Adaptive Low-Rate Wireless Videophone Schemes

Lajos Hanzo, *Senior Member, IEEE*, and Jürgen Streit

Abstract—**The design and performance of a range of wireless videophone transceivers are presented. Highly bandwidth efficient, fixed but arbitrarily programmable rate, perceptually weighted Discrete Cosine Transform (DCT) based video codecs are proposed for quarter common intermediate format (QCIF) videophone sequences. Perceptually weighted cost/gain controlled motion compensation and quad-class DCT-based compression is applied with and without run-length coding. Specifically, we propose video codecs having transmission rates in the range of 5–11.36 kbps and preselected the 11.36 kbps Codec 1, the 8.52 kbps Codec 2 and the 8 kbps Codec 2a, for which we designed the intelligent reconfigurable Systems 1–5. After sensitivity-matched binary Bose–Chaudhuri–Hocquenghem (BCH) forward error correction (FEC) coding the data rate associated with Codec 1 and Codec 2a became 20.32 kbps, while that of Codec 2 was 15.24 kbps. Throughout these systems a partial forced update (PFU) technique was invoked in order to keep transmitter and receiver aligned amongst hostile channel conditions. When using Codec 1 in System 1 and coherent pilot symbol assisted 16-level quadrature amplitude modulation (16-PSAQAM), an overall signalling rate of 9 kBd was yielded. Over lower quality channels the 4QAM mode of operation had to be invoked, which required twice as many time slots to accommodate the resulting 18 kBd stream. The system's robustness was increased using Automatic Repeat Requests (ARQ), inevitably reducing the number of users supported, which was between 6 and 19 for the various systems. In a bandwidth of 200 kHz, similarly to the Pan-European GSM mobile radio system's speech channel, using System 1 for example 16 and 8 videophone users can be supported in the 16QAM and 4QAM modes, respectively. All system features are summarized in Table III.**

I. Introduction and Motivation

In RECENT YEARS, there has been an increased research activity in the field of videophony [33], [2]–[4], [6]–[9], [27], in particular for mobile channels. The motion pictures expert group [10], [11] (MPEG) and the p $\times$ 64 kbps CCITT H261 video codecs have been contrived for high-rate, low bit error rate fixed channels. Although the new MPEG 4 working group's activities target mobile videophony [12], to date there are no appropriate video standards for mobile videophony over existing standard radio systems.

Some authors have investigated the deployment of the H261 codec over mobile channels [5], [7] while others [2], [6], [33] have contrived a range of different run-length coded variable rate arrangements, such as subband coded schemes [9], [14]–[16], fractal-based [8] or discrete cosine transformed [27] systems. A common feature of most previously proposed systems is that they typically generate a variable output rate.

Manuscript received January 24, 1995; revised June 5, 1995. This work was supported by EPSRC, U.K., Contract GR/J46845. This paper was recommended by Associate Editor L. Wu.

The authors are with the Department of Electronics and Computer Science, University of Southampton, SO17 1BJ, U.K.

IEEE Log Number 9413981.

In contrast to the above variable rate schemes, in this treatise we set out to contrive a range of programmable constant-rate videophone codecs, which can adjust their coding rate in order to accommodate their stream in a conventional speech channel, such as for example that of the Pan-European GSM system [59], the Japanese PDC [19], and the American IS-54 [20] as well as IS-95 systems [21]. The corresponding speech rates are 13, 6.7, 8, and 9.6 kbps and the proposed video codecs are capable of operating at a scanning rate of 10 frames/s, while maintaining such low bit rates. Their further advantage is that they are capable of programmable multirate operation in the 3rd generation adaptive multimedia, multimode terminal of the near future, which is currently under intensive study within the European Community's Fourth Framework Programme in the field of Advanced Communications Technologies and Services (ACTS) [13].

In this short treatise we cannot consider the performance of the proposed video codecs in all of the above 2nd and 3rd generation mobile radio systems. Our main goal is to describe the design philosophy of our prototype video codecs and document their performance using three different characteristic fixed bit rates within the above mentioned typical speech coding rate range. Furthermore, we will devote special attention to transmission robustness issues and devise and evalute a range of error control measures. In particular, a variety of unequal protection coded modulation [26], [25] and Automatic Repeat Request (ARQ) techniques are invoked and assessed in robustness, complexity, and bandwidth efficiency terms and the benefits of a particular multimode scenario are analyzed. The aspects of devising reliable transceiver reconfiguration protocols and analysing the associated switching transients as well as signalling issues are left for future studies.

Specifically, the video source rate was fixed to 8, 8.52, or 11.36 kbps, and this stream was then transmitted using an intelligent transceiver, which can configure itself as a robust but less bandwidth efficient scheme or can double its bandwidth efficiency at the cost of requiring better channel conditions. Note that although the 11.36 kbps video rate differs from the 13 kbps GSM speech rate, some of the systems proposed accommodate exactly 16 or 8 videophone users, as the full- and half-rate GSM speech system.

System Concept: According to Shannon's pioneering work [22], which was further exposed in [23] and [24], in case of lossless coding the lowest achievable source coded rate is given by the source entropy [28]. Such an ideal source encoder produces a completely uncorrelated sequence, where all symbols are mutually independent and have the same significance or error sensitivity. Any further rate reduction implies that some distortion is inflicted. Since our source codecs operate well below the source entropy, the design philosophy hinges around the principle as to how best the total distortion is distributed over the source message in the time- or frequency-domain in order to minimise its subjective effects.

When using Shannon's ideal source codecs and channel codecs over memoryless AWGN channels, where bit errors occur randomly, there is no advantage in treating source and channel coding jointly. Our nonideal source codecs however produce sequences, which still retain correlation and unequal error sensitivity. Over fading mobile channels this problem is aggravated by the bursty error statistics, which can only be randomized using infinite memory channel interleavers inflicting infinite delays. In this situation source-matched channel coding [9], [18], [23], which takes account of the source significance information [23] (SSI) brings substantial advantages in terms of reducing the required minimum channel SNR.

Joint coding and modulation in the form of trellis coded modulation (TCM) or block coded modulation (BCM) was also proposed in the literature in order to reduce the required channel SNR [51], [52], while in [18] and [25] source-matched joint source/channel coding and modulation was introduced. In this treatise we will follow a similar design philosophy in order to achieve best videophone performance over fading channels.

The schematic of the proposed transceiver is portrayed in Fig. 7 and this treatise follows the same structure. Speech source coding issues are not considered here, the reader is referred to [34] and [33] for the choice of the appropriate speech codec. Channel coding issues are addressed in [50], while a detailed discussion of modulation is given in [26]. Section II outlines the design of a variety of programmable, but fixed-rate video source codecs and analyzes their bit sensitivity. Section III details modulation and transmission aspects, which is followed by the description of the source-matched transceiver in Section IV. The system's performance is characterized in Section V, before offering some conclusions in Section VI.

II. Videophone Codecs

2.1. Codec 1

Let us initially focus our attention on the proposed discrete cosine transform [28] (DCT) based video codec depicted in Fig. 1, which was designed for hostile mobile channels. The codec uses 176 × 144 pixels Quarter Common Intermediate Format (QCIF) images scanned at 10 frames/s. For the sake of communications convenience and simple networking our aim was to develop a fixed-rate codec which is able to dispense with an adaptive feed-back-driven bit-rate control buffer. Therefore a constant bit-rate source codec was required, which in Codec 1 forced us to avoid using efficient variable-rate compaction algorithms, such as Huffman coding. This was achieved by fixing both the number of 8 × 8 blocks to be motion-compensated and those to be subjected to DCT to 30 out of 22 × 18 = 396. The selection of these blocks is based on a gain-controlled approach, which will be highlighted next.

In order to curtail error propagation across image frames the codec was designed to switch between intraframe and interframe modes of operation. In the **intraframe mode** the encoder transmits the coarsely quantized block averages for the current frame, which provides a low-resolution initial frame required for the operation of the interframe codec at both the commencement and during later stages of communications in order to prevent encoder/decoder misalignment. The **interframe mode** of operation is based on a combination of gain-controlled motion compensation and gain-controlled DCT coding as seen in Fig. 1.

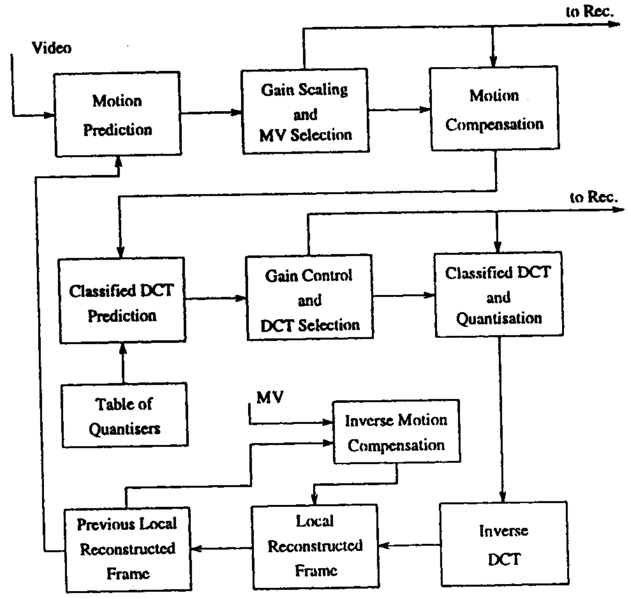


Fig. 1. Video encoder schematic.

Gain Controlled Motion Detection: At the commencement of the encoding procedure the motion compensation (MC) scheme determines a motion vector (MV) for each of the 8 × 8 blocks. The MC search window is fixed to 4 × 4 pels around the center of each block. Before the actual motion compensation takes place the codec tentatively determines the potential benefit of the compensation in terms of motion compensated error energy reduction. In order to emphasize the subjectively, more important eye and mouth region of the videophone images the potential gains for each motion compensated block are augmented by a factor of two in the center of the screen. Then the codec selects the thirty blocks resulting in the highest scaled gain, and motion compensation is applied only to these blocks, whereas for all other so-called passive blocks the codec applies simple frame differencing.

Gain Controlled Quadruple-Class DCT: Pursuing a similar approach, gain control is also applied to the DCT-based compression. Every block is DCT transformed and quantized. Because of the nonstationary nature of the motion compensated error residual (MCER) the energy distribution characteristics of the DCT coefficients vary. Therefore four different sets of DCT quantizers are available, as examplified in Fig. 2. All four bit allocation schemes are tentatively invoked in order to select the best set of quantizers resulting in the highest energy compaction gain. Ten bits are allocated for each quantizer, each of which are trained Max-Lloyd quantizers catering for a specific frequency-domain energy

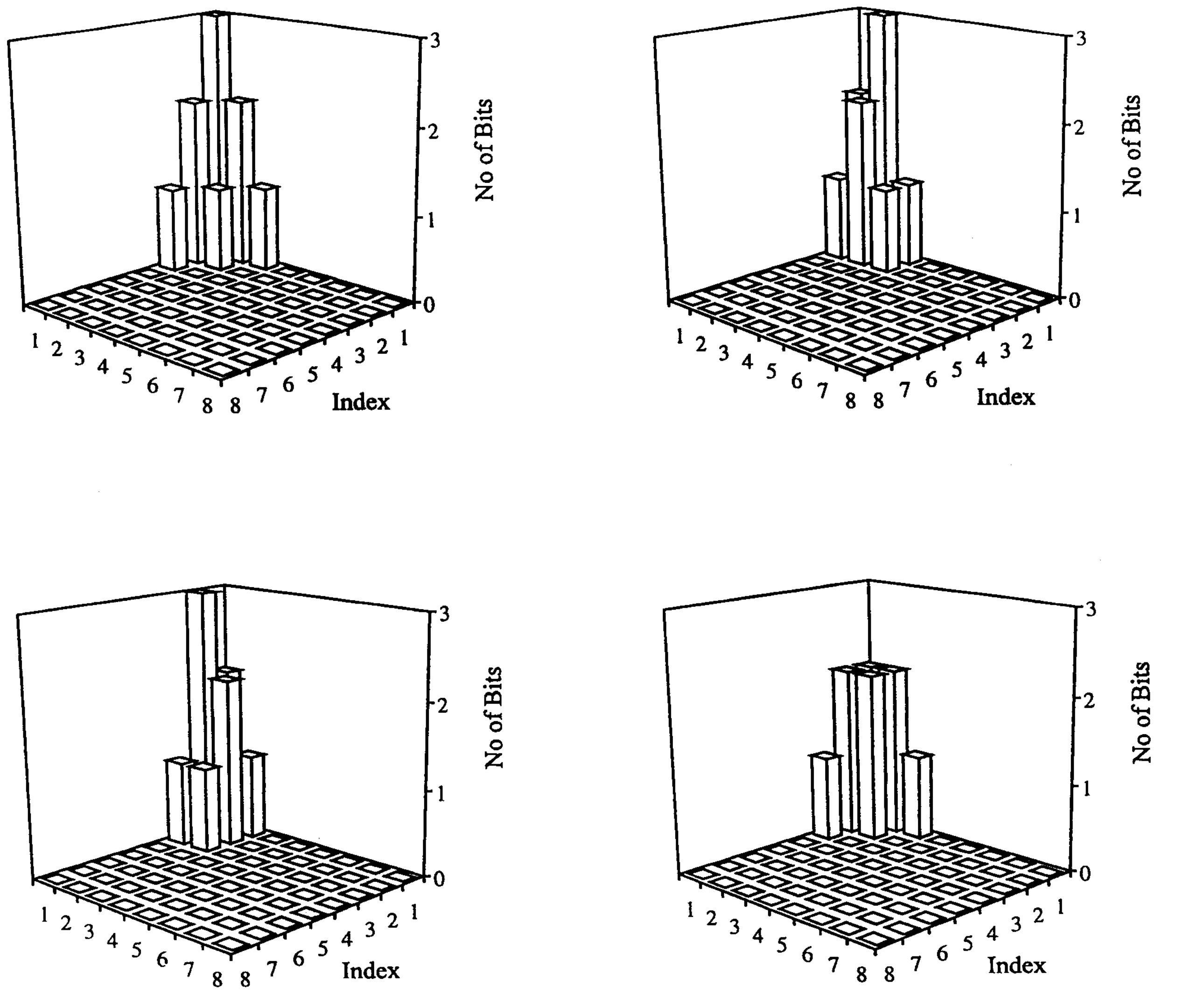


Fig. 2. Quad-class DCT quantization schemes.

distribution class. Again, the energy compaction gain values are scaled to emphasise the eye and mouth region of the image and the DCT coefficients of the thirty highest-compression blocks are transmitted to the decoder.

Partial Forced Update: The disadvantage of interframe codecs is their vulnerability to channel errors. Every channel error results in a misalignment between the reconstructed frame buffer of the encoder and decoder. The errors accumulate and do not decay, unless a leakage-factor or a partial forced update (PFU) technique is employed. In our proposed codec in every frame 22 out of the 396 blocks, scattered over the entire frame, are periodically updated using the 4-b quantized block means, which are partially overlayed on to the contents of the reconstructed frame buffer. The overlaying is performed such that the block's contents in the local buffer is weighted by 0.7 and superimposed on to the received block average, which is scaled by 0.3. The bit-rate contribution of this PFU process is a moderate $22 \times 4 = 88$ bits per QCIF frame and it refreshes about 5.6% of each frame.

Bit Allocation Strategy: The bit allocation scheme was designed to deliver 1136 b per frame, which begins with a 22-b frame alignment word (FAW). This is necessary to assist the video decoder's operation in order resume synchronous operation after loss of frame synchronization over hostile fading channels. The partial intraframe update refreshes only 22 out of 396 blocks every frame. Therefore every 18 frames or 1.8 s the update refreshes the same blocks. This periodicity is signalled to the decoder by transmitting the inverted FAW. A MV is stored using 13 b, where 9 b are required to identify one of the 396 the block indexes using the enumerative method and 4 b for encoding the 16 possible combinations of the X and Y displacements. The 8×8 DCT-compressed blocks use a total of 21 b, again 9 for the block index, 10 for the DCT coefficient quantizers, and 2 b to indicate which of the four quantizer has been applied. The total number of bits becomes $30 \cdot (13 + 21) + 22 \cdot 4 + 22 + 6 = 1136$, where six dummy bits were added in order to obtain a total of 1136 b suitable in terms of bit packing requirements for the specific forward error correction block codec used. The video codec's peak signal-to-noise ratio (PSNR) performance is portrayed in Fig. 3 for the well-known 'Miss America' sequence and for

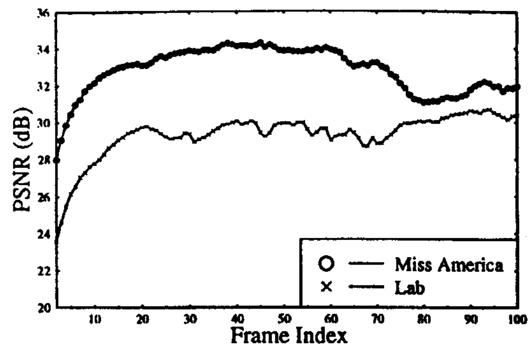


Fig. 3. PSNR performance of the 11.36 kbps Codec 1.

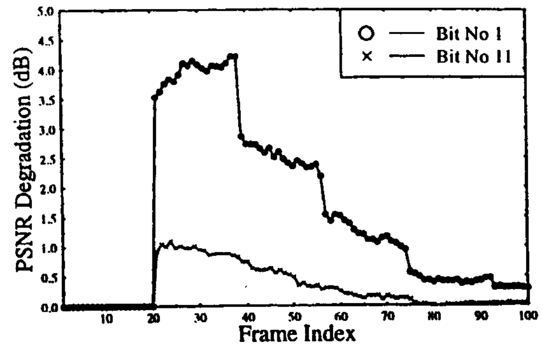


Fig. 4. PSNR degradation profile for Bits 2 and 11 of the MV in Codec 1.

TABLE I
BIT ALLOCATION TABLE

FAW	PFU	MV Index	MV	DCT Index	DCT	Padding	Total
22	22×4	30×9	30×4	30×9	30×12	6	1136

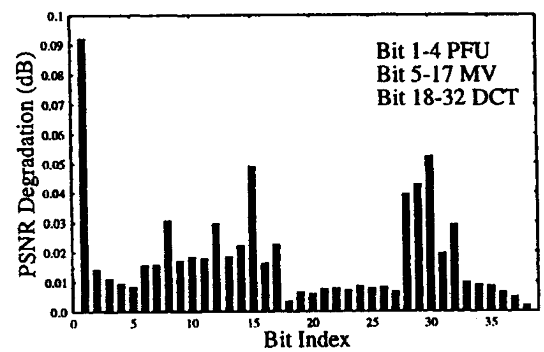


Fig. 5. Integrated PSNR bit sensitivities of Codec 1.

a high-activity sequence referred to as the 'Lab sequence.'[1] For 'Miss America' an average PSNR of about 33 dB was maintained, which was associated with pleasant videophone quality. The bit allocation scheme is summarized in Table I and the complexity of this codec is about 50 Mflops, which can be reduced to about 25 Mflops without significant performance penalty. In our further discourse we will refer to the above scheme as Codec 1. After addressing the bit sensitivity issues of Codec 1 we will propose a lower bit rate but more error sensitive arrangement, Codec 2, and analyze their advantages and disadvantages.

Source Sensitivity: In order to apply source-sensitivity matched protection the video bits were subjected to sensitivity analysis. In [9] we have consistently corrupted a single bit of a video coded frame and observed the image peak signal-to-noise ratio (PSNR) degradation inflicted. Repeating this method for all bits of a frame provided the required sensitivity figures and on this basis bits having different sensitivities can be assigned matching FEC codes. This technique, however, does not take adequate account of the phenomenon of error propagation across image frame boundaries. Therefore in this treatise we propose to use the method suggested in [17], where we corrupted each bit of the same type in the current frame and observed the PSNR degradation for the consecutive frames due to the error event in the current frame. As an example, Fig. 4 depicts the PSNR degradation profile in case of corrupting all 'No 1' Bits, the most significant bit (MSB) of the PFU and all 'No 11' Bits, one of the address bits of the MV, in frame 21. In the first case, the MSB of all PFU blocks are corrupted causing a scattered pattern of artifacts across the image. Those blocks will be replenished by the PFU exactly every 18 frames, revealed in the 'staircase' effect in Fig. 4. The impact of the corrupted MV is randomly distributed across the frame and hence, mitigated continuously by the PFU.

In order to quantify the overall sensitivity of any specific bit we have integrated (summed) the PSNR degradations over the consecutive frames, where they have had a measurable effect and averaged these values for all the occurrences of the corresponding bit errors. These results are shown in Fig. 5 for the 13 MV bits and 21 DCT bits of an 8×8 block, as well as for the 4 partial forced update bits.

2.2. *Codec 2*

In an attempt to improve the bandwidth efficiency of Codec 1 and to explore the range of design trade-offs, we have studied the statistical properties of the various parameters of Codec 1 in order to identify any persisting residual redundancy. We found that the motion activity table and the table of DCT-active blocks were potentially amenable to further data compression using run length coding (RLC). Therefore we set out to contrive a range of run length coded video codecs with bit rates as low as 5, 8, and 10 kbps, which we refer to as Codec 2.

The schematic diagram of Codec 2 is akin to that of Codec 1 shown in Fig. 7, but the above mentioned coding tables are further compressed by RLC. Similarly to Codec 1, the operation of Codec 2 is also initialized in the **intraframe mode,** where the encoder transmits the coarsely quantized block averages for the current frame. This provides a low-resolution initial frame required for the operation of the motion compensated **interframe** codec at both the commencement and during later stages of communications in order to prevent encoder/decoder misalignment. However, for the sake of maintaining a total bit rate R in the range of 5–10 kb/s for our 176×144 pixel CCITT standard QCIF images at a scanning rate of 10 frames/s we now limited the number of encoded bits per frame in Codec 2 to 500, 800, and 1000 b/frame, respectively. In order to

[1] The MA sequence encoded at 11.36 kbps can be viewed under the address www:/whirligig.ecs.soton.ac.uk/~jss.

transmit all block averages with a 4-b resolution, as in Codec 1, while not exceeding the above stipulated maximum bit rate, we fixed the initial intraframe block size to 10×10, 12×12, or 14×14 pixels for the above three target bit rates. The intraframe block size in Codec 1 was 10×10 pixels.

However, in the motion-compensation (MC) we retained the block-size of 8×8 and the search window size of 4×4 around the center of each block. Furthermore, the previously proposed gain-controlled MC and quad-class DCT quantization was invoked. 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 of the 396 blocks, marking it as passive or active. These tables are compressed using the elements of a two stage quad tree (QT) as follows.

First the 396-entry activity table containing the binary flags is grouped in 2×2 blocks and a four bit symbol is allocated to those blocks which contain at least one active flag. These four-bit 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 packetized to a symbol and then run length encoded. As a result, a typical 396-b active/passive table containing 30 active flags can be compressed to less than 150 b. The motion vectors do not lend themselves to run length encoding.

If at this stage of the encoding process the number of bits allocated to the compressed motion- and DCT-activity tables as well as to the active MV's exceeds half of the total number of available bits/frame, some of the blocks satisfying the initial motion-active criterion will be relegated to the motion-passive class. This process takes account of the subjective importance of various blocks and 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 frame. The DCT blocks are handled using a similar procedure. Depending on the actual fixed-length transmission burst and the free buffer space, a number of active DCT blocks is chosen and the corresponding compressed tables are determined. If the total bit count overspills the transmission burst or if there are too many bits left unused, a different number of active blocks is estimated and new tables are determined.

The PSNR versus frame index performance of a 5, 8, and 10 kbps RLC scheme is shown for the 'Miss America' sequence in Fig. 6 and the average results are summarized is Table II. Although due to the low-resolution intraframe mode at the commencement of communications it takes a few frames for the image to reproduce fine details, this effect is not objectionable. This is because the subjectively more important center of the screen is processed first. Fig. 6 demonstrates that at 5 kbps the codec operates at its limits and hence it takes a long time before the steady-state PSNR value is reached. However, at rates at or above 8 kbps a pleasant quality is maintained leading to an average PSNR in excess of 30 dB, which is exceeded in the center of the image. Based on these findings, in the run length coded System 2 we have opted for an **8.52 kbps** implementation of **Codec 2,** generating 852 b per frame and maintaining an average PSNR of about 33.3 dB for the MA sequence. We also note that in some of the proposed systems an **8 kbps** reduced-rate version of Codec 2 will be invoked, which we refer to as **Codec 2a.** Before we continue with the description of the source-matched transceiver schemes it must be emphasized that, in contrast to Codec 1 where no RLC is employed, if the RL-coded activity table bits are corrupted, the rest of that frame will be completely corrupted. Hence automatic repeat request (ARQ) techniques are preferred in the systems employing the RL-coded Codec 2. The sensitivity of the remaining bits is similar to that of the corresponding Class Two bits of Codec 1.

TABLE II
AVERAGE PSNR PERFORMANCE OF CODEC 2 FOR THE 'MISS AMERICA' AND 'LAB' SEQUENCES

Sequence	'Miss America'	'Lab'
5 kb/s	30.26 dB	21.87 dB
8 kb/s	33.29 dB	24.34 dB
10 kb/s	33.52 dB	26.91 dB

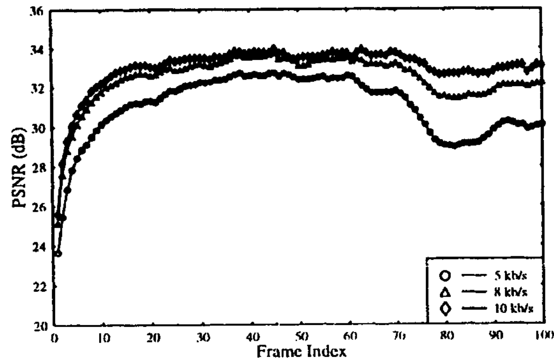


Fig. 6. PSNR versus frame index performance of Codec 2 for the 'Miss America' sequence.

III. MODULATION AND TRANSMISSION

Over mobile channels constant envelope modulation techniques, such as for example Gaussian Minimum Shift Keying (GMSK) used in the Pan-European GSM system [59] has successfully been applied. In contrast, until quite recently QAM research was mainly focused at applications over AWGN channels [35]. However, fuelled by the drive towards ever higher bandwidth efficiency and facilitated by advances, such as noncoherent star QAM [45], coherent pilot symbol assisted modulation [55] and the transparent tone in band [56], [57] (TTIB) technique, during the last few years its employment has also become realistic over mobile channels [36]–[48]. In order to achieve high bandwidth efficiency, QAM encodes information on both the phase and magnitude of the complex transmitted signal and hence it requires a linear transceiver, which suffer from low power efficiency [53], [54]. However, in low-power pico- or microcellular applications this is not a serious limitation, since the power consumption of the high-complexity digital circuitry is more crucial. In fact, due to its reduced signalling rate such a transceiver may be able to operate in a nondispersive scenario, without a channel equalizer, which reduces the power consumption.

The innate sensitivity of QAM against co-channel interference in an interference limited scenario is mitigated by the partitioning walls in indoors pico-cells and can be further reduced using the channel segregation algorithm proposed in [49]. Instead of tolerance-sensitive linear-phase Nyquist filtering nonlinear filtering (NLF) joining time-domain signal transitions with a smooth curve can be employed [26]. In case of coherent detection better performance can be achieved than using lower-complexity noncoherent differential modems. In order to phase-coherently recover the orthogonal quadrature carriers at the receiver, which will assist to recover the transmitted data, the Transparent-tone-in-band (TTIB) principle [56]–[58] or Pilot Symbol Assisted Modulation (PSAM) [55], [48] can be invoked.

Differentially coded noncoherent QAM modems [45] have typically low complexity than their coherent counterparts, but they inflict a characteristic 3 dB differential coding SNR penalty over AWGN channels, which persists also over Rayleigh channels. Hence they require higher SNR and SIR values than the more complex coherent schemes [26]. Therefore in our video transceiver second-order switched-diversity assisted coherent Pilot Symbol Assisted Modulation (PSAM) using the maximum-minimum-distance square QAM constellation is used [26].

IV. Source-Matched Transceiver

4.1. System 1

System Concept: The system's schematic is portrayed in Fig. 7, where the source encoded video bits generated by Codec 1 are split in two sensitivity classes and sensitivity matched channel coding/modulation is invoked. The proposed system was designed for mobile packet video telephony and it had two different modes of operation, namely 4-level and 16-level quadrature amplitude modulation (QAM) [26]. Our intention was to contrive a system, where the more benign propagation environment of indoors cells would benefit from the prevailing higher signal-to-noise ratio (SNR) by using bandwidth efficient 16QAM and thereby requiring only half the number of packets compared to 4QAM. When the portable station (PS) is handed over to an outdoors microcell or roams in a lower SNR region towards the edge of a cell, the base station (BS) instructs the PS to lower its number of modulation levels to 4 in order to maintain an adequate robustness under lower SNR conditions. Let us now focus our attention on specific details of System 1.

Sensitivity-Matched Modulation: Best robustness against channel errors is achieved, if sensitivity-matched forward error correction coding is used. Similarly to our approach in [18], Wei [25] has also suggested to use unequal protection multilevel coded modulation in order to achieve high bandwidth efficiency. Following similar principles here, in our proposed videophone schemes we will exploit that 16-level pilot symbol assisted quadrature amplitude modulation [26] (16-PSAQAM) provides two independent 2-b subchannels having different bit error rates (BER). Specifically, the BER of the higher integrity C1 subchannel is a factor 2–3 times lower than that of the lower quality C2 subchannel. Both subchannels support the transmission of two bits per symbol. This implies that the 16-PSAQAM scheme inherently caters for sensitivity-matched protection, which can be fine-tuned using appropriate FEC codes to match the source requirements. This property is not retained by the 4QAM scheme, but the required different protection for the source coded bits can be ensured using appropriately matched channel codecs.

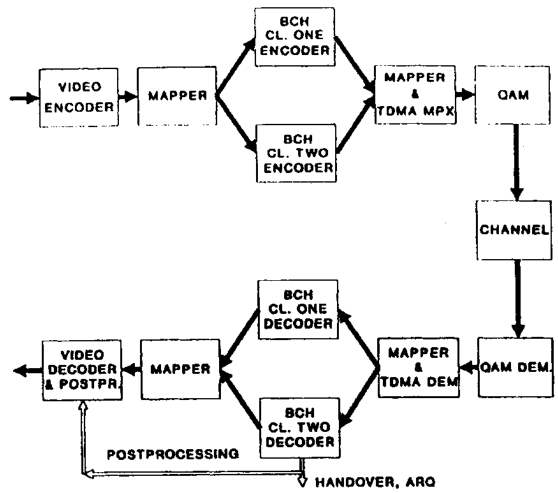


Fig. 7. System's schematic.

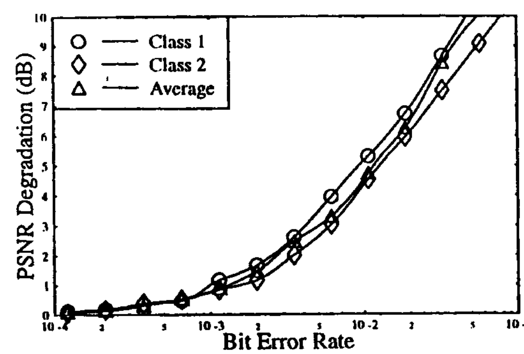


Fig. 8. PSNR versus BER degradation of Codec 1 for class one and two.

Source Sensitivity: In order to find the appropriate FEC code for our video codec, its output stream was split in two equal sensitivity classes, Class One and Two according to our findings in Fig. 5. Note that the notation Class One and Two introduced here for the more and less sensitive video bits is different from the higher and lower integrity C1 and C2 modulation channels. Then the PSNR degradation of both Class One and Two as well as the average PSNR degradation was evaluated for a range of BER values in Fig. 8. These results showed that a factor two lower BER was required by Class One bits than by Class Two bits, in order to maintain similar PSNR degradations in the range of 1–2 dB. These integrity requirements conveniently coincided with the integrity ratio of the C1 and C2 subchannels of our 16-PSAQAM modem [26]. Hence we can apply the same FEC protection to both Class One and Two source bits and direct Class One bits to the C1 16-PSAQAM subchannel, while Class Two bits to the C2 subchannel.

TABLE III
SUMMARY OF SYSTEM FEATURES

Feature	System 1	System 2	System 3	System 4	System 5
Video Codec	Codec 1	Codec 2	Codec 1	Codec 2a	Codec 2a
Video rate (kbps)	11.36	8.52	11.36	8	8
Frame Rate (fr/s)	10	10	10	10	10
C1 FEC C2 FEC Header FEC	BCH(127,71,9) BCH(127,71,9) BCH(127,50,13)	BCH(127,50,13) BCH(127,92,5) BCH(127,50,13)	BCH(127,71,9) BCH(127,71,9) BCH(127,50,13)	BCH(127,50,13) BCH(127,50,13) BCH(127,50,13)	(BCH(127,50,13) (BCH(127,50,13) (BCH(127,50,13)
FEC-coded Rate (kbps)	20.36	15.24	20.36	20.36	20.36
Modem	4/16-PSAQAM	4/16-PSAQAM	4/16-PSAQAM	4/16-PSAQAM	4/16-PSAQAM
ARQ	None	Cl. One	Cl. One & Two	Cl. One & Two	None
User Signal. Rate (kBd)	18 or 9	6.66	18 or 9	18 or 9	18 or 9
System Signal. Rate (kBd)	144	144	144	144	144
System Bandwidth (kHz)	200	200	200	200	200
No. of Users	8 or 16	(21-2)=19	6 or 14	6 or 14	8 or 16
Eff. User Bandwidth (kHz)	25 or 12.5	10.5	33.3 or 14.3	33.3 or 14.3	25 or 12.5
Min. AWGN SNR (dB)	7 or 15	15	6 or 13	5 or 11	8 or 12
Min. Rayleigh SNR (dB)	12 or 20	25	7 or 20	7 or 14	15 or 16

Forward Error Correction: Both convolutional and block codes can be successfully used over mobile radio links [50], but in our proposed scheme we have favored binary Bose–Chaudhuri–Hocquenghem (BCH) codes. BCH codes combine a good burst error correction capability with reliable error detection, a facility useful to invoke image post-enhancement, to monitor the channel's quality and to control handovers between traffic cells. The preferred $R = 71/127 \approx 0.56$-rate BCH(127,71,9) code can correct 9 errors in a block of 127 b, a correction capability of about 7.1%. The number of channel coded bits per image frame becomes $1136 \times 127/71 = 2032$, while the bit rate is 20.32 kbps at an image frame rate of 10 frames/s.

Transmission Format: The transmission packets are constructed using one Class One BCH(127,71,9) code, one Class Two BCH(127,71,9) code, and a stronger BCH(127,50,13) is allocated to the packet header, yielding a total of 381 b per packet. In case of 16QAM these are represented by 96 symbols and after adding 11 pilot symbols using a pilot spacing of $P = 10$ as well as 4 ramp symbols to ensure smooth power amplifier ramping the resulting 111-symbol packets are transmitted over the radio channel. Eight such packets represent a whole image frame and hence the signalling rate becomes 111 symb/12.5 ms $\approx$ 9 kBd. When using a time division multiple access (TDMA) channel bandwidth of 200 kHz, such as in the Pan-European second generation mobile radio system known as GSM and a modulation excess bandwidth of 38.8%, the signalling rate becomes 144 kBd. This allows us to accommodate 144/9 = 16 users, which coincides with the number of so-called half-rate speech users supported by the GSM system [59].

When the prevailing channel SNR does not allow 16QAM communications, 4QAM must be invoked. In this case the 381-b packets are represented by 191 2-b symbols and after adding 20 pilot symbols and 4 ramp symbols the packet-length becomes 225 symb/12.5 ms, yielding a signalling rate of 18 kBd. In this case the number of videophone users supported by System 1 becomes 8, as in the full-rate GSM speech channel. The system also facilitates mixed-mode operation, where 4QAM users must reserve two slots in each 12.5 ms TDMA frame towards the fringes of the cell, while in the central section of the cell 16QAM users will only require one slot per frame in order to maximize the number of users supported. Assuming an equal proportion of 4 and 16QAM users the average number of users per carrier becomes 12. The equivalent user bandwidth of the 4QAM PS's is 200 kHz/8 = 25 kHz, while that of the 16QAM users is 200 kHz/16 = 12.5 kHz.

For very high quality mobile channels or for conventional telephone lines 64-QAM can be invoked, which further reduces the required bandwidth at the cost of a higher channel SNR demand. However, the packet format of this mode of operation is different from that of the 16 and 4QAM modes and hence requires a different slot length. The 381-b payload of the packet is represented by 64 6-b symbols, four ramp symbols are added along with 14 pilot symbols, which corresponds to a pilot spacing of $P = 5$. The resulting 82-symbol/12.5 ms packets are transmitted at a signalling rate of 6.6 kBd, which allows us to host 22 videophone users. The user bandwidth becomes 200 kHz/22 $\approx$ 9.1 kHz.

The above-mentioned features of the 16/4QAM System 1 along with the characteristics of a range of other systems about to be introduced in the next section are summarized in Table III.

Clearly, the required signalling rate and bandwidth are comparable to those of most state-of-art mobile radio speech links, which renders our scheme attractive for mobile video telephony in the framework of existing mobile radio systems. Furthermore, this rate can also be readily accommodated by conventional telephone subscriber loops.

4.2. System 2

In order to improve the bandwidth efficiency of Codec 1 we introduced the run-length coded Codec 2. Hence Codec 2 became more vulnerable against transmission errors than Codec 1 and their effect is particularly objectionable, if the run length coded activity table bits are corrupted. Therefore in System 2, which was designed to incorporate Codec 2, the more sensitive run length coded activity table bits are protected by the powerful binary Bose–Chaudhuri–Hocquenghem BCH(127,50,13) codec, while the less vulnerable remaining bits by the weaker BCH(127,92,5) code. Note that the overall coding rate of $R = (50 + 92)/(127 + 127) \approx 0.63$ is identical to that of System 1, but the RL-coded Class One bits are more strongly protected. At a fixed coding rate this inevitably assumes a weaker code for the protection of the less vulnerable Class Two bits. The 852 b/100 ms video frame is encoded using six pairs of such BCH code words, yielding a total of $6 \cdot 254 = 1524$ b, which is equivalent to a bit rate of 15.24 kbps.

As in System 1, the more vulnerable run length and BCH(127,50,13) coded Class One bits are then transmitted over the higher integrity C1 16QAM subchannel. The less sensitive BCH(127,92,5) coded Class Two DCT coefficient bits are conveyed using the lower-integrity C2 16QAM subchannel. This arrangement is favored in order to further emphasize the integrity differences of the BCH codecs used, which is necessitated by the integrity requirements of the video bits.

The transmission burst is constructed by adding an additional BCH(127,50,13) code word for the packet header and the resulting 381 b are again converted to 96 16QAM symbols, and pilot as well as ramp symbols are added. In System 2 six such packets represent a video frame, hence the single-user signalling rate becomes 666 symb/100 ms, which corresponds to 6.66 kBd. This allows us to accommodate now Integer[144kBd/6.66] = 21 such users, if no time slots are reserved for packet re-transmissions. This number will have to be reduced in order to accommodate ARQ's.

Automatic Repeat Request: ARQ techniques have been successfully used in data communications [29]–[32] in order to render the bit and frame error rate arbitrarily low. However, due to their inherent delay and the additional requirement for a feed-back channel for message acknowledgement they have not been employed in interactive speech or video communications. In our packet video system however there exists a full duplex control link between the BS and PS, which can be used for acknowledgements and the short TDMA frame length ensures a low packet delay, hence ARQ can be invoked.

In System 2 when the more powerful BCH codec conveying the more sensitive run-length coded Class One bits over the C1 16QAM subchannel is overloaded by channel errors, we re-transmit these bits only using robust 4QAM. Explicitly, for the first transmission attempt (TX1) we use contention-free Time Division Multiple Access (TDMA). If an ARQ-request occurs, the re-transmitted packets will have to contend for a number of earmarked time slots similarly to Packet Reservation Multiple Access (PRMA) [26]. The intelligent base station (BS) detects these events of packet corruption and instructs the portable stations (PS) to re-transmit their packets during the slots dedicated to ARQ-packets. Reserving slots for ARQ-packets reduces the number of video users supported depending on the prevailing channel conditions, as we will show in the Results Section, Section V.

Although the probability of erroneous packets can be reduced by allowing repeated re-transmissions, there is a clear trade-off between the number of maximum transmission attempts and the BCH-coded frame error rate (FER). In order to limit the number of slots required for ARQ-attempts, which potentially reduce the number of video users supported, in System we invoke ARQ only, if the more sensitive run-length coded Class One bits transmitted via the C1 16-PSAQAM channel and protected by the BCH(127,50,13) codec are corrupted. Furthermore, we re-transmit only Class One bits, but in order to insure a high success rate, we use 4-PSAQAM, which is more robust than 16-PSAQAM. Since only half of the information bits are re-transmitted, they can be accommodated within the same slot interval and same bandwidth, as the full packet. If there are only C2 bit errors in the packet, it is not re-transmitted, which implies that typically there will be residual Class Two errors. In order to limit the number slots dedicated to re-transmissions we limited the number of transmission attempts to three, which implies that a minimum of two slots per frame must be reserved for ARQ. In order to maintain a low system complexity we dispense with any contention mechanism and allocate two time slots to that particular user, whose packet was first corrupted within the TDMA frame. Further users cannot therefore invoke ARQ, since there are no more unallocated slots. A further advantage is that in possession of three copies of the transmitted packet majority decisions can be invoked, if all three packets became corrupted. The basic features of System 2 designed to accommodate Codec 2 are also summarized in Table III.

4.3. Systems 3–5

In order to explore the whole range of available trade-offs we have contrived three further systems, namely Systems 3–5.

System 3 uses the same video and FEC codecs as well as modems as System 1, but it allows a maximum of three transmission attempts in case of C1 BCH(127,71,9) decoding errors. If there are only C2 errors no ARQ is invoked. Employing ARQ in System 1 constitutes a further trade-off in terms of reducing the number of subscribers supported by two, while potentially improving the communications quality at a certain BER, or allowing an expansion of the range of operating channel SNR towards lower values.

System 4 employs a run-length coded source compression scheme similar to Codec 2, but with a slightly reduced bit rate of 8 kbps or 800 b per frame, which we refer to as **Codec 2a.** This system followed the philosophy of System 3, but Class One and Two bits were protected by the more powerful BCH(127,50,13) code instead of the BCH(127,71,9) scheme. The slightly reduced video rate of 8 kbps was imposed in order to be able to accommodate the BCH(127,50,13) code in both 16QAM subchannels, while maintaining the same 20.36 kbps overall rate, as Systems 1 and 3. Clearly, System 4 will allow us to assess, whether it is a worthwhile complexity investment to introduce run-length coding in Codec 1 in order to reduce the source bit rate and whether the increased error sensitivity of Codec 2a can be compensated for by accommodating the more complex and more powerful BCH(127,50,13) codec.

In order to maximise the number of video subscribers supported, the performance of System 4 can also be studied without ARQ techniques. We will refer to this scheme as **System 5.** Again, these system features are summarized in Table III. Having designed the video transceivers their performance results are presented in the next section.

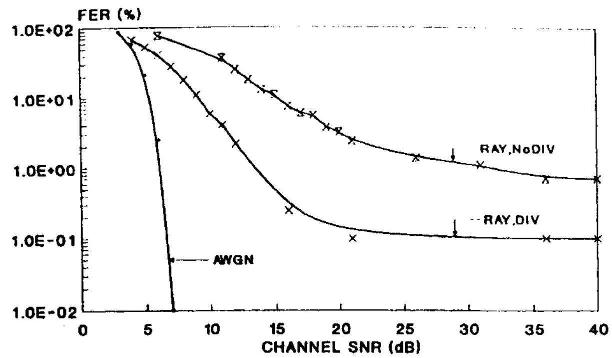


Fig. 9. 4QAM BCH(127,71,9) FER versus channel SNR performance of the 18 kBd mode of operation of System 1 over AWGN and Rayleigh (RAY) channels with diversity (DIV) and without diversity (NoDIV).

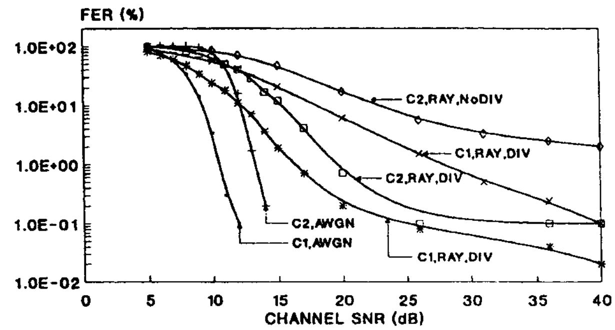


Fig. 10. 16QAM C1 and C2 BCH(127,71,9) FER versus channel SNR performance of the 9 kBd mode of operation of System 1 over AWGN and Rayleigh (RAY) channels with diversity (DIV) and without diversity (NoDIV).

V. System Performance

5.1. Performance of System 1

In our experiments the signalling rate was 144 kBd, while the propagation frequency and the vehicular speed were 1.8 GHz and 30 mph, respectively. For pedestrian speeds the fading envelope fluctuates less dramatically and hence our experimental conditions constitute a GSM-like urban worst-case scenario.

Here we characterise the performance of the transceiver in terms of the BCH(127,71,9) coded frame error rate (FER) versus channel signal-to-noise ratio (SNR), as portrayed in Figs. 9 and 10 in case of the 4QAM and 16QAM modes of operation of System 1. In these figures we displayed the FER over both AWGN and Rayleigh channels, in case of the latter both with and without diversity. Note that for near-unimpaired video quality the FER must be below 1%, but preferably below 0.1%. This requirement is satisfied over AWGN channels for SNR's in excess of about 7 dB for 4QAM. In case of 16QAM and AWGN channels the C1 and C2 FER's are reduced to about 0.1% for SNR's above 13 and 15 dB, respectively. Observe in the figures that over Rayleigh-fading (RAY) channels with diversity (DIV) the corresponding FER values are increased to about 15 dB for 4QAM and 20 dB for 16QAM, while without diversity (NoDIV) further increased SNR values are necessitated.

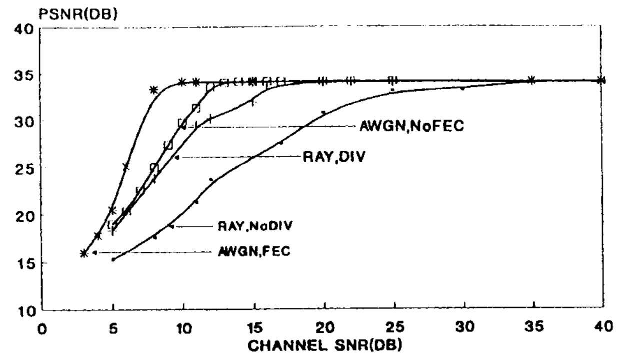


Fig. 11. 4QAM PSNR versus channel SNR performance of System 1 in its 18 kBd mode of operation over various channels.

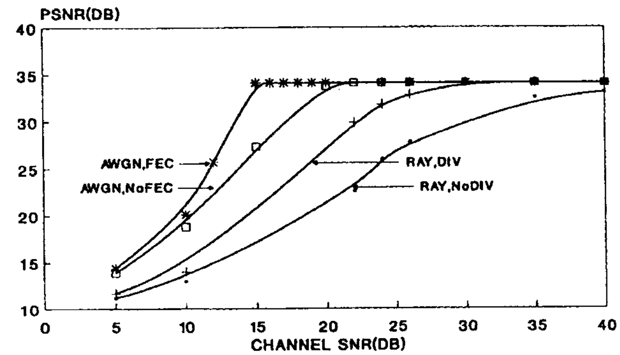


Fig. 12. 16QAM PSNR versus channel SNR performance of System 1 in its 9 kBd mode of operation over various channels.

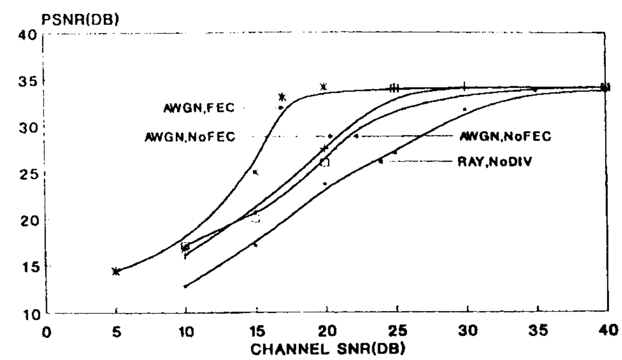


Fig. 13. 64QAM PSNR versus channel SNR performance of System 1 in its 6.6 kBd mode of operation over various channels.

The overall video PSNR versus channel SNR (ChSNR) performance of System 1 is shown in Figs. 11 and 12 for the 4QAM and 16QAM modes of operation, respectively. The PSNR versus ChSNR characteristics of the 6.6 kBd 64QAM arrangement are also given for the sake of completeness in Fig. 13. Observe in the above PSNR versus channel SNR figures that the AWGN performance was evaluated also without forward error correction (FEC) coding in order to indicate the expected performance in a conventional AWGN environment, such as telephone or satellite channels without FEC coding.

Due to its limited bandwidth efficiency gain, high SNR requirement and incompatible slot structure we recommend the 64QAM system for applications, where the bandwidth is at absolute premium and in our further discourse we favor the

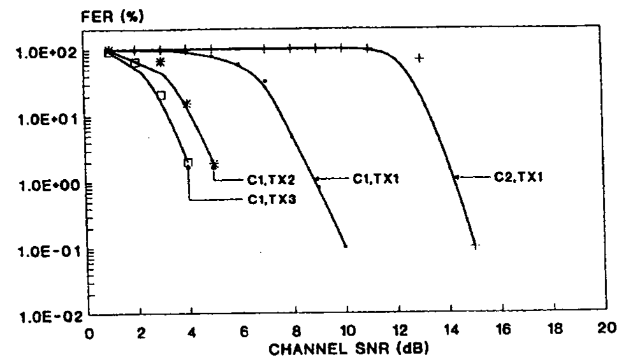


Fig. 14. BCH(127,50,13) and BCH(127,92,5) FER versus channel SNR performance of System 2 over AWGN channels.

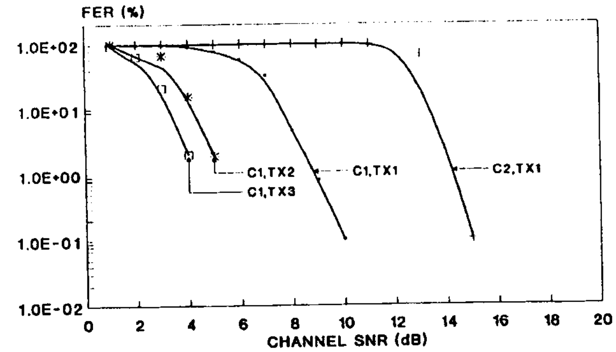


Fig. 15. BCH(127,50,13) and BCH(127,92,5) FER versus channel SNR performance of System 2 over Rayleigh channels without diversity.

16QAM/4QAM modes of System 1. The corresponding figures suggest that best performance was achieved over AWGN channels with FEC, requiring a channel SNR of about 15 and 7 dB in case of the 16QAM and 4QAM modes of operation, respectively, in order to achieve an unimpaired image quality associated with a PSNR value of about 34 dB. Without FEC coding over AWGN channels these SNR values had to be increased to about 20 and 12 dB, respectively. Over Rayleigh channels with second order diversity the system required ChSNR values of about 15 and 25 dB in the 4QAM and 16QAM modes in order to reach an image PSNR within 1 dB of its unimpaired value of 34 dB. This 1 dB PSNR degradation threshold will be used in all scenarios to characterise the near unimpaired image quality. Lastly, without diversity over Rayleigh channels SNR's of about 25 and 33 dB were needed for near-unimpaired PSNR performance in the 4QAM and 16QAM modes.

5.2. *Performance of System 2*

FER versus ChSNR: In order to evaluate the overall video performance of System 2, 100 frames of the MA sequence were encoded and transmitted over both the best-case Additive White Gaussian Noise (AWGN) channel and the worst-case narrowband Rayleigh-fading channel. The BCH(127,50,13) and BCH(127,92,5) decoded frame error rate (FER) was evaluated for both the C1 and C2 bits after the first transmission attempt (TX1) over AWGN and Rayleigh channels with and without second-order diversity, as seen in Figs. 14–16. These figures also portray the C1 FER after the second (TX2) and third (TX3) transmission attempts, which were carried out using 4QAM in order to maximise the success rate of the C1 bits, representing the vulnerable run-length coded activity table.

Over AWGN channels a C1 FER of less than 1% can be maintained for channel SNR's in excess of about 5 dB, if three transmission attempts are allowed, although at such low SNR's the C2 bit errors are inflicting an unacceptably high video degradation. The corresponding C2 FER over AWGN channels becomes sufficiently low for channel SNR's above about 14–15 dB in order to guarantee unimpaired video communications, which is significantly higher than that required by the C1 subchannel. Over the Rayleigh channel, but without diversity a channel SNR of about 12 dB was required

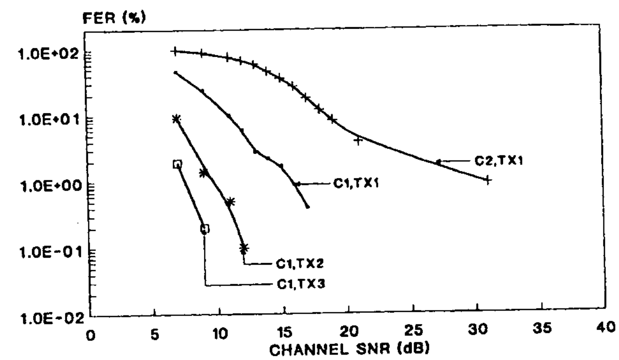


Fig. 16. BCH(127,50,13) and BCH(127,92,5) FER versus channel SNR performance of System 2 over Rayleigh channels with diversity.

with maximum three transmissions in order to reduce the C1 FER below 1% or FER $= 10^{-2}$, as shown in Fig. 15. But the C2 FER curved flattened out for high SNR values, which resulted in a severe 'leakage' of erroneous C2 bits and this resulted in a somewhat impaired video performance. When diversity reception was used, the minimum required SNR value necessary to maintain a similar C1 FER was reduced to around 10 dB, while the C2 FER became adequately low for SNR's in excess of about 20–25 dB, as demonstrated by Fig. 16.

Slot Occupancy: As mentioned before, the ARQ attempts require a number of reserved time slots, for which the re-transmitting MS's have to contend. When the channel SNR is too low, there is a high number of re-transmitted packets contending for too low a number of slots. The slot occupancy increase versus channel SNR performance, which was defined as the ratio of original packets to total transmitted packets, is portrayed for a range of scenarios in Fig. 17. For SNR values in excess of about 10, 15, and 25 dB, when using 16QAM over AWGN as well as Rayleigh channels with and without diversity, respectively, the slot occupancy was increased due to re-transmissions only marginally. Therefore reserving two time slots per frame for a maximum of two re-transmission attempts ensures a very low probability of packet collision during ARQ operations, while reducing the number of subscribers supported by two. In a simplistic approach this would imply that for a channel SNR value, where the FER is below 1% and assuming 20 users the reserved ARQ slots will be only

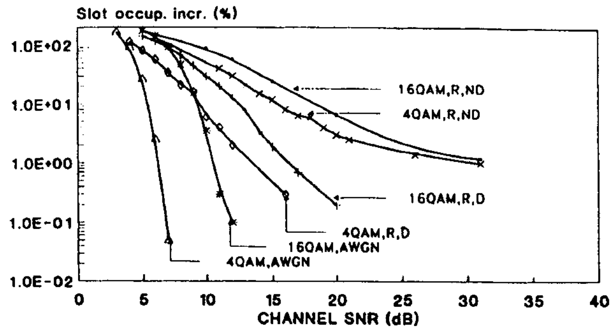


Fig. 17. Slot occupancy increase versus channel SNR performance of the proposed transceivers over AWGN and Rayleigh channels with (D) and without (ND) diversity.

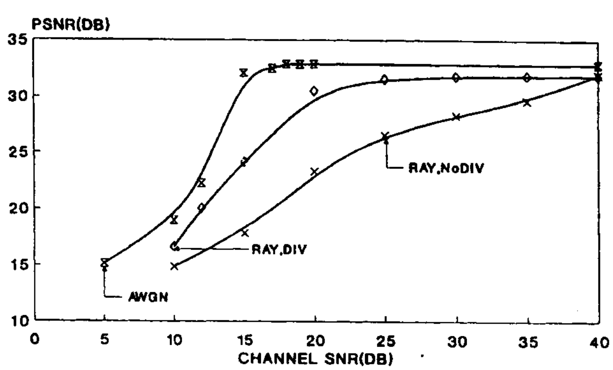


Fig. 18. PSNR versus channel SNR performance of the 6.6 kBd System 2 over various channels.

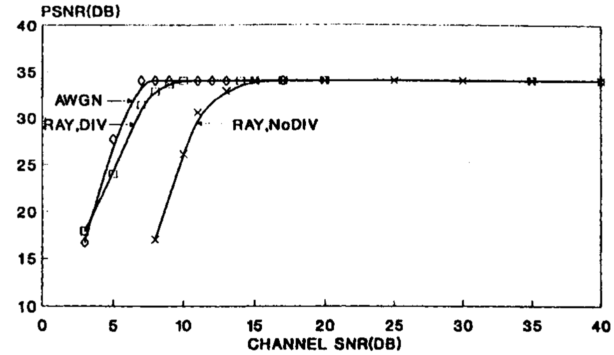


Fig. 19. PSNR versus channel SNR performance of the 18 kBd 4QAM mode of System 3 over various channels using three transmission attempts.

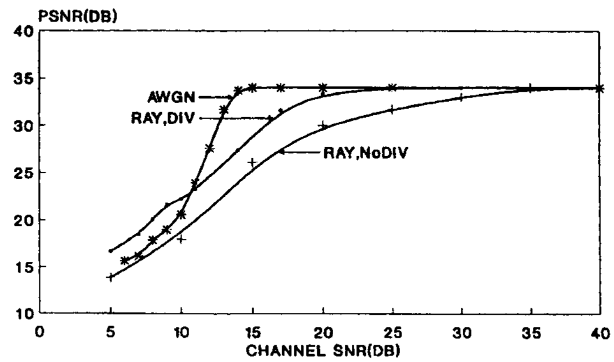


Fig. 20. PSNR versus channel SNR performance of the 9 kBd 16QAM mode of System 3 using three transmission attempts over various channels.

occupied in about every fifth frame. However, we cannot earmark less than two slots for two additional transmission attempts. The 4QAM slot occupancy is even more modest at a given channel SNR than that of the 16QAM schemes, as suggested by Fig. 17.

The PSNR versus ChSNR performance of System 2 is characterized by Fig. 18. Observe that over AWGN channels ChSNR values in excess of 15 dB are required for unimpaired video performance. Over Rayleigh channels with diversity about 20 dB ChSNR is necessitated for an unimpaired PSNR performance, while without diversity the PSNR performance seriously suffers from the leaking Class Two bit errors. Overall, the 6.6 kBd System 2 has a lower robustness than the 9 kBd System 1, since its behavior is predetermined by the initially transmitted Class Two video bits, which were protected by the weaker BCH(127,92,5) code. Recall that System 1 used the BCH(127,71,9) code in both the C1 and C2 subchannels. In fact the performance of the 6.6 kBd System 2 is more similar to that of the 6.6 kBd 64QAM system characterized in Fig. 13, which does not use ARQ. Retransmission attempts to improve the integrity of the initially received Class One bits and ensure an adequate integrity for these vulnerable bits, but without enhancing the quality of the initial 16QAM C2 subchannel, System 2 cannot outperform System 1. Furthermore, System 2 is inherently more complex than System 1 and only marginally more bandwidth efficient. Therefore in contriving the remaining systems we set out to improve the noted deficiencies of System 2.

5.3. Performance of Systems 3–5

Clearly, our experience with System 2 suggested that it was necessary to re-transmit both Class One and Two bits, if the overloading of the C1 FEC codec indicated poor channel conditions. This plausible hypothesis was verified using System 3, which is the ARQ-assisted System 1. This allowed us to assess the potential benefit of ARQ's in terms of the minimum required channel SNR, while its advantages in terms of FER reduction were portrayed in Figs. 14–16. The corresponding PSNR curves of System 3 are plotted in Figs. 19 and 20 for its 18 kBd 4QAM and 9 kBd 16QAM modes, respectively. Re-transmission was invoked only, if the C1 FEC decoder was overloaded, but in these cases both the C1 and C2 subchannels were re-transmitted. Comparison with Figs. 11 and 12 revealed very substantial ChSNR reduction over Rayleigh channels, in particular without diversity. This was due to the fact that in case of a BCH frame error by the time of the second or third transmission attempt the channel typically emerged from a deep fade. Over AWGN channels the channel conditions during any further ARQ attempts were similar to those during the previous ones, hence ARQ offered more limited ChSNR reduction. The minimum required ChSNR values for the 4QAM mode over AWGN and Rayleigh channels with and without diversity are 7, 8, and 13 dB, while for the 16QAM mode 13, 18, and 27 dB, respectively, as also shown in Table III.

In System 4 the employment of ARQ was more crucial than in System 3, since the corrupted run-length coded activity

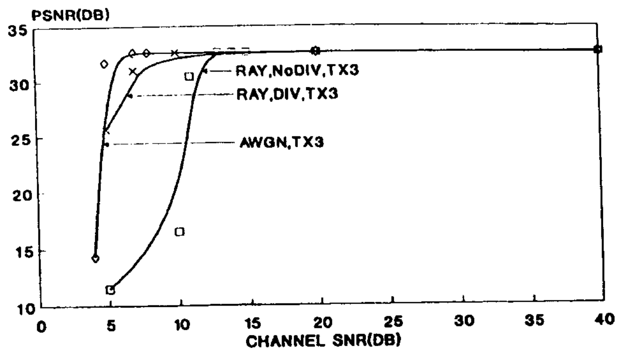


Fig. 21. PSNR versus channel SNR performance of the 18 kBd 4QAM mode of System 4 using three transmission attempts over various channels.

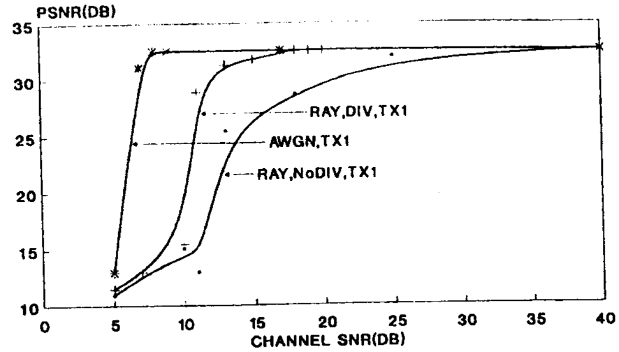


Fig. 23. PSNR versus channel SNR performance of the 18 kBd 4QAM mode of System 5 over various channels.

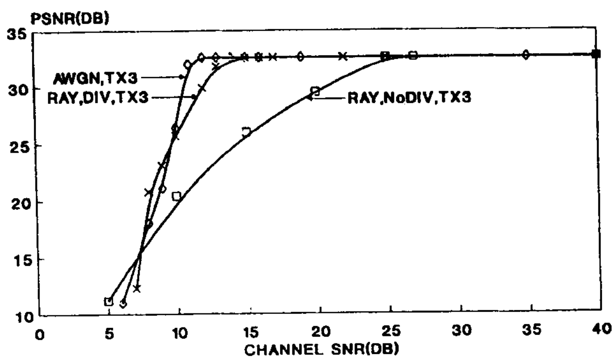


Fig. 22. PSNR versus channel SNR performance of the 9 kBd 16QAM mode of System 4 using three transmission attempts over various channels.

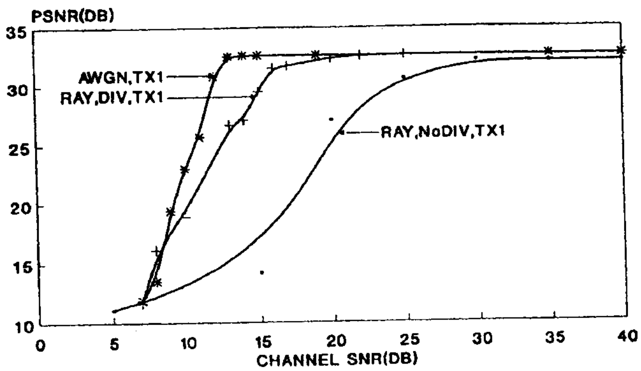


Fig. 24. PSNR versus channel SNR performance of the 9 kBd 16QAM mode of System 5 over various channels.

tables would inflict severe quality degradations for the whole frame. The corresponding PSNR curves are portrayed for the 18 and 9 kBd 4QAM and 16QAM operating modes in Figs. 21 and 22 over various channels, which can be contrasted with the results shown for System 5 without ARQ in Figs. 23 and 24. Again, over Rayleigh channels the ChSNR requirement reductions due to ARQ are substantial, in particular without diversity, where the received signal typically emerges from a fade by the time ARQ takes place. Over AWGN channels the benefits of ARQ are less dramatic, but still significant. This is because during re-transmission each packet faces similar propagation conditions, as during its first transmission. The required ChSNR thresholds for near-perfect image reconstruction in the 4QAM mode of System 4 are about 6, 8, and 12 dB over AWGN and Rayleigh channels with and without diversity, which are increased to 11, 14, and 27 dB in the 16QAM mode. In contrast, System 5 necessitates ChSNR's of 8, 13, and 25 dB as well as 12, 16, and 27 dB under the previously stated conditions over 4QAM and 16QAM, respectively.

VI. Summary and Conclusions

A range of bandwidth efficient, fixed-rate mobile videophone transceivers have been presented, which retain the features summarized in Table III. The video source rate can be fixed to any arbitrary value in order be able to accommodate the videophone signal by conventional 2nd generation mobile radio speech channels, such as for example that of the Pan-European GSM system [59], the Japanese PDC [19], and the American IS-54 [20] as well as IS-95 systems [21] at bit rates between 6.7 and 13 kbps.

In System 1 the 11.36 kbps Codec 1 was used, which has a lower rate than the 13 kbps speech rate of the GSM system. After BCH(127,71,8) coding the channel rate becomes 20.32 kbps. When using an adaptive transceiver, which can invoke 16QAM and 4QAM depending on the channel conditions experienced, the signalling rate becomes 9 and 18 kBd, respectively. Accordingly, 16 or 8 videophone users can be accommodated in the GSM bandwidth of 200 kHz, which implies user bandwidths of 12.5 and 25 kHz, respectively. Over line-of-sight AWGN channels SNR values of about 15 and 7 dB are required, when using 16QAM and 4QAM, respectively, in order to maintain unimpaired PSNR values of about 34 dB. An increased channel SNR of about 20 and 12 dB is needed over the diversity-assisted Rayleigh scenario.

In System 2 we have opted for an 8.52 kbps videophone codec, maintaining a PSNR of about 33 dB for the MA sequence. The source-coded bit stream was sensitivity-matched binary BCH(127,50,13) and BCH(127,92,5) coded and transmitted using pilot assisted 16QAM. Due to the lower source coded rate of 8.52 kbps of Codec 2 the single-user signalling rate of System 2 was reduced to 6.66 kBd, allowing us to accommodate $21 - 2 = 19$ video-telephone users in the 200 kHz GSM bandwidth. If the signal-to-interference ratio (SIR) and signal-to-noise ratio (SNR) values are in excess of about 15, and 25 dB over the AWGN and diversity-assisted nondispersive Rayleigh fading channels, respectively, pleasant videophone quality is maintained. The implementation complexity of System 1 is lower than that of System 2, while System 2 can accommodate more users, although it is less robust, as also demonstrated by Table III. This is due to the fact that only the C1 bits are re-transmitted.

On the basis of our experience with System 2 the fully ARQ-assisted System 3 was contrived, which provided a better image quality and a higher robustness, but was slightly less bandwidth efficient than Systems 1 and 2 due to reserving two time slots for ARQ. Furthermore, the question arose, whether it was better to use the more vulnerable run-length coded Codec 2a with stronger and more complex FEC protection, as in Systems 4 and 5, or the slightly higher rate Codec 1 with its weaker and less complex FEC was preferable. In terms of robustness System 5 proved somewhat more attractive than System 4, although the performances of the non-ARQ based System 1 and System 5 are rather similar.

Overall, using schemes similar to the proposed ones mobile videotelephony is becoming realistic over existing mobile speech links, such as the Pan-European GSM system [59], the Japanese PDC [19], and the American IS-54 [20] as well as IS-95 systems [21] at bit rates between 6.7 and 13 kbps. Our future work in this field will be targeted at improving the complexity/quality balance of the proposed schemes using a variety of other video codecs, such as parametrically assisted quad-tree and vector quantized codecs. A further important research area to be addressed is devising reliable transceiver reconfiguration algorithms.

ACKNOWLEDGMENT

The helpful suggestions of the anonymous Reviewers are gratefully acknowledged.

REFERENCES

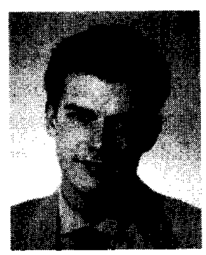
[1] "Speech and image coding," *IEEE J. Select. Areas Commun.*, Special Issue, vol. 10, no. 5, pp. 793–976, June 1992.
[2] *IEEE IEEE Trans. Circuits Syst. Video Technol.*, Special Issue Very Low Bit Rate Video Coding, vol. 4, no. 3, pp. 213–357, June 1994.
[3] B. Girod *et al.*, Guest Eds., "Special issue on image sequence compression," *IEEE Image Process.*, vol. 3, no. 5, Sept. 1994.
[4] *Image communications*. New York: Elsevier, (Special Issue on MPEG4 Techniques), to appear in July 1995.
[5] M. Khansari, A. Jalali, E. Dubois, and P. Mermelstein, "Robust low bit-rate video transmission over wireless access systems," in *Proc. Int. Commun. Conf. ICC'94*, pp. 571–575.
[6] R. Mann Pelz, "An un-equal error protected $p \times 8$ kbit/s video transmission for DECT," in *Proc. VTC'94*, Stockholm, Sweden, June 8–10, 1994, pp. 1020–1024.
[7] K. Watanabe *et al.*, "A study on transmission of low bit-rate coded video over radio links," in *Proc. VTC'94,*, pp. 1025–1029.
[8] L. Hanzo *et al.*, "A mobile speech, video and data transceiver scheme," in *Proc. VTC'94*, Stockholm, Sweden, June 8–10, 1994, pp. 452–456.
[9] R. Stedman, H. Gharavi, L. Hanzo, and R. Steele, "Transmission of subband-coded images via mobile channels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 1, pp. 15–27, Feb. 1993.
[10] ISO/IEC 11172 MPEG 1, International Standard, "Coding of moving pictures and associated audio for digital storage media up to about 1.5 Mbit/s," Parts 1–3.
[11] ISO/IEC CD 13818 MPEG 2, International Standard, "Information technology, generic coding of moving video and associated audio information," Parts 1–3.
[12] P. Sheldon, J. Cosmas, and A. Permain, "Dynamically adaptive control system for MPEG-4," in *Proc. 2nd Int. Wkshp. Mobile Multimedia Commun.*, Bristol, U.K., Apr. 11–13.
[13] "Advanced communications technologies and services (ACTS)," Workplan, European Commission, DGXIII-B-RA946043-WP, Aug. 1994.
[14] J. W. Woods, Ed., *Subband Image Coding*. Norwell, MA: Kluwer, 1991.
[15] H. Gharavi, "Subband coding of video signal," in *Subband Image Coding*, J. W. Woods, Ed. Norwell, MA: Kluwer, 1991, pp. 229–271.
[16] K. N. Ngan and W. L. Chooi, "Very low bit rate video coding using 3D subband approach," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 3, pp. 309–316, June 1994.
[17] L. Hanzo and J. P. Woodard, "An intelligent cordless voice terminal for indoors communications," submitted to *IEEE Trans. Vehic. Technol.*, 1994.
[18] L. Hanzo, R. Steele, and P. M. Fortune, "A subband coding, BCH coding and 16-QAM system for mobile radio communication," *IEEE Trans. Vehic. Technol.*, vol. 39, no. 4, pp. 327–340, Nov. 1990.
[19] "Public Digital Cellular (PDC) Standard," Res. and Develop. Ctr. for Radio Systems, Japan, RCR STD-27.
[20] "Dual-mode subscriber equipment—Network equipment compatibility specification," Telcommun. Industry Assoc. (TIA) Interim Standard IS-54, Wash., DC, 1989.
[21] "Mobile station—Base station compatibility standard for dual-mode wideband spread spectrum cellular system," EIA/TIA Interim Standard IS-95, Wash., DC, July 1993.
[22] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379-423 and 623–656, June and Oct. 1948.
[23] J. Hagenauer, "Quellengesteuerte Kanalcodierung fuer Sprach- und Tonuebertragung im Mobilfunk," in *Aachener Kolloquium: Signaltheorie,* Mobile Kommunikationssysteme, Mar. 23–25, 1994, pp. 67–76.
[24] A. J. Viterbi, "Wireless digital communication: A view based on three lessons learned," *IEEE Communications Mag.*, pp. 33–36, Sept. 1991.
[25] L. F. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, vol. 41, no. 10, pp. 1439–1450, Oct. 1993.
[26] W. T. Webb and L. Hanzo, *Modern Quadrature Amplitude Modulation: Principles and Applications for Wireless Communications*. London: IEEE-Pentech, 1994.
[27] L. Hanzo, R. Stedman, R. Steele, and J. C. S. Cheung, "A portable multimedia communicator scheme," in *Chapter in Multimedia Technologies and Future Applications,* R. I. Damper *et al.,*, Ed. London: Pentech, 1994, pp. 31–54.
[28] A. K. Jain, *Fundamentals of digital image processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[29] G. M. Djuknic and D. L. Schilling, "Performance analysis of an ARQ transmission scheme for meteor burst communications," *IEEE Trans. Commun.*, vol. 42, pp. 268–271, Feb./Apr. 1994.
[30] L. de Alfaro and A. R. Meo, "Codes for second and third order GH-ARQ schemes," *IEEE Trans. Commun.*, vol. 42, pp. 988–910, Feb./Apr. 1994.
[31] T.-H. Lee, "Throughput performance of a class of continuous ARQ strategies for burst-error channels," *IEEE Trans. Vehic. Technol.*, vol. 41, no. 4, pp. 380–386, Nov. 1992.
[32] S. Lin, D. J. Costello, and M. J. Miller, "Automatic-repeat-request error-control schemes," *IEEE Comm. Mag.*, pp. 5–17, Dec. 1984.
[33] *IEEE J. Select. Areas Commun.*, Speech and image coding, vol. 10, no. 5, pp. 793–976, June 1992.
[34] R. A. Salami and L. Hanzo *et al.*, "Speech coding," in *Mobile Radio Communications*, R. Steele, Ed. London: Pentech, 1992, pp. 186–346.
[35] G. D. Forney *et al.*, "Efficient modulation for band-limited channels," *IEEE J. Select. Areas Commun.*, vol. SAC-2, no. 5, pp. 632–647, Sept. 1984.
[36] K. Feher, "Modems for emerging digital cellular mobile systems," *IEEE Trans. Vehic. Technol.*, vol. 40, no. 2, pp. 355–365, May 1991.
[37] M. Iida and K. Sakniwa, "Frequency selective compensation technology of digital 16-QAM for microcellular mobile radio communicatin systems," in *Proc. VTC'92,* Denver, CO, pp. 662–665.
[38] R. J. Castle and J. P. McGeehan, "A multilevel differential modem for narrowband fading channels," in *Proc. VTC'92,* Denver, CO, pp. 104–109.
[39] D. J. Purle, A. R. Nix, M. A. Beach, and J. P. McGeehan, "A preliminary performance evaluation of a linear frequency hopped modem," in *Proc. VTC'92,* Denver, CO, pp. 120–124.
[40] Y. Kamio and S. Sampei, "Pervormance of reduced complexity DFE using bidirectional equalizing in land mobile communications," in *Proc. VTC'92,* Denver, CO, pp. 372–376.
[41] T. Nagayasu, S. Sampei, and Y. Kamio, "Performance of 16-QAM with decision feedback equalizer using interpolation for land mobile communications," in *Proc. VTC'92,* Denver, CO, pp. 384–387.

[42] E. Malkamaki, "Binary and multilevel offset QAM, spectrum efficient modulation schemes for personal communications," in *Proc. VTC'92,* Denver, CO, pp. 325–378.
[43] Z. Wan and K. Feher, "Improved efficiency CDMA by constant envelope SQAM," in *Proc. VTC'92,* Denver, CO, pp. 51–55.
[44] H. Sasaoka, "Block coded 16-QAM/TDMA cellular radio system using cyclical slow frequency hopping," in *Proc. VTC'92,* Denver, CO, pp. 405–408.
[45] W. T. Webb, L. Hanzo, and R. Steele, "Bandwidth-efficient QAM schemes for Rayleigh-fading channels," *IEE Proc.,* vol. 138, no. 3, pp. 169–175, June 1991.
[46] T. Sunaga and S. Sampei, "Performance of multi-level QAM with post-detection maximal ratio combining space diversity for digital land-mobile radio communications," *IEEE Trans. Vehic. Technol.,* vol. 42, no. 3, pp. 294–302, Aug. 1993.
[47] P. K. Ho, J. K. Cavers, and J. L. Varaldi, "The effects of constellation density on trellis-coded modulation in fading channels," *IEEE Trans. Vehic. Technol.,* vol. 42, no. 3, pp. 318–326, Aug. 1993.
[48] J. K. Cavers and J. Varaldi, "Cochannel interference and pilot symbol assited modulation," *IEEE Trans. Vehic. Technol.,* vol. 42, no. 4, pp. 407–414, Nov. 1993.
[49] M. Frullone *et al.,* "Investigation on dynamic channel allocation strategies suitable for PRMA schemes," in *1993 IEEE Int. Symp. Cicuits Syst.,* Chicago, IL, May 1993, pp. 2216–2219.
[50] K. H. H. Wong and L. Hanzo, "Channel coding," in *Mobile Radio Communications,* R. Stelle, Ed. London: IEEE-Pentech, 1992, pp. 347–488.
[51] E. Biglieri and M. Luise, "Coded modulation and bandwidth-efficient transmission," in *Proc. Fifth Tirrenia Int. Wkshp.,* Sept. 8–12, 1991.
[52] *IEEE Communications Mag.,* Special Issue on Coded Modulation, vol. 29, no. 12, Dec. 1991.
[53] A. S. Wright and W. G. Durtler, "Experimental performance of an adaptive digital linearized power amplifier," *IEEE Trans. Vehic. Technol.,* vol. 41, no. 4, pp. 395–400, Nov. 1992.
[54] M. Faulkner and T. Mattson, "Spectral sensitivity of power amplifiers to quadrature modulator misalignment," *IEEE Trans. Vehic. Technol.,* vol. 41, no. 4, pp. 516–525, Nov. 1992.
[55] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Vehic. Technol.,* vol. 40, no. 4, pp. 686–693, Nov. 1991.
[56] A. Bateman and J. P. McGeehan, "Feedforward transparent tone in band for rapid fading protection in multipath fading," in *IEE Int. Conf. Commun.,* 1986, vol. 68, pp. 9–13.
[57] A. Bateman, "Feedforward transparent tone in band: Its implementation and applications," *IEEE Trans. Vehic. Technol.,* vol. 39, no. 3, pp. 235–243, Aug. 1990.
[58] J. K. Cavers, "The performance of phase locked transparent tone in band with symmetric phase detection," *IEEE Trans. Commun.,* vol. 39, pp. 1389–1399, Sept. 1991.
[59] L. Hanzo and J. Stefanov, "The Pan-European digital cellular mobile radio system—Known as GSM," in *Mobile Radio Commun.,* R. Stelle, Ed. London: Pentech, 1992, ch. 8, pp. 677–773.

Lajos Hanzo (M-91–SM'92) graduated with the degree in electronics in 1976 and the Ph.D. degree in 1983 from the Technical University of Budapest Hungary.

During his 19 year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the United Kingdom. Since 1986, he has been with the Department of Electronics and Computer Science, University of Southampton, U.K., and has been a Consultant to Multiple Access Communications Ltd., U.K.

Dr. Hanzo has co-authored two books on mobile radio communications, published more than 80 research papers, and was awarded a number of distinctions. He is a member of the IEE.



Jürgen Streit was born in Cologne, Germany in 1968. He received the Dipl.-Ing. degree in electronic engineering from the Aachen University of Technology in 1993.

Since 1992, he has been with the Department of Electronics and Computer Science, University of Southampton, U.K., working with the mobile multimedia research group. He is currently working towards the Ph.D. degree in image coding.