above arguments and constrained by the high bandwidth efficiency requirement in this treatise we have opted for non-coherently detected 16-QAM.

#### 4.2. NON-COHERENT STAR 16-QAM

The pivotal point of differentially coded non-coherent QAM demodulation is that of finding a rotationally symmetric QAM constellation, where all constellation points are rotated by the same amount. Such a rotationally symmetric 'star-constellation' was proposed in [23]. A disadvantage of the proposed star 16-QAM (16-StQAM) constellation is its lower average energy. While square 16-QAM had an average phasor energy of  $10d^2$ , the 16-StQAM halves this value to  $5d^2$ , where  $2d$  is the phasor spacing of the I and Q components. This implies a 3 dB disadvantage over Gaussian channels, but via Rayleigh fading channels this SNR penalty becomes less.

Our differential encoder obeys the following rules. The first bit  $b_1$  of a four-bit symbol is differentially encoded onto the phasor magnitude, yielding a ring-swap for an input logical one and maintaining the current magnitude, i.e., ring for  $b_1 = 0$ . Bits  $(b_2, b_3, b_4)$  are then differentially Gray-coded onto the phasors of the particular ring pin-pointed by  $b_1$ . Accordingly,  $(b_2, b_3, b_4) = (0, 0, 0)$  implies no phase change,  $(0, 0, 1)$  a change of  $45^\circ$ ,  $(0, 1, 1)$  a change of  $90^\circ$ , etc.

The corresponding non-coherent differential 16-StQAM demodulation is equally straightforward, having decision boundaries at a concentric ring of radius 2 and at phase rotations of  $(22.5^\circ + n.45^\circ)$   $n = 0 \dots 7$ . Assuming received phasors of  $P_t$  and  $P_{t+1}$  at consecutive sampling instants of  $t$  and  $t + 1$ , respectively, bit  $b_1$  is inferred by evaluating the condition:

$$\left| \frac{P_{t+1}}{P_t} \right| \geq 2. \quad (6)$$

If this condition is met,  $b_1 = 1$  is assigned, otherwise  $b_1 = 0$  is demodulated. Bits ( $b_2, b_3, b_4$ ) are then recovered by computing the phase difference

$$\Delta\Theta = (\Theta_{t+1} - \Theta_t) \pmod{2\pi} \quad (7)$$

and comparing it against the decision boundaries  $(22.5^\circ + n \cdot 45^\circ)$   $n = 0 \dots 7$ . Having decided which rotation interval the received phase difference  $\Delta\Theta$  belongs to, Gray-decoding delivers the bits ( $b_2, b_3, b_4$ ).

From our previous discourse it is plausible that the less dramatic the fading envelope and phase trajectory fluctuation between adjacent signaling instants, the better this differential scheme works. This implies that lower vehicular speeds are preferred by this arrangement, if the signaling rate is fixed. Therefore the modem's performance improves for low pedestrian speeds, when compared to typical vehicular scenarios. Alternatively, for a fixed vehicular speed higher signaling rates are favourable, since the relative amplitude and phase changes introduced by the fading channel between adjacent information symbols are less drastic. For a full treatise on QAM the interested reader is referred to [45]. After this discussion on appropriate QAM techniques let us now consider ways of reducing the QAM BER using FEC techniques.

## 5. Channel Coding

Both convolutional and block codes have been successfully used to combat the bursty channel errors. Cox, Hagenauer et al. [38] proposed rate compatible punctured convolutional (RCPC) codes [39] to provide bit sensitivity matched FEC protection for a subband speech codec using a  $R=1/2$  rate mother code. In our proposed packet radio transceiver we opted for binary Bose–Chaudhuri–Hocquenghem (BCH) block codes, since we found that the subjective speech quality of BCH-coded speech was often preferable to RCPC coded speech due to longer unimpaired speech segments, even if the objective Segmental Signal to Noise Ratio (SEGSNR) and Bit Error Rate (BER) performances of the RCPC and block codes were similar [40]. This was because the speech quality is typically more strongly dependent on the frame error rate (FER) than on decoded BER. Furthermore, powerful block codes also have reliable error detection capability, which can be advantageously exploited in order to invoke speech post enhancement in case of error events [44].

A set of appropriate FEC codes is constituted by the BCH5=BCH(63,36,5), BCH2=BCH(63,51,2) and BCH1=BCH(63,57,1) codes, correcting 5, 2 and 1 bits per 63-bit frame, respectively. Accordingly, the most sensitive class 1 (C1) 36 speech bits are protected by the powerful BCH(63,36,5) code, while the less vulnerable 51 class 2 (C2) and 57 class 3 (C3) bits are encoded by the BCH(63,51,2) and BCH(63,57,1) codes, respectively. The total number of protected bits is 144. The packet header conveying control information is also BCH(63,36,5) coded, hence  $4 \cdot 63 = 252$  bits per 30 ms are transmitted. After adding six ramp-symbols in order to assist the transceiver in its attempt to mitigate spurious adjacent channel emissions the total bit rate becomes 8.6 kbit/s, yielding a signaling rate of 2.15 kBd.

The above FEC scheme has the advantage of curtailing error propagation across speech frame boundaries and over-bridging deep channel fades for typical urban vehicular speeds. For example, for a vehicular speed of 30 mph or 13.3 m/s the traveling distance is 39.9 cm/30 ms speech frame. For a propagation frequency of 1.9 GHz, as in the future PCS the wavelength is about 15 cm, and therefore interleaving over an interval of about 40 cm ensures adequate error



Table 2. Summary of PRMA/TDMA parameters

| PRMA parameter       |          |
|----------------------|----------|
| Channel rate         | 20 kBd   |
| Speech rate          | 2.15 kBd |
| Video rate           | 4.3 kBd  |
| Frame duration       | 30 ms    |
| Total no. of slots   | 9        |
| No. of PRMA slots    | 7        |
| No. of TDMA slots    | 2        |
| Slot duration        | 3.33 ms  |
| Header length        | 63 bits  |
| Maximum speech delay | 32 ms    |
| Speech perm. prob.   | 0.6      |

randomisation for the FEC scheme to work efficiently. However, for pedestrians PSs there is a danger of idling in deep fades, in which case a switch-diversity scheme is essential.

The 852 bits video frame is encoded using 12 BCH(127,71,9) code words, yielding a total of 1524 bits. A pair of such codewords form a video packet of 254 bits, which is expanded by four ramp symbols to deliver a 258-bit/30 ms video packet at a corresponding rate of 2.15 kBd. Six such video packets are needed to deliver the 1524-bit BCH-coded video frame, but during the 90 ms video frame repetition time there are only three 30 ms PRMA frames. This implies that two reserved time-slots per PRMA frame are required for video users. This is equivalent to a video signaling rate of  $2 \times 2.15 = 4.3$  kBd. Having resolved the choice of FEC codecs let us now consider how PRMA can be used to maximise the number of users supported.

## 6. Packet Reservation Multiple Access

PRMA is a relative of slotted ALOHA contrived for conveying speech signals on a flexible demand basis via time division multiple access (TDMA) systems. PRMA was documented in a series of excellent treatises by Goodman et al [41], while a PRMA-assisted adaptive differential pulse code modulation (ADPCM) transceiver was proposed in reference [42]. The voice activity detector (VAD) [42] queues the active speech spurts to contend for an up-link TDMA time-slot for transmission to the BS. Inactive users' TDMA time slots are offered by the BS to other users, who become active and are allowed to contend for the un-used time slots with a less than unity permission probability. This measure prevents previously colliding users from consistently keep colliding in their further attempts to attain a time-slot reservation. If several users contend for an available slot, neither of them will be granted it, while if only one user requires the time slot, he can reserve it for its future communications. When many users are contending for a reservation, the collision probability is increased and hence a speech packet might have to contend for a number of consecutive slots, until its maximum contention delay of typically 32 ms expires. In this case the speech packet must be dropped, but the packet dropping probability must be kept below 1%, a value inflicting minimal degradation in perceivable speech quality in contemporary speech codecs.

The transmitted Baud rate of our transceiver was fixed to 20 kBd, in order for the PRMA signal to fit in a 30 kHz channel slot, as in the IS-54 system, when using a modem excess

bandwidth of 50 % [45]. Hence our transceiver can accommodate TRUNC(20 kD/2.1 kD) = 9 time slots, where TRUNC represents truncation to the nearest integer. The slot duration was  $30 \text{ ms} / 9 \approx 3.33 \text{ ms}$  and one of the PRMA users was transmitting speech signals recorded during a telephone conversation, while all the other users generated negative exponentially distributed speech spurts and speech gaps with mean durations of 1 and 1.35 s. The PRMA parameters used are summarised in Table 2. Observe that two time slots are reserved for a videophone users and 7 slots are dedicated to PRMA for the speech users.

The performance of communications systems is often evaluated in terms of the teletraffic carried, while maintaining a set of communications quality measures. In conventional TDMA mobile systems the grade of service (GOS) degrades due to speech impairments caused by call blocking, hand-over failures and speech frame interference engendered by noise, as well as co- and adjacent-channel interference. In PRMA-assisted systems calls are not blocked due to the lack of an idle time-slot, but the packet dropping probability is increased gracefully. Hand-overs will be performed in the form of contention for an idle time slot provided by the specific BS offering the highest signal quality amongst the potential target BSs.

The specific physical up-link to the BS offering the best signal quality during decoding the packet header is not likely to substantially degrade during the life-time of an active speech spurt having a typical mean duration of 1s or some thirty consecutive 30 ms speech frames. If, however the link degrades before the next active spurt is due for transmission, the subsequent contention phase is likely to establish link with another BS. Hence this process will have a favourable effect on the channel's quality, effectively simulating a diversity system having independent fading channels and limiting the time spent by the MS in deep fades, thereby avoiding channels with high noise or interference.

This advantageous property can be exploited to train a self-adjusting adaptive system using the channel segregation scheme proposed for PRMA systems in reference [31]. Accordingly, each BS evaluates and ranks the quality of its idle physical channels constituted by the unused time slots on a frame-by-frame basis and identifies a certain number of slots,  $N$ , with the highest quality i.e. lowest noise and interference. These high-quality, low-interference channels are segregated for contention, while the lower quality idle slots contaminated by noise and interference are temporarily disabled. Hence upon a new access request the BS is likely to receive a signal having low interference, which maximizes the chances of successful packet decoding, unless a collision caused by a simultaneous MS attempt to attain a reservation has occurred. When a successful, uncontended reservation takes place, the BS promotes the highest quality disabled time slot to the set of  $N$  segregated channels, unless its quality is unacceptably low. It appears plausible that if  $N$  is high, the packet dropping probability becomes low, but the physical channels constituted by the time slots might become heavily interfered, while if  $N$  is low, we have a packet dropping-dominated scenario, which equally limits the GOS.

Clearly, the main cause of GOS degradation in PRMA systems is limited to speech packet corruption due to noise or interference and packet dropping [43]. They both result in different subjective speech or GOS degradation, which we will attempt to quantitatively compare in terms of the objective segmental signal to noise ratio (SEGSNR) degradation. Quantifying these GOS degradations in relative terms in contrast to each other will allow us to appropriately split the acceptable overall degradation between packet dropping and packet corruption. With the system elements described we now focus our attention on the amalgamated PCS transceiver proposed.

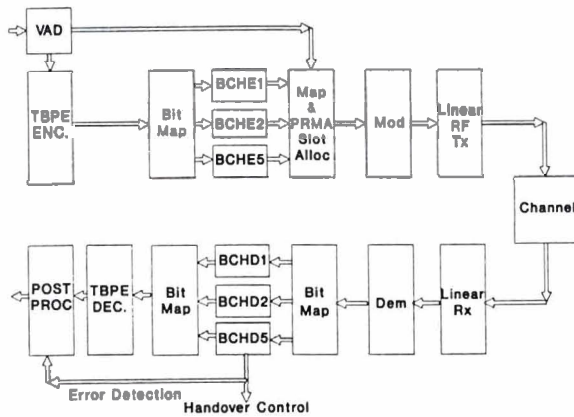


Fig. 4. Transceiver schematic.

## 7. The Proposed PCS Transceiver

The block diagram of the proposed PCS transceiver is shown in Fig. 4. The TBPE encoder outputs a 4.8 kbit/s bit stream, which is mapped in three bit sensitivity classes, C1, C2 and C3. The bits belonging to these classes are FEC encoded along with the packet header conveying control information by the BCHE1 = BCH(63,36,5), BCHE2 = BCH(63,51,2) and BCHE1 = BCH(63,57,1) encoders, respectively. The 258 bits/30ms=8.6 kbit/s FEC-coded speech packets are then transmitted at 2.15 kBd using differentially encoded 16-QAM. Only active speech spurts are queued by the VAD for transmission. If there is no other PS attempting to acquire a slot reservation, the PS is allocated this particular time slot for its future communication. In case of collision further contention is enabled with a less than unity permission probability, until either a slot is reserved or the speech packet's life-span expires. The microcellular PSs receive the slot-status near-instantaneously, implying negligible propagation delays. If this cannot be ensured, adaptive time frame alignment must be used, as proposed for the Pan-European GSM system [7].

The 852 bits per frame video encoded stream is BCH(127,71,9) coded to 1524 bits/frame and transmitted at 4.3 kBd, which is equivalent to the signaling rate of two speech users. Clearly, for video telephony two time-slots are required. The video transceiver obeys the structure of Fig. 4, simply the TBPE speech encoder must be replaced by the DCT video codec and no bit mapping is invoked, since a single-class BCH(127,71,9) codec is used.

The receiver seen in Fig. 4 carries out the inverse functions of the transmitter. The error detection capability of the strongest BCH(63,36,5) decoder is exploited to initiate handovers and to invoke speech post-processing [44], if the FEC decoder happens to be overloaded due to interference or collision. The system elements of Fig. 4 were simulated and the main transceiver parameters are summarised in Table 3, while the system performance will be characterised in the following section.

## 8. Results and Discussion

The system performance was evaluated for the worst-case narrowband Rayleigh-fading channel characterised by a propagation frequency of 1.9 GHz, vehicular speed of 30 mph and signaling rate of 20 kBd. The packet dropping probability versus number of users curve of the proposed system is portrayed in Fig. 5. Observe that about 10-11 users can be supported

Table 3. Speech transceiver parameters

| 1                     | 2                         | 3                         | 4                           | 5                           | 6                        | 7                         | 8                       |
|-----------------------|---------------------------|---------------------------|-----------------------------|-----------------------------|--------------------------|---------------------------|-------------------------|
| Speech Rate<br>(kbps) | PRMA Source<br>Rate (kBd) | TDMA User<br>Bandw. (kHz) | No of TDMA<br>Users/Carrier | No of PRMA<br>Users/Carrier | No of PRMA<br>Users/slot | PRMA User<br>Bandw. (kHz) | Min SNR<br>and SIR (dB) |
| 4.8                   | 2.15                      | 4.3                       | 7                           | 10                          | 1.43                     | 3                         | 24                      |

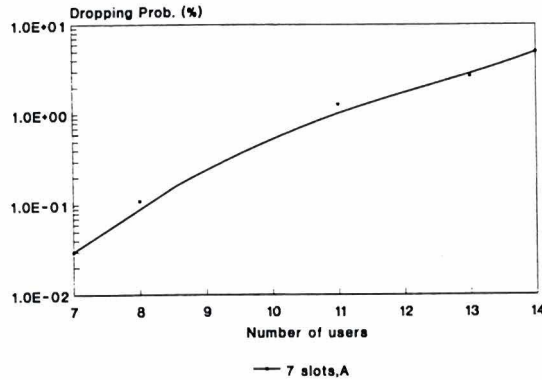


Fig. 5. Packet dropping versus number of speech users performance.

by our 7-slot PRMA scheme with  $P_{\text{drop}} < 1\%$ , a value inflicting almost negligible speech degradation.

The performance of our 16-QAM transceiver with and without second-order selection diversity and BCH1, BCH2 and BCH5 FEC coding is depicted in Figs. 6 and 7. When no diversity is invoked, at a channel SNR of about 28 dB the most important C1 bits protected by the BCH5 codec have a BER of about 0.1%, while the less sensitive C2 and C3 bits attain BERs of about 0.5 and 1%, respectively. These values are sufficiently low for nearly un-impaired speech transmission. When second-order diversity is used, these target BERs are achieved around 24 dB channel SNR. Although the lowest integrity C3 BCH1 codec does not provide a reduced BER for the least significant bits, since it is often overloaded, it ensures periods of un-impaired transmission for the most robust bits, which has a favourable subjective effect on the perceived speech quality. The 16-QAM modem is sensitive to co-channel interference, requiring SIR values similar to the minimum channel SNR necessitated [45]. Therefore it is advantageous to use the channel segregation algorithm proposed in references [31, 43], unless the system is used in an in-doors environment, where the partitioning walls and floors naturally contribute towards the interference attenuation.

The overall objective SEGSNR degradation (SEGSNR-DEG) versus channel SNR performance of our diversity-assisted PCS transceiver is displayed in Fig. 8 parameterized with the number of PRMA users supported. While for 7–10 users no speech degradation can be observed, if the channel SNR is in excess of about 24 dB, for 12 users the SEGSNR-DEG due to PRMA packet dropping becomes noticeable, although not subjectively objectionable. In case of 14 users, however, there is a consistent SEGSNR-DEG of about 1 dB due to the 4–5% packet dropping probability seen in Fig. 5. Without diversity about 5 dB higher channel

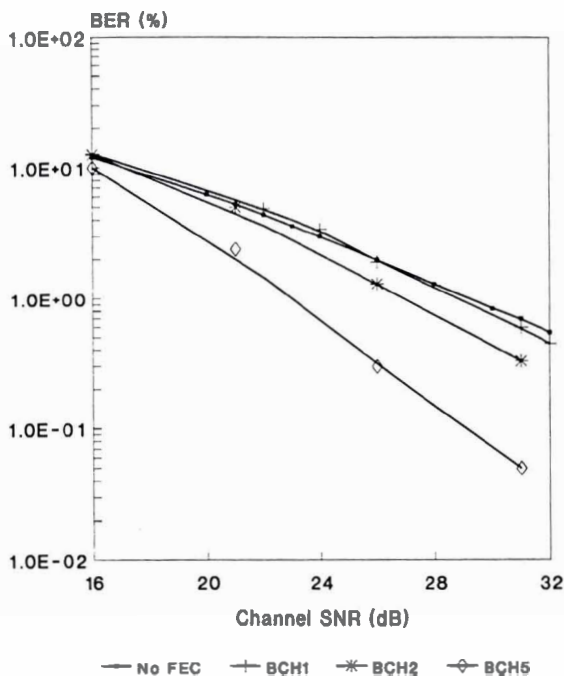


Fig. 6. 16-QAM BER versus channel SNR performance at 30 mph without diversity using no FEC and BCH1, BCH2 and BCH5 coding.

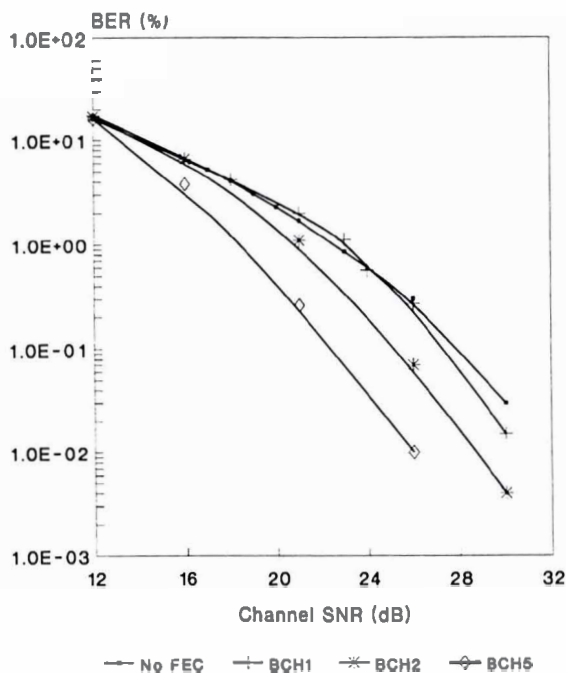


Fig. 7. 16-QAM BER versus channel SNR performance at 30 mph with diversity using no FEC and BCH1, BCH2 and BCH5 coding.

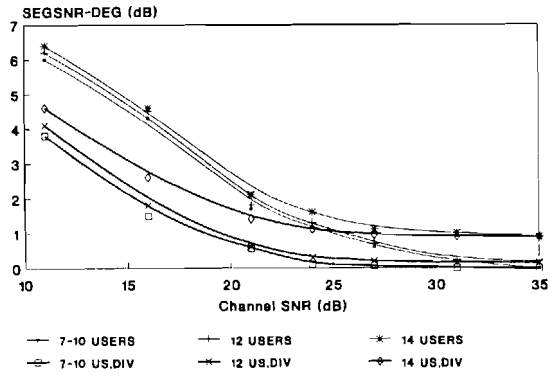


Fig. 8. SEGSNR-DEG versus channel SNR performance of the proposed 16-StQAM transceiver at 30 mph with and without diversity parameterized with the number of PRMA users supported.

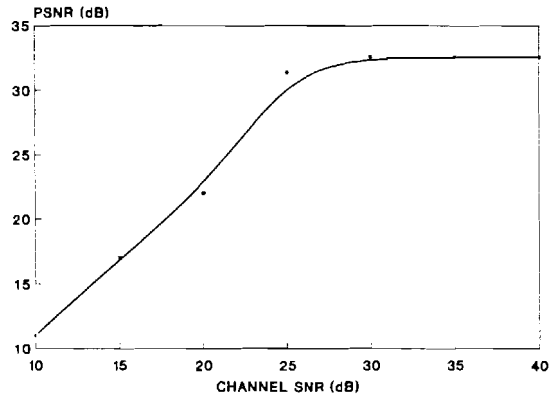


Fig. 9. PSNR versus channel SNR performance of the proposed diversity-assisted videophone scheme.

SNR is necessitated in order to achieve a similar performance to that of the diversity-assisted scheme.

The PSNR versus channel SNR performance of the diversity-assisted video transceiver is portrayed in Fig. 9, where in harmony with the voice transceiver a channel SNR of about 22–25 dB is required for near-unimpaired video quality. Without diversity the video scheme lacks robustness, since the corrupted run-length coded activity tables affect the whole of each video frame.

### 9. Summary and Conclusions

The potential of a bandwidth-efficient 2.15 kBd PRMA-assisted TBPE/BCH/16-QAM scheme has been investigated for deployment in the future PCS under the assumption of benign channel conditions. Within the 30 kHz IS-54 bandwidth about 10 voice users plus a video telephone user can be supported, if channel SNR and SIR values in excess of about 24 dB can be maintained. The main transceiver features are summarised in Table 3. The system performance can be further improved at the cost of higher implementational complexity, when using a more sophisticated pilot symbol assisted, block-coded coherent square 16-QAM modem. Future work will be targeted at improving the speech quality, implementational complexity, bit rate, bandwidth occupancy and error resilience trade-off achieved invoking

the adaptive transceiver re-configuration algorithms we used in reference [46] and Chapters 13, 17 and 18 of reference [45].

## Acknowledgement

The authors are grateful to the Education and Physical Sciences Research Council (EPSRC), Swindon, UK for their financial support of the research project GR/J46845.

## References

1. IEEE Communication Magazine, Dec. 1992, Vol. 30 No. 12, Special Issue, PCS: The Second Generation, pp. 30–133, Guest Editor: R. Steele.
2. IEEE Communications Magazine, Special Issue on Coded Modulation, Vol. 29, 1991.
3. E. Biglieri and M. Luise, Coded Modulation and Bandwidth-Efficient Transmission, Proc. of the Fifth Tirrenia Intern. Workshop, 8–12 Sept. 1991, Elsevier, Netherlands, 1992.
4. R. Steele, Speech codecs for personal communications, IEEE Communication Magazine, Nov. 1993 Vol. 31 No. 11, Feature Topic: Speech Processing, Guest Editors: N. Jayant and R.V. Cox.
5. M.R. Schroeder and B.S. Atal, "Code-Excited Linear Prediction (CELP): High Quality Speech at Very Low Bitrates", Proc. of ICASSP'85, pp. 937–941, 1985.
6. R. Steele (Ed.), "Mobile Radio Communication", IEEE Press-Pentech Publishers, 1992.
7. L. Hanzo and J. Stefanov, The Pan-European digital cellular mobile radio system – known as GSM, Chapter 8, pp. 677–768, in R. Steele (Ed.), "Mobile Radio Communication", IEEE Press-Pentech Publishers, 1992.
8. R.A. Salami, Binary pulse excitation: a novel approach to low complexity CELP coding, in "Advances in speech coding", Kluwer Academic Publishers 1991.
9. I.A. Gerson and M.A. Jasiuk, "Vector sum excitation linear prediction (VSELP) speech coding at 8 kbps". Proc. ICASSP'90, Albuquerque, New Mexico, USA, pp. 461–464, 1990.
10. J.-P. Adoul et al., "Fast CELP coding based on algebraic codes". Proc. ICASSP'87, pp. 1957–1960, 1987.
11. L. Hanzo, R. Salami, R. Steele and P.M. Fortune, "Transmission of Digitally Encoded Speech at 1.2 Kbd for PCN", IEE Proc.-I, Vol. 139, pp. 437–447, 1992.
12. C.N. Campopiano and B.G. Glazer, "A Coherent Digital Amplitude and Phase Modulation Scheme", IRE Trans. Commun. Systems, Vol. CS-10, pp. 90–95, 1962.
13. G.D. Forney et al., "Efficient Modulation for Band-Limited Channels", IEEE Journal on Selected Areas in Communications, Vol. SAC-2, pp. 632–647, 1984.
14. K. Feher, "Modems for Emerging Digital Cellular Mobile Systems", IEEE Tr. on VT, Vol. 40, pp. 355–365, 1991.
15. M. Iida and K. Sakniwa, "Frequency Selective Compensation Technology of Digital 16-QAM for Microcellular Mobile Radio Communication Systems", Proc. of VTC'92, Denver, Colorado, pp. 662–665, 1992.
16. R.J. Castle and J.P. McGeehan, "A Multilevel Differential Modem for Narrowband Fading Channels", Proc. of VTC'92, Denver, Colorado, pp. 104–109, 1992.
17. D.J. Purle, A.R. Nix, M.A. Beach and J.P. McGeehan, "A Preliminary Performance Evaluation of a Linear Frequency Hopped Modem", Proc. of VTC'92, Denver, Colorado, pp. 120–124, 1992.
18. Y. Kamio and S. Sampei, "Performance of Reduced Complexity DFE Using Bidirectional Equalizing in Land Mobile Communications", Proc. of VTC'92, Denver, Colorado, pp. 372–376, 1992.
19. T. Nagayasu, S. Sampei and Y. Kamio, "Performance of 16-QAM with Decision Feedback Equalizer Using Interpolation for Land Mobile Communications", Proc. of VTC'92, Denver, Colorado, pp. 384–387, 1992.
20. E. Malkamaki, "Binary and Multilevel Offset QAM, Spectrum Efficient Modulation Schemes for Personal Communications", Proc. of VTC'92, Denver, Colorado, pp. 325–378, 1992.
21. Z. Wan and K. Feher, "Improved Efficiency CDMA by Constant Envelope SQAM", Proc. of VTC'92, Denver, Colorado, pp. 51–55, 1992.
22. H. Sasaoka, "Block Coded 16-QAM/TDMA Cellular Radio System Using Cyclical Slow Frequency Hopping", Proc. of VTC'92, Denver, Colorado, pp. 405–408, 1992.
23. W.T. Webb, L. Hanzo and R. Steele, "Bandwidth-Efficient QAM Schemes for Rayleigh-Fading Channels", IEE Proceedings, Vol. 138, pp. 169–175, 1991.
24. L. Hanzo, R. Steele and P.M. Fortune, "A Subband Coding, BCH Coding and 16-QAM System for Mobile Radio Speech Communications", IEEE Vehicular Technology, Vol. 39, pp. 327–339, 1990.
25. P.M. Fortune, L. Hanzo and R. Steele, On the Computation of 16-QAM and 64-QAM Performance in Rayleigh-Fading Channels, IEICE Tr. on Comms., Vol. E75-B, pp. 466–475, 1992.
26. A.K. Jain, "Fundamentals of Digital Image Processing", Prentice-Hall, 1989.



27. A.S. Wright and W.G. Durtler, "Experimental Performance of an Adaptive Digital Linearized Power Amplifier", *IEEE Tr. on VT*, Vol. 41, pp. 395–400, 1992.
28. M. Faulkner and T. Mattson, "Spectral Sensitivity of Power Amplifiers to Quadrature Modulator Misalignment", *IEEE Tr. on VT*, Vol. 41, pp. 516–525, 1992.
29. P.B. Kennington et al., "Broadband Linear Amplifier Design for a PCN Base-Station", *Proc. of 41st IEEE VTC*, pp. 155–160, 1991.
30. R.J. Wilkinson et al., "Linear Transmitter Design for MSAT Terminals", 2nd Int. Mobile Satellite Conference, June 1990.
31. M. Frullone, G. Riva, P. Grazioso and C. Carciofy, "Investigation on Dynamic Channel Allocation Strategies Suitable for PRMA Schemes", 1993 IEEE Int. Symp. on Circuits and Systems, Chicago, pp. 2216–2219, 1993.
32. S.P. Stapleton and F.C. Costescu, "An Adaptive Predistorter for a Power Amplifier Based on Adjacent Channel Emissions", *IEEE Tr. on VT*, Vol. 41, pp. 49–57, 1992.
33. S.P. Stapleton, G.S. Kandola and J.K. Cavers, "Simulation and Analysis of an Adaptive Predistorter Utilizing a Complex Spectral Convolution", *IEEE Tr. on VT*, Vol. 41, pp. 387–394, 1992.
34. J.K. Cavers, "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels", *IEEE Tr. on VT*, Vol. 40, pp. 686–693, 1991.
35. A. Bateman and J.P. McGeehan, "Feedforward transparent tone in band for rapid fading protection in multipath fading", *IEE Int. Conf. Comms.* Vol. 68, pp. 9–13, 1986.
36. A. Bateman, "Feedforward transparent tone in band, Its implementation and applications", *IEEE Trans. Veh. Tech.* Vol. 39, pp. 235–243, 1990.
37. J.K. Cavers, "The performance of phase locked transparent tone in band with symmetric phase detection", *IEEE Trans. on Comms.*, Vol. 39, pp. 1389–1399, 1991.  
pp. 191–195, 1977.
38. R.V. Cox, J. Hagenauer, N. Seshadri and C.E.W. Sundberg, "Subband Speech Coding and Matched Convolutional Channel Coding for Mobile Radio Channels", *IEEE Tr. on Sig. Proc.*, Vol. 39, pp. 1717–1731, 1991.
39. J. Hagenauer, "Rate-Compatible Punctured Convolutional Codes (RCPC) and Their Applications", *IEEE Tr. on Comms.*, Vol. 36, pp. 389–400, 1988.
40. R.A. Salami, K.H.H. Wong, R. Steele and D. Appleby, "Performance of Error-Protected Binary Pulse Excitation Coders at 11.4 Kb/s over Mobile Radio Channels", *Proc. of ICASSP'90*, pp. 473–476, 1990.
41. S. Nanda, D.J. Goodman and U. Timor, "Performance of PRMA: A Packet Voice Protocol for Cellular Systems", *IEEE Tr. on VT*, Vol. 40, pp. 584–598, 1991.
42. L. Hanzo, J.C.S. Cheung, W. Webb and R. Steele, A packet reservation multiple access assisted cordless telecommunication scheme, *IEEE Tr. on VT.*, Vol. 43, pp. 234–244, 1994.
43. M. Frullone, G. Falciasacca, P. Grazioso, G. Riva and A. M. Serra, "On the performance of packet reservation multiple access with fixed and dynamic channel allocation", *IEEE Tr. on Veh. Techn.*, Vol. 42, pp. 78–86, 1993.
44. D.J. Goodman, G.B. Lockhart, O.J. Wasem, W.C. Wong, "Waveform substitution techniques for recovering missing speech segments in packet voice communications", *IEEE Tr. on ASSP*, Vol. 34, pp. 1440–1448, 1986.
45. W.T. Webb and L. Hanzo, "Modern Quadrature Amplitude Modulation, Principles and Applications for Fixed and Wireless Channels", Research Monograph, IEEE Press-Pentech Press, 1994.
46. R. Steele and W.T. Webb, "Variable rate QAM for data transmissions over mobile radio channels", Keynote paper, *Wireless 91*, Calgary Alberta, July 1991.



**Lajos Hanzo** was awarded a Ph.D. in telecommunications and has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the



Department of Electronics and Computer Science, University of Southampton, UK and has been a consultant to Multiple Access Communications Ltd., UK. He co-authored two books on mobile radio communications, published about eighty research papers and was awarded a number of distinctions. His main research interests are in the field of mobile multimedia communications, including speech, audio and video source coding, channel coding, modulation, transmission as well as the joint optimisation of these links. He is Senior member of the IEEE.



Jurgen Streit was born in Cologne, Germany, on March 30, 1968. He received his Dipl.-Ing. Degree from the University of Aachen, Germany in 1993. Since 1992 he has been with the Department of Electronics and Computer Science, University of Southampton, UK, where he completed his M.Sc. project and currently works as research assistant towards the degree of Ph.D. His research interest are low bitrate video coding in mobile communications.

Redwan Salami was born in Tyre, Lebanon, on April 15, 1961. He received the B.Sc. degree in Electrical Engineering from Al-Fateh University, Tripoli, Libya and the M.Sc. and Ph.D. degrees in Electronics from the University of Southampton, UK in 1987 and 1990, respectively. In 1990 he joined the Department of Electrical Engineering, University of Sherbrooke, Quebec, Canada, where he is currently a senior researcher. His research interests include speech coding, digital communications and digital mobile radio systems.



**William Webb** has gained a Ph.D. in the field of QAM techniques from the University of Southampton, UK, while working for Multiple Access Communications Ltd., UK, where he was promoted to the position of technical director. He has published over twenty papers on QAM for which he has been awarded the IERE Premium, and holds three patents on QAM. He currently works as a mobile radio consultant for Smith Systems Engineering Ltd., UK.