Dual-Mode Vector-Quantized Low-Rate Cordless Videophone Systems for Indoors and Outdoors Applications

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

***Abstract*— Dual-mode reconfigurable wireless videophone transceivers are proposed for noise-, rather than interference-limited indoors and outdoors applications and their video quality, bit rate, robustness, and complexity issues are analyzed. A suite of fixed, but arbitrarily programmable low-rate, perceptually weighted vector quantized (VQ) codecs with and without run-length compression (RLC) are contrived for quarter common intermediate format (QCIF) videophone sequences. The 11.36-kb/s Codec 1 is Bose–Chaudhuri–Hochquenghem (BCH) (127,71,9) coded to a rate of 20.32 kb/s and this arrangement is comparatively studied along with the 8-kb/s Codec 2 and BCH(127,50,13) scheme, which has the same 20.32-kb/s overall rate. The source-sensitivity matched Systems 1–6 characterized in Table IV were contrived to comparatively study the range of system design options. For example, using Codec 1 in System 1 and coherent pilot symbol assisted 16-level quadrature amplitude modulation (16-PSAQAM), an overall signaling rate of 9 kBd was yielded, if the noise-limited channel had a signal-to-noise ratio (SNR) in excess of about 22 dB in the vicinity of the basestation or in indoors scenarios. In contrast, over lower quality outdoors channels near the fringes of the cell, the more robust 4-QAM mode of operation had to be invoked, which required twice as many time slots to accommodate the resulting 18-kBd stream and hence, reduced the total number of users supported. The robustness of Systems 2–4, and 6 was increased using automatic repeat requests (ARQ), again, inevitably reducing the number of users supported, which was between 6 and 16. In a bandwidth of 200 kHz, similarly to the Pan-European GSM mobile radio system's speech channel, using Systems 1, 3, 4, or 5, for example, 16 and eight videophone users can be supported in the 16- and 4-QAM modes, respectively, while in dual-mode cells the number of users is between eight and 16. The basic system characteristics are highlighted in Table IV.**

***Index Terms*—Fixed-rate video coding, QAM video communications, vector-quantized video coding, wireless video telephony.**

I. Introduction

MOTIVATED by the rapid emergence of mobile video telephony as a major research area [1]–[11], in this treatise[1] we set out to contrive and study a range of dual-mode vector-quantized wireless videophone schemes, which can configure themselves as a more bandwidth-efficient 9-kBd 4-b/symbol transceiver for benign noise-limited indoor applications or as a more robust 18-kBaud scheme for outdoors scenarios. Such intelligent multimode transceivers [12] (IMT) or software radios [13] are currently intensively studied under the auspices of the European Community's Fourth Framework Programme on Advanced Communications Technologies and Services (ACTS) [14] with the motivation of ensuring an optimum combination of transceiver parameters under time-variant teletraffic and propagation conditions.

Focusing our attention on video communications issues, the CCITT standard H261 codec was designed for benign additive white Gaussian noise (AWGN) channels, but there have been a number of successful attempts to invoke it over mobile channels [5], [7]. The motion pictures expert group (MPEG) standardized another discrete cosine transformed (DCT) [25], [26] codec for higher delay distributive video applications, namely the MPEG2 codec, which, due to its innate delay and vulnerability against channel errors has a limited suitability for mobile videophony. The standardization of the new H.263 recommendation and the associated audio-video multimedia multiplexing scheme specified in the H.324 recommendation is also near to completion, although due to the vulnerability of the multiplexing scheme the audio and video bits are easily misinterpreted. Hence, the robustness of the H.263 and H.324 schemes must be enhanced for wireless applications. The ambitious MPEG4 initiative is likely to have a similar compression ratio to the H.263 codec for all five possible video resolutions spanning from the 128 $\times$ 96 pixel so-called subquarter common intermediate format (SQCIF) to the 16 $\times$ 1408 $\times$ 1152 high-definition television-like resolution. An impediment of the current MPEG4 proposals is that their bit-rate fluctuation versus time is typically about 50% of the nominal mean bit rate, unless specific bit packetization algorithms are invoked. This bit-rate fluctuation naturally can be smoothed by adaptive buffering, but only at the cost of increased delay, which impinges on its suitability for interactive videophony. In contrast to existing video codecs, a key feature of the proposed codecs is that they maintain a fixed bit rate and, hence, are suitable for videophony over existing and future mobile radio systems.

The pivotal issue in designing such videophone schemes is finding the most attractive video codec in terms of video quality, bit rate, implementational complexity, robustness against channel errors, etc. [1]. In object-based analysis by synthesis coding [2] attempts are made to characterize each moving

Manuscript received March 13, 1995; revised March 11, 1996. This work was supported by the EPSRC, U.K., under Contract GR/K74043.

J. Streit is with the Department of Electrical and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K. (e-mail: jss@ecs.soton.ac.uk).

L. Hanzo is with the Department of Electrical and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

Publisher Item Identifier S 0018-9545(97)03132-0.

[1]The exposition of this treatise is augmented by a software demonstration package downloadable from the www site http://www-mobile.ecs.soton.ac.uk portraying a range of videophone sequences encoded at the rates of 8, 9.6, 11.36, and 13 kb/s using various coding techniques.

object by their shape, texture, and motion characteristics, which can be combined with more conventional methods in order to further reduce the bit rate. In knowledge-based or object-based coding [3] *a priori* knowledge as regards to the shape of the object to be encoded can be exploited in order to represent for example a person's head by means of its so-called wire-frame model. The motion of an object, such as the head can be classified as global and local, corresponding to head movements and various facial expression, respectively. These translations can be represented by fractal transformations of previously coded so-called range blocks. Analysis by synthesis coding [4] typically attempts to synthesize the image to be encoded a number of times in order to arrive at the subjectively most attractive quality/bit-rate tradeoff. This inevitably increases the complexity and raises the question of finding an appropriate objective quality measure, which exhibits a high correlation with the subjective quality.

Apart from DCT-based codecs [9], [25], subband coded [11], [15]–[17] and fractal-based [8] schemes have also been successfully employed to reduce the bit rate, although the latter exhibited a high implementational complexity and more modest compression ratio. Motivated by the flexibility of subband codecs in terms of controlling the allocation of the quantification distortion to the various spatial frequency bands at the early stages of our videophony research program we designed a 55-kbit/s CIF subband codec, which was embedded in a 22 kBaud single-user 16-QAM transceiver [11]. Following this experience, our work evolved further by recognizing some of the limitations of the subband codec in terms of bit-rate reduction and motion activity tracking. The 55-kbit/s average rate was achieved with the aid of run-length coding, which is known to exhibit extreme sensitivity against channel errors, since a single error is sufficient to corrupt an entire video frame. In order to smooth the bit-rate fluctuation, adaptive-feedback buffering was used, which increased the codec's delay. Since over wireless channels the typical bit error rate (BER) is high, our conclusion was that below a threshold signal-to-noise ratio (SNR) the system's performance broke down. As regards to motion activity tracking, we understood that in head-and-shoulders videophony there is typically a high proportion of low-activity background segments, which are amenable to higher compression ratios.

Based on the above experience, the above mentioned activity tracking was exploited by defining motion active and DCT-active blocks in [9], which led to a very high bandwidth efficiency, while maintaining good video quality at 8.52 and 11.36 kbits/s and 10 frames/s. Although the bit rate was constant, the 8.52-kb/s codec achieved this extreme low rate at the price of higher vulnerability due to variable-length coding, while the 11.36-kb/s codec had a higher robustness against channel errors. A common feature of both schemes was a slight blurring due to the bandwidth-limiting nature of the DCT-masking. In a further attempt to address this problem and to document the performance of potential benchmarkers under identical circumstances we also studied an 11.36-kb/s quadtree (QT) codec [10]. We will report on the performance of the above schemes after considering our best candidate codec, namely the vector-quantized scheme.

Explicitly, this contribution is aimed at mitigating this low-pass effect, improving the coding efficiency and robustness, while comparatively studying a range of vector-quantized constant-rate benchmarker videophone codecs, whose encoded stream can be transmitted over conventional mobile radio speech links, such as for example the European DECT scheme, the American IS-54 or IS-95 systems, the Japanese Handyphone arrangement known as PHP or the Pan-European GSM system [57]. In this study we opted for an identical transmission system to that of [9], which can be accommodated in a 200-kHz GSM channel, allowing us to relate our findings to a benchmarker system. The 11.36-kb/s codec refrained from using efficient, but vulnerable variable-length coding, while the 8-kb/s codec employed variable-length coding. The vector-quantized video stream was transmitted using a source-sensitivity matched reconfigurable transceiver similar to that of [9], which has a robust, but less bandwidth efficient mode of operation or can halve its bandwidth requirement, if better channel conditions are maintained. Again, similarly to [9], the 11.36/8-kb/s video source rate was chosen in order to be able to accommodate 16 or eight videophone users in a bandwidth of 200 kHz, as in case of the full- and half-rate GSM speech systems [57] and to directly compare the codec and system performances.

Joint coding and modulation, known as trellis-coded modulation (TCM) or block–coded modulation (BCM) [49], [50] exhibits typically higher coding gain than that achieved by them in isolation. However, in case of practical source codecs, which retain some residual redundancy and exhibit unequal sensitivity against channel errors source-sensitivity matched channel coding [11], [19], [21] must be invoked. References [19] and [23] proposed source-sensitivity matched joint source/channel coding and modulation schemes and in this treatise we will follow a similar design philosophy, where the video encoded stream is mapped in two different protection classes, namely class one and two. Then it is binary Bose–Chaudhuri–Hochquenghem (BCH) coded [48], before they are multiplexed to a time division multiple access (TDMA) slot and transmitted using quadrature amplitude modulation (QAM) [24]. We note that only one TDMA slot will be reserved for benign channel conditions, when 16 QAM is used, while two slots are required for users roaming in more hostile environments near the cell fringes, who are communicating using 4 QAM.

Section II highlights the design of the 11.36/8-kb/s video codecs, which is followed by their bit sensitivity analysis, while Section IV justifies the choice of the vector quantized codec on the basis of comparing it with other identical-rate benchmarkers. Section V discusses modulation, forward error correction (FEC) coding and other transmission issues, including the design of the source-sensitivity matched transceiver. The system's performance is analyzed in Section VI, while Section VII draws some conclusions.

II. Vector-Quantized Video Codecs

A. The Codebook Design

Let us commence our detailed discourse considering the VQ video codec. The codebook design is a crucial issue for every

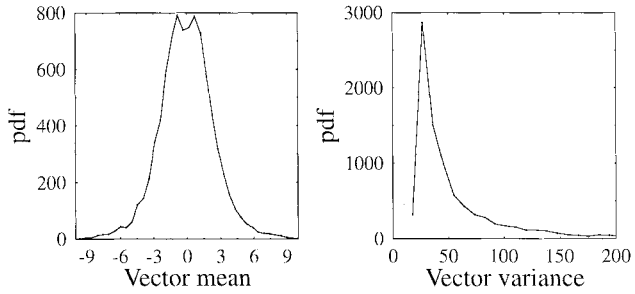

Fig. 1. Statistical properties of the training sequence.

VQ codec. Unfortunately, there is no practical algorithm that leads to the optimal codebook C for a given training sequence S. Practical algorithms are usually based on a two-stage approach. First, an initial codebook is derived, which is then improved in a second stage using the so-called generalized Max–Lloyd algorithm [26].

The choice of the VQ training sequence is also critical, since it should contain data, which is characteristic of a wide range of inputs. In our experiments, we calculated the displaced frame difference (DFD) signal for various input sequences, while only allowing for 35 active motion vectors out of the 396 8×8 blocks of a quarter common intermediate format (QCIF) frame to be generated. This presumes that the VQ will be used in very low bit-rate image coding, where an active/passive block classification is necessary. The DFD energy was analyzed and the 35 blocks containing the highest energy were copied into the training sequence. The size of the training sequence was chosen to be at least ten times the size of the codebook in order to guarantee a sufficient statistical independence between the codebook and the training sequence. The statistical properties of the training sequence are portrayed in terms of the probability density function (PDF) of its mean and variance in Fig. 1. The mean and variance are highly peaked between $[-2 \cdots 2]$ and $[20 \cdots 40]$, respectively. Note in the figure that the narrow, but highly peaked PDF of the DFD mean underlines the efficiency of MC, since in the vast majority of cases its mean is close to zero and its variance implies a low DFD energy.

Initially the codebook was designed using the so-called "pruning" method [58], where a small codebook is derived form a large codebook or from a training sequence. The algorithm is initialized with a single-vector codebook. Then the vectors of the training sequence are compared to those in the codebook using the component-wise squared and summed distance as a similarity criterion. If this distance term exceeds a given threshold, the vector is classified as "new" and accordingly copied into the codebook. Hence, the threshold controls the size of the codebook, which contains all "new" vectors. In our case this method did not result in an adequate codebook, since the performance of a large codebook of 512 entries, which had an unacceptably high-complexity, was found poor.

In a second attempt, we made use of the "pairwise nearest neighbor" (PNN) algorithm [58]. This approach shrinks a given codebook step by step until a desired size is achieved. Initially each vector is assigned to a separate cluster. Then for two candidate clusters the distortion penalty incurred by merging these two clusters is determined. This is carried out for all possible pairs of clusters and finally the pair with the minimum distortion penalty is merged to a single cluster. This process is continued until the codebook is shrunk to the desired size. The algorithm's complexity increases drastically with the codebook size and this technique became impractical for our large training set. Hence, we simplified this algorithm by limiting the number of tentative cluster combinations as follows. Instead of attempting to merge each possible pair of clusters, where the number of combinations exhibits a quadratic expansion with increasing codebook size, in our approach a single cluster is preselected and combined with all the others.

Following this technique the codebook length was initially reduced and then the full PNN algorithm was invoked in order to create a range of codebooks having sizes between four and 1024. In the second step of the codebook generation we used the generalized Max–Lloyd algorithm to enhance our initial codebooks, which resulted in a limited reduction of the average codebook distortion. This underlines the fact that our initial codebooks used by the PNN algorithm were adequate and the suboptimum two-stage approach caused only a negligible performance loss. As an example, our 16-entry codebook constituted by 8×8 vectors is depicted in Fig. 2, where the vector components were shifted by 127 and multiplied by 30 in order to visually emphasize their differences.

B. VQ Codec Outline

The schematic of the proposed image codec is displayed in Fig. 3. As mentioned, the video codec operates using 176×144 pixels QCIF images scanned at 10 frames/s. A constant bit-rate source codec was required, in order to be able accommodate its bit stream in conventional mobile radio speech channels, such as for example that of the GSM system [57]. Hence, we fixed both the number of 8×8 blocks to be motion-compensated and those to be vector quantitative to a value depending on the required bit rate.

1) Intraframe Mode: Since the first frame used in motion compensation is unknown to the decoder, the 4-b encoded block averages are transmitted first in the intraframe mode, which are also used by the local decoder. In order to mitigate the effects of transmission errors partial forced update (PFU) of both the encoder's and decoder's reconstructed buffers is also carried out on a similar basis, using the 4-b encoded block averages of a certain number of blocks. The number of the PFU blocks per frame is also bit-rate dependent and it is automatically determined by our programmable codec. For example, in our 11.36-kb/s prototype 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.

Gain Controlled Motion Compensation: Initially the motion compensation (MC) scheme determines a motion vector

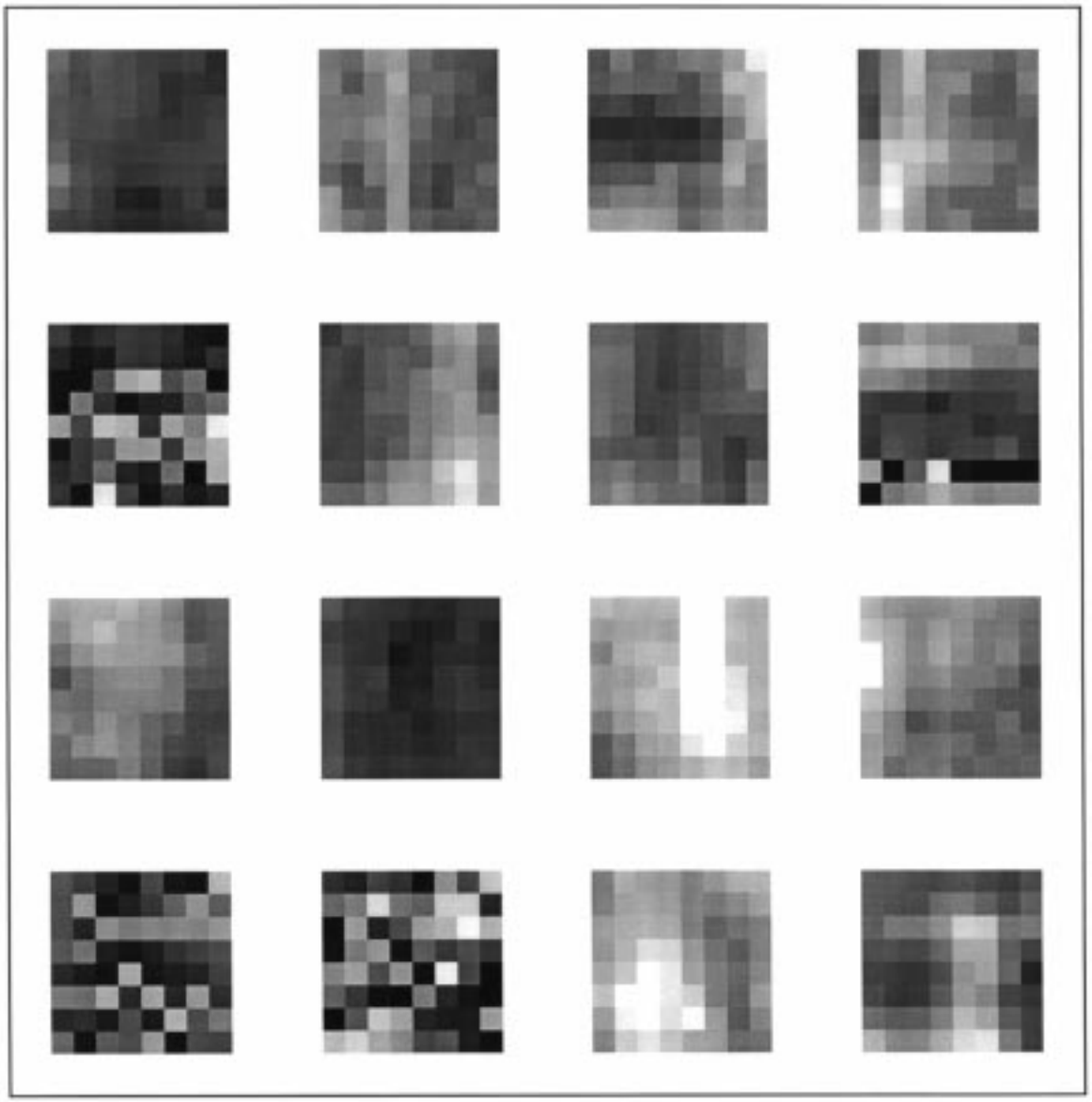

Fig. 2. Enhanced sample codebook with 16 8 × 8 vectors.

(MV) for each of the 8 × 8 blocks. The MC search window is fixed to 4 × 4 pels around the center of each block. In our cost-gain controlled approach the codec tentatively determines the achievable gain of the compensation in terms of DFD energy reduction. The gains in the most important eye and mouth region are weighted by a factor of two. Then the codec identifies the required number of blocks resulting in the highest scaled DFD gain, and motion compensation is applied only to these blocks. For the remaining motion-passive blocks frame differencing is employed. Observe in Fig. 3 that an optional "table compression" algorithm can be employed in order to further improve the codec's bandwidth efficiency, which will be described during our further discourse.

C. Displaced Frame Difference Coding

1) Potential Coding Schemes: The DFD signal is often encoded in the spatial frequency domain either after subband splitting [11] or after DCT [9]. Although both of these techniques lead to attractive codec implementations, they tend to loose high-frequency fine resolution due to the extreme paucity of bits. An alternative is to encode the DFD signal in the time-domain using QT coding [10] by optimizing the bitallocation for a whole video frame, but unfortunately the QT code is more vulnerable to channel errors than the DCT coefficients. In this contribution we expanded our former work [9]–[11] to consider the time-domain representation of the DFD signal in contrast to the above benchmarkers. Gersho and Gray considered a variety of vector quantized schemes [58] for a range of applications. In this treatise we comparatively study a wide variety of VQ-based codecs for constant-rate wireless videophony over existing and future intelligent wireless systems at a rate of around 10 kb/s.

a) Vector quantification: The DFD signal is then passed on to the VQ scheme. A second gain-and bit-rate-controlled algorithm decides for which of the blocks the VQ provides a high coding gain, i.e., which of the blocks are deemed

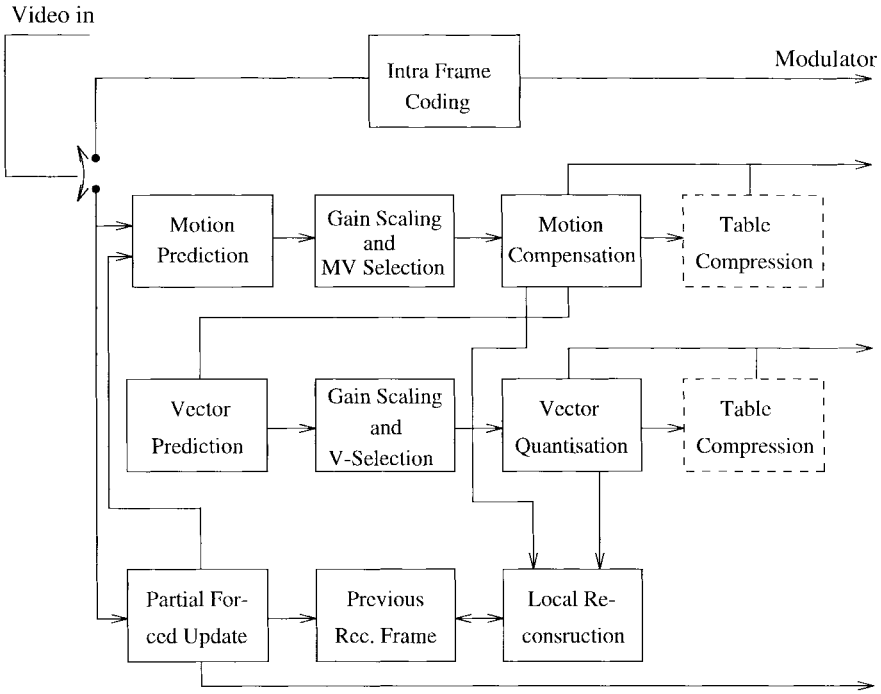

Fig. 3. Basic schematic of the VQ.

to be active and these indexes are then passed on to the receiver. Explicitly, the active VQ block indexes are stored and transmitted after optional activity table compression, as seen in Fig. 3, which will be described at a later stage. The encoded frame is locally reconstructed and fed back to the "previous reconstructed frame buffer" in order to assist in the next stage of motion compensation.

For every block the VQ searches through the entire codebook and compares the match by summing the squared differences for each pixel within the two blocks under consideration. Therefore 191 floating point operations (Flops) are necessary to perform one block match. In case of a 128-entry codebook for an image frame, a total computational complexity of 396 blocks $\times$ 128 codebook entries $\times$ 191 Flop $=$ 9.6 Mflops is incurred. Similarly, a 256-entry codebook implies a complexity of 19.3 Mflops.

In order to explore the range of design tradeoffs we tested several VQ-schemes in which we incorporated adaptive codebooks, classified VQ [59] and scaling VQ. However, before commenting on these endeavors let us first report our findings as regards to the practical range of codebook sizes. As mentioned, we opted for a fixed, but programmable bit-rate scheme and varied the codebook size between four and 1024. As Fig. 4 reveals, best peak signal-to-noise ratio (PSNR) performance was achieved using codebook sizes in the range of 128–512. For these investigations we used a locally recorded high-activity head-and-shoulders videophone sequence, which we refer to as the "lab sequence," since the well-known low-activity Miss America (MA) sequence was inadequate for evaluating the VQ performance. Observe that the 256-entry codebook results in the best PSNR performance. This corresponds to a VQ data rate of 0.14 b per pixel or 8 b per 8 $\times$ 8 block. However, a codebook size of 128 is preferable, as it halves the codec's complexity without significantly reducing the quality.

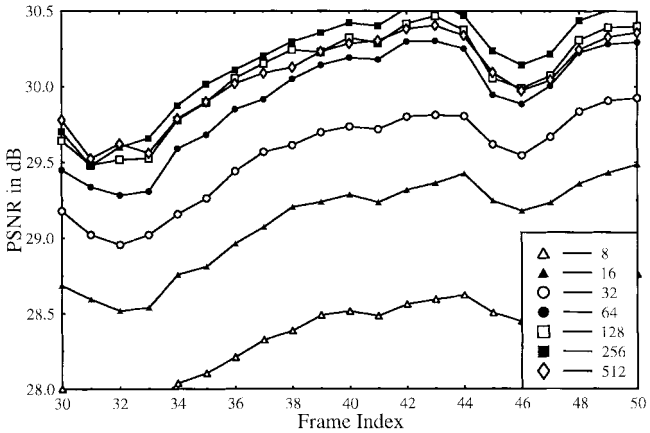

Fig. 4. PSNR performance for the "lab sequence" and various codebook sizes at a constant bit rate.

2) Mean and Shape Gain VQ (MSVQ): In an attempt to increase the VQ quality without increasing the codebook size, we also experimented with a mean- and variance-normalized VQ scheme. The approach is based on normalizing the incoming blocks to zero mean and unity variance in order to create a larger variety of codebook entries without a significant increase of the complexity. As seen in Fig. 5, the MSVQ codec first removes the quantized mean of a given block and then it normalizes the block variance to near-unity variance by its quantized variance. These two operations require $3n^2 - 2$ Flop per block, while virtually expanding the codebook size by the product of the number of reconstruction levels of the mean and variance quantizers. Hence, a 2-b quantizer for both the mean and variance would virtually expand the codebook to 16 times its original size. The appropriate reconstruction levels were obtained by employing a Max–Lloyd quantizer.

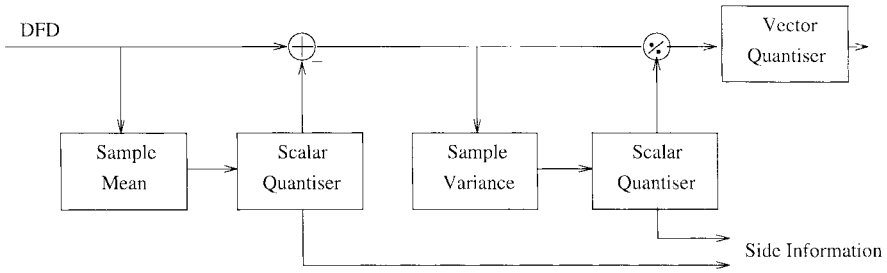

Fig. 5. Block diagram for a MSVQ.

TABLE I
RESULTS FOR THE MSVQ AT 8 kb/s FOR A 64- AND 256-ENTRY CODEBOOK

Mean / Variance quantiser levels	1	2	4	8	16	32
Av. PSNR 64-entry MSVQ	28.22	28.35	27.09	26.61	26.64	26.62
Av. PSNR 256-entry MSVQ	28.6	28.16	27.68	27.46	27.56	27.49

We assessed the performance of the MSVQ using codebook sizes of 64 and 256 and quantizers, which ranged from 1–32 reconstructions levels, while maintaining a constant bit rate. Using a one-level MSVQ scheme corresponds to the baseline VQ with no scaling at all. As demonstrated by Table I, the MSVQ codec's performance was slightly degraded in comparison to the baseline codec, except for the 64-entry codebook with a two-level scaler, where a minor improvement was achieved. It is interesting to note, however, that this scheme has an almost identical performance to that of the 256-entry one-level scaler arrangement, while retaining a lower complexity. From these endeavors, we concluded that our codebooks obtained by the modified PNN algorithm are close to the optimum codebooks and therefore the virtual codebook expansion cannot improve the performance significantly. The advantage of the 64-entry MSVQ scheme was that we maintained the performance of the 256-entry VQ arrangement with a complexity reduction of around 70%.

3) Adaptive VQ (AVQ): In adaptive VQ, the codebook is typically updated with vectors that occur frequently, but are not represented in the codebook. The increased flexibility of the method is unfortunately often associated with an increased vulnerability against channel errors. Our adaptive approach is depicted in Fig. 6. The codec is based on two codebooks implemented as first-in-first-out (FIFO) pipelines, one of which is called the active codebook and the other is referred to as passive. The VQ itself has only access to the active codebook and can choose any of the vectors contained in that codebook. Vectors selected by the VQ are then written back to the beginning of the pipeline. After each encoding step, a certain number of vectors are taken from the end of the pipeline and fed-back to the passive codebook. The same number of vectors are taken from the end of the pipeline in the passive codebook and entered at the beginning of the active codebook. This method forces a continuous fluctuation between the two codebooks, but allows frequently used vectors to remain "active." The principle is as vulnerable to channel errors as the fully adaptive system, but a forced realignment at regular intervals is possible, as the total pool of vectors is finite and known to both the encoder and decoder.

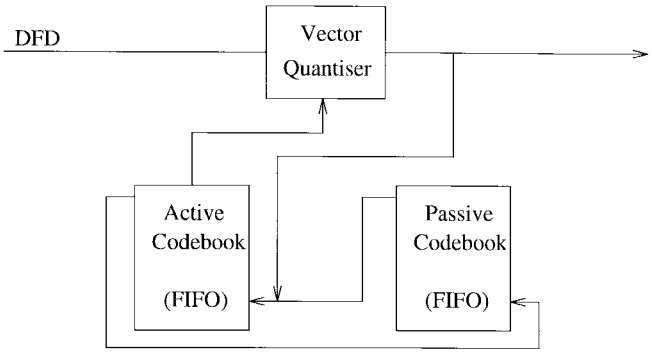

Fig. 6. Block diagram for the codebook adaptive VQ.

Our performance evaluation experiments were based again on the "lab sequence" and the 256-sized codebook, which was split into an active codebook of 64 vectors and a remaining passive codebook of 192 vectors. The variable in this scenario was the number of updated vectors after each encoding step, which was fixed to four, eight, 16, and 32 vectors per step. The corresponding PSNR versus frame index results are depicted in Fig. 7 along with the performance of the standard 256-entry VQ. As with all other experiments, the comparison was based on a constant bit rate. Using this concept we were able to increase the performance of the 64-entry VQ to a level similar to that of the standard 256-entry VQ, when applying a high vector exchange rate of 32 vectors per step. The lower vector exchange rates still improved the performance, when compared to a standard 64-entry VQ, but the PSNR gain was more modest, amounting to about 0.2 dB. Hence, adaptive VQ with a high vector replenishment rate of 32 vectors is also an attractive candidate to reduce the complexity of the VQ at a concomitant increased vulnerability to channel errors.

4) Classified VQ (CVQ): Classified VQ [59] has been introduced as a means of reducing the VQ complexity and to preserve the edge integrity of intraframe coded images. As the block diagram in Fig. 8 outlines, a CVQ codec is a combination of a classifier and an ordinary VQ using a series of codebooks. The incoming block is classified into

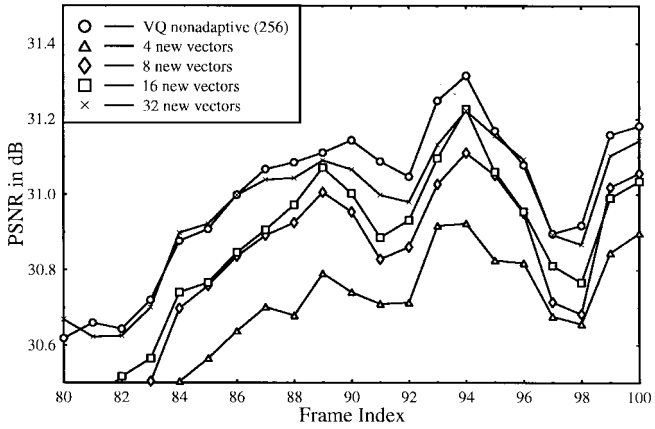

Fig. 7. VQ performance for the "lab sequence" with and without adaptive codebooks at 10 kb/s.

one of n classes and, hence, only the corresponding reduced-size codebook c_n is searched, thereby reducing the search complexity. In [59] the image blocks to be vector quantized were classified as shade, midrange, mixed and edge blocks, with the latter ones further sorted according to their gradient across the block.

In our case the VQ's input was the DFD signal, for which the above classification algorithm is not applicable. In our approach a smaller codebook, derived from the same training sequence as the unclassified codebook C, can be viewed as a set of centroids for the codebook C. Hence, codebook C can be split into n codebooks c_n by assigning each vector in C to one of the centroids. The encoding procedure then consists of a two-stage VQ process and may be seen as a tree-structured VQ (TSVQ) [59]. The classifier is a VQ in its own right, using a codebook filled with the n centroids. Once the closest centroid has been found, the subcodebook containing the associated vectors is selected for the second VQ step. We carried out various experiments, based on the optimum 256-entry codebook, with classifiers containing $n = 16$, 64, and 128 centroids, as revealed in Fig. 9. The PSNR performance loss due to using this suboptimum two-stage approach is less than 0.3 dB. This is amazing, as the number of block comparisons was reduced from 256 for the standard VQ to about 64 for the CVQ. The complexity reduction cannot be exactly quantified, as the codebooks c_n do not necessarily contain N/n vectors.

5) Algorithmic Complexity: The mean squared error (MSE) of the pattern matching process is defined as

$$d_{\text{mse}} = \sum_{i=1}^{b} \sum_{j=1}^{b} [x(i, j) - y(i, j)]^2 \tag{1}$$

where the block dimensions are given by $b \times b$ pixels, $\vec{x}$ is the two-dimensional input vector and $\vec{y}$ represents the codebook vector. Equation (1) can be expanded as

$$d_{\text{mse}} = \sum_{i=1}^{b} \sum_{j=1}^{b} x(i, j)^2$$

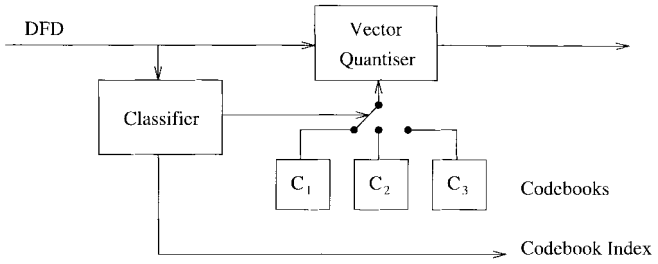

Fig. 8. Block diagram of a classified VQ.

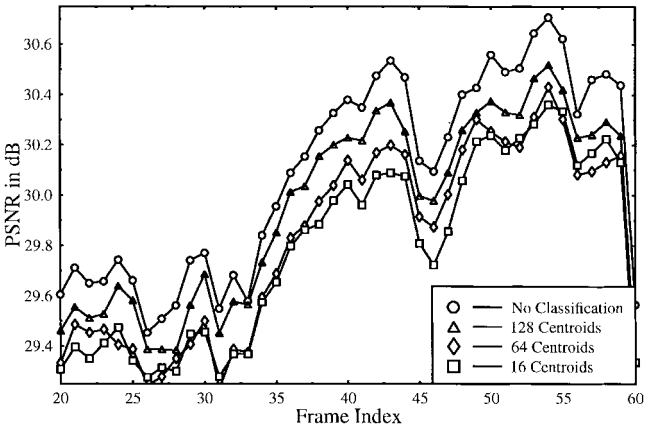

Fig. 9. PSNR versus frame index performance of classified VQ codecs when using the "lab sequence."

$$+ \sum_{i=1}^{b} \sum_{j=1}^{b} y(i, j)^2 - 2 \sum_{i=1}^{b} \sum_{j=1}^{b} x(i, j) y(i, j). \tag{2}$$

The first term of (2) depends only on the input vector and, hence, it is constant for every comparison, not affecting the nearest neighbor decision [60]. The second term depends only on the codebook entries and therefore it can be determined prior to the encoding process. The last term of (2) depends on both the incoming vector and the codebook entries. The evaluation of (2) implies a complexity of $2b^2+1$ flops, whereas that of (1) is associated with $3b^2 - 1$ flops, corresponding to a complexity reduction of about 30%. When using a CVQ codec invoking the above technique, the overall codec complexity can be reduced to around one Mflop per frame or 10 Mflops at a frame rate of 10 frame/s.

D. Bit Allocation Strategy

Finally, in order to be able to accommodate the video encoded bit stream in a conventional mobile radio speech channel we contrived two VQ-based prototype codecs. Codec 1 generated a video source rate of 11.36 kb/s or 1136 b/frame and dispensed with the optional motion- and VQ-activity table compression shown in dashed lines in Fig. 3. This ensured a higher innate robustness than in case of the run-length compressed 8-kb/s Codec 2. The bit allocation scheme of both codecs is shown in Table II, which will be detailed during our following discourse.

TABLE II
BIT-ALLOCATION TABLE

Codec	FAW	PFU	MV Index	MV	VQ Index	VQ	Padding	Total
Codec 1	22	22×4	38×9	38×4	31×9	31×8	5	1136
Codec 2	22	22×4	-	< 500 (VLC)	-	VLC	VLC	800

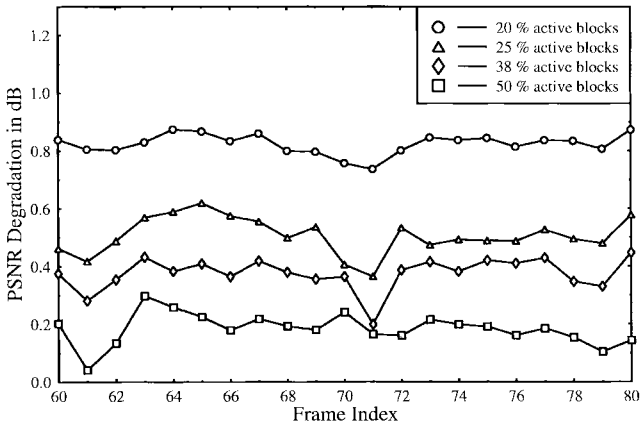

Fig. 10. PSNR degradation versus frame index performance of active/passive classified VQ codecs.

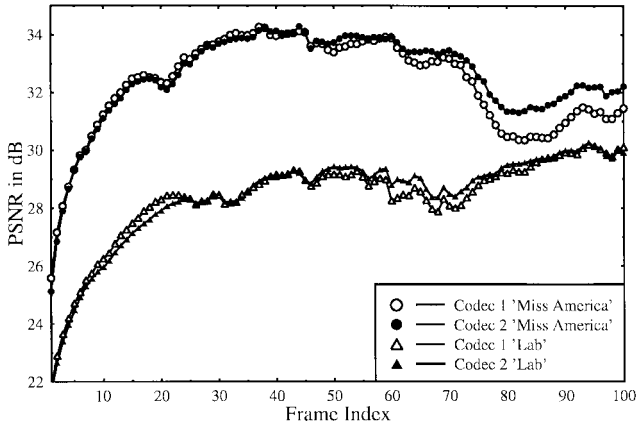

Fig. 11. PSNR versus frame index performance of the 11.36-kb/s Codec 1 and the 8-kb/s Codec 2 for the "Miss America" sequence.

Both codecs are based on the classified VQ codec as it offers the best compromise in terms of quality, robustness, and complexity. We used a centroid codebook size of 16 and an overall codebook size of 256, which lead to an overall codec complexity of around 15 Mflops when using the following active/passive classification. Initially the frame difference signal is computed for all 396 8 × 8 blocks and only a certain fraction of the high-energy frame difference blocks between 20–50% is deemed active on this basis, which then undergo further processing as detailed below. This initial classification results in a significant complexity reduction at the price of a minor PSNR performance degradation, which is quantified by Fig. 10 for active block fractions of 20–50%, when used in the 8-kb/s prototype Codec 2 to be described later in this section.

Codec 1 did not take advantage of the optional runlength encoding for the active/passive tables for the sake of increased robustness. The codec output contains the frame alignment word (FAW), the PFU, the MV's, and the VQ data. The 22-b FAW is required to support the video decoder's operation in order to regain synchronous operation after loss of frame synchronization. The partial intraframe update refreshes 22 8 × 8-sized blocks out of the 396 blocks per frame. Therefore every 18 frames or 1.8 s the update refreshes the same blocks. This periodicity is signaled to the decoder by transmitting the inverted FAW. In order to comply with the bit-rate requirement of 11.36 kb/s, a total of 38 MV's are stored using 13 b each, where 9 b are required to identify one of the 396 indexes using the enumerative method and 4 b/s for encoding the 16 possible combinations of the X and Y displacements. Each of the 31 active 8 × 8 VQ blocks uses a total of 17 bits, again nine for the block index and eight for the codebook index. The total number of bits becomes $22 + 22 \cdot 4 + 38 \cdot 13 + 31 \cdot 17 + 5 = 1136$, where five 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 PSNR performance is portrayed in Fig. 11 for the well-known MA sequence and for a high-activity sequence referred to as the "lab sequence."[2] For "MA" an average PSNR of about 33 dB was maintained, which was associated with pleasant videophone quality.

Again, the bit allocation scheme is summarized in Table II and the complexity of this codec is about 15 Mflops when applying the previously introduced active/passive block classification prior to further processing. Recall that this classification process determines the energy of every block prior to the MC or VQ encoding steps and only the 100 blocks associated with the largest energy values are subjected tentatively to full MC and VQ.

In our further discourse we will refer to the above scheme as Codec 1. After proposing a lower bit rate, but more error sensitive arrangement, Codec 2, we will address their bit sensitivity issues and analyze their advantages and disadvantages.

Codec 2 capitalizes on the fact that the tables of active MV and active VQ-block indexes retain some redundancy, which can be removed by variable-length coding (VLC), as suggested by the dashed blocks of Fig. 3. In order to be able to transmit all block averages in the intraframe mode with a 4-b resolution, as in Codec 1, while not exceeding the 800-b/frame budget we fixed the intra frame block size to 12 × 12. Again, the full bit allocation scheme is portrayed in Fig. 2.

However, in the MC we retained the block-size of 8 × 8 and the search window size of 4 × 4 around the center of each block. This method of classifying the blocks as motion-active and motion-passive results in an active/passive table, which consists of a 1-b flag for each of the 396 blocks, marking it as passive or active. These tables are compressed using the elements of a two stage QT as follows.

[2]The MA sequence encoded at various bit rates can be viewed under the WWW address http://www-mobile.ecs.soton.ac.uk.

TABLE III
AVERAGE PSNR PERFORMANCE OF CODECS 1 AND 2 FOR THE "MISS AMERICA" AND "LAB" SEQUENCES

Codec / Sequence	'Miss America'	'Lab'
Codec 1 (11.36 kb/s)	32.56 dB	28.56 dB
Codec 2 (8 kb/s)	32.62 dB	28.60 dB

First, the 396-entry activity table containing the binary flags is grouped in 2×2 blocks and a 4-b symbol is allocated to those blocks that contain at least one active flag. These 4-b 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-activity tables as well as to the active MV's exceeds half of the total number of bits available for the encoding of each frame, which is 800 in case of the 8-kb/s Codec 2, then some of the blocks satisfying the initial motion-active criterion will be relegated to the motion-passive class.

The VQ blocks are handled using a similar procedure. Depending on the required number of bits per frame and the free buffer space, an appropriate number of active VQ blocks is chosen and the corresponding compressed tables are determined. If the total number of bits "overspills," or if there are too many bits left unused, a different number of active blocks is estimated and new tables are determined. The disadvantage of such a high compression scheme is the strongly increased vulnerability of the transmission burst. If a transmission error corrupts one of the RL encoded tables, it is likely that a codeword of a different length is generated and decoding becomes impossible. This will force the decoder to drop the entire frame.

The average PSNR performances of Codecs 1 and 2 are summarized for the MA sequence and for the higher activity "lab sequence" in Table III.

III. BIT SENSITIVITY

In order to apply source-sensitivity matched protection the video bits were subjected to sensitivity analysis. In [11] we have consistently corrupted a single bit of a video coded frame and observed the image 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 [18], 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.

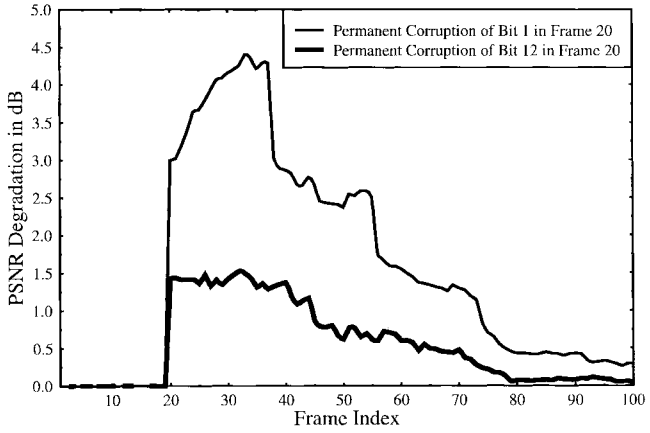

Fig. 12. PSNR degradation profile for Bits 1 and 12 of the in Frame 21.

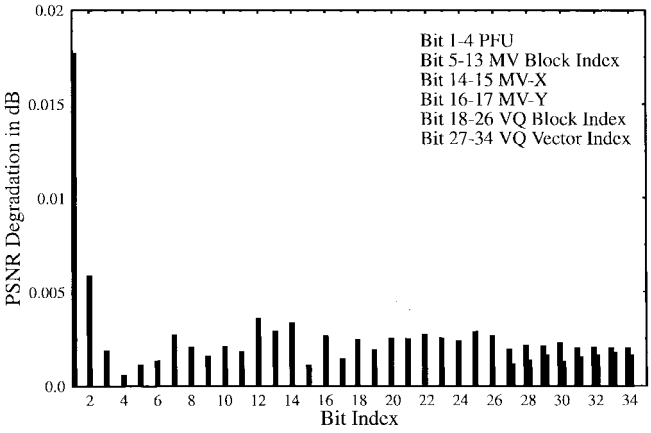

Fig. 13. Integrated PSNR bit sensitivities of Codec 1.

As an example, Fig. 12 depicts the PSNR degradation profile in case of corrupting all "No 1" bits, the most significant bit (MSB) of the PFU and all "No 12" Bits, one of the address bits of the MV, in each block of frame 21. In the first case, the MSB's 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. 12. The impact of the corrupted MV is randomly distributed across the frame and it is also 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 possible occurrences of the corresponding bit errors. These results are shown in Fig. 13 for the 13 MV bits and 17 VQ bits of an 8×8 block, as well as for the four partial forced update bits. The VQ-index sensitivity depends on the codebook-contents as well as on the codebook order. We decreased the sensitivity of the VQ indexes by reordering the codebook using the simulated annealing algorithm.

Accordingly, starting from a random codebook order, we invoked a recursive algorithm selecting two codebook entries at random and evaluating the potential gain in terms of codebook robustness, if the candidate entries were swapped. At the

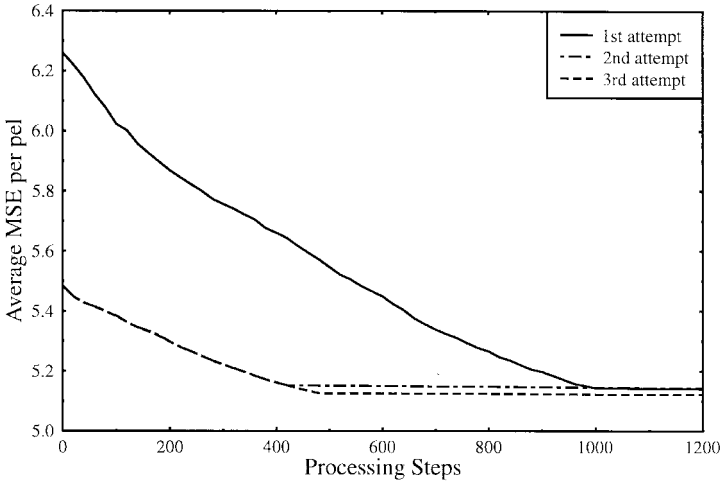

Fig. 14. Improvement of the codec robustness due to reordering the codebook entries.

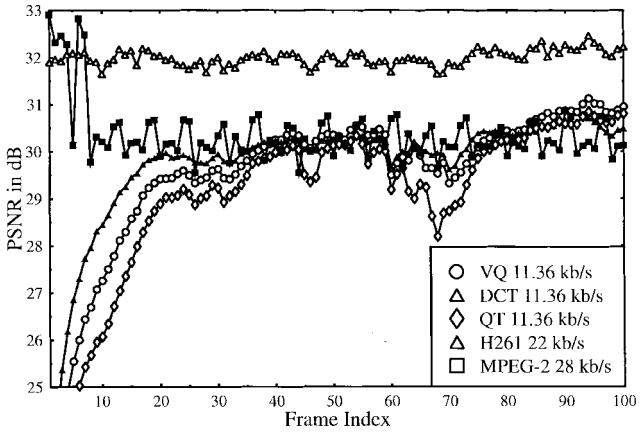

Fig. 15. PSNR versus frame index performance of the proposed adaptive codecs and two standard codecs.

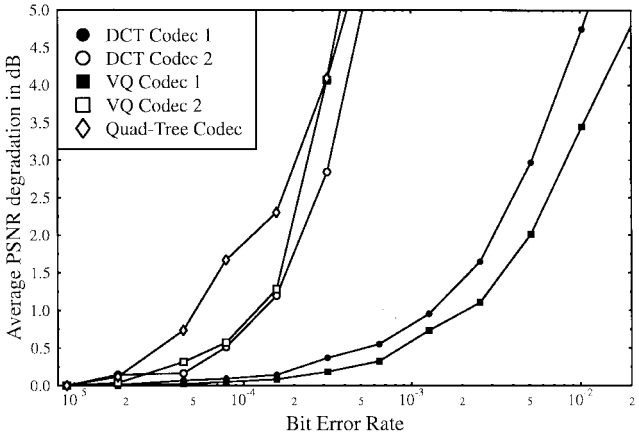

Fig. 16. PSNR degradation versus BER for the proposed codecs.

commencement of our codebook rearranging algorithm we allowed codebook entry swaps even if they decreased the overall codebook robustness. This assisted the algorithm in emerging from a local minimum. Throughout the optimization process we decreased the probability of such "disadvantageous" swaps and allowed the algorithm to settle to a minimum. In Fig. 14 we the demonstrated the increase of the CB robustness for three "simulated annealing" attempts starting from different initial assignments. As seen in the figure, all attempts resulted in a robustness increase of up to 20%. The impoved robustness is revealed by comparing the two sets of PSNR degradation bars associated with the corresponding bit positions 27–34 in Fig. 13. Having characterized a range of VQ-based video codecs let us now relate their performance to that of the DCT- and QT-based benchmarker of [9], [10] as well as to the CCITT H.261 and the MPEG2 standard codecs.

IV. Comparison of Benchmark Codecs

Let us now compare the 11.36-kb/s QT scheme of [10], the DCT codec arrangements of [9] and the currently proposed VQ codecs in quality and error resilience terms. As a comparative basis we will also refer to two widely used standard codecs, namely the MPEG-2 and H261 codecs, which have a fluctuating bit rate and hence, are unsuited for videophony over fixed-rate mobile channels. In our PSNR comparison shown in Fig. 15 we adjusted the parameters of the H261 and MPEG-2 codecs to provide a similar video quality associated with a similar average PSNR performance to our fixed-rate wireless video codecs. The corresponding bit rate was about twice as high for the two standard codecs, exhibiting a random fluctuation for the H261 codec. The MPEG codec exhibits three different characteristic bit rates, corresponding to the so-called intra (I), bidirectional (B), and predicted (P) frames in decreasing order from around 8000 b/frame, to about 1800 and 1300, respectively.

Scrutinizing the above five codecs in Figs. 15 and 16 shows that our codecs achieve a similar PSNR performance to the MPEG-2 and the H.261 codecs at less than half the bit rate, although our fixed-rate DCT and VQ codecs reach their steady-state video quality due to the fixed bit-rate limitation only after some 2 s. Furthermore, the delay of our codecs and that of the H-261 codec is 100 ms, while that of the MPEG-2 codec extends to several frames due to the P-frames. The vector quantized codec constitutes the best compromise in terms of quality, compression ratio and computational demand, closely followed by the DCT codecs [9].

Specifically, the proposed VQ codec achieved a similar DFD representation quality and video quality to the DCT codec of [9] by allocating an 8-b codebook index to all VQ-active blocks, whereas the DCT codec used a total of 12 b per active DFD block, yet introducing slightly more low-pass effects. Therefore the proposed VQ codec was able to support 38, rather than 30 motion-active and DFD-active blocks, as we have seen in [9] for the DCT codec, providing an overall improved activity tracking and video quality.

In robustness terms we concluded that best error resilience was achieved by the currently proposed VQ Codec 1, followed by the DCT Codec 1 of [9], while the corresponding Codec 2 schemes and the QT codec of [10] were more error sensitive. The VQ codec's lower error sensitivity in comparison to that of the DCT codec, was partly a consequence of the reduced codebook index sensitivity due to using simulated annealing. This exhibited itself also in Fig. 13. In contrast, in the DCT codec [9] the corruption of the low-frequency DCT coefficients introduced large video degradations. Due to its variable-length

nature the QT codec is more prone to errors than the VQ and DCT Codecs 1 and 2.[3] Having identified the best VQ video codec let us now turn to transmission issues.

V. Transmission Issues

A. Choice of Modulation

In recent years the increasing teletraffic demands led to compact frequency reuse patterns, such as micro- and pico-cellular structures, which exhibit more benign propagation properties than conventional macro-cells. Specifically, due to the low power constraint and smaller coverage area the transmitted power is typically channeled into street canyons. This mitigates both the dispersion due to far-field, long-delay reflections as well as cochannel interferences. This is particularly true for indoors channels, where interferences are further mitigated by the room partitions. Under these favorable propagation conditions multilevel modulation schemes have been invoked in order to further increase the teletraffic delivered, as in the American IS-54 and the Japanese digital cellular (JDC) system.

In this contribution, we created a range of reconfigurable source-sensitivity matched videophone transceivers, which have two modes of operation. Namely, a more robust, but less bandwidth efficient four-level QAM (4QAM) [24], [33] mode of operation for outdoors applications, accommodating eight users in a bandwidth of 200 kHz, as in the GSM system [57] and a less robust, but more bandwidth efficient 16-QAM mode, supporting 16 users in indoors cells or near the BS. In other words, indoors cells exploit the prevailing higher SNR and signal-to-interference ratio (SIR) invoking 16 QAM and thereby requiring only half as many TDMA slots as 4 QAM. When the portable station (PS) is roaming in a lower-SNR outdoors cell, the intelligent basestation (BS) reconfigures the system to operate at 4 QAM.

Recently, transmission techniques, such as differentially coded noncoherent star QAM (StQAM) [43], coherent pilot symbol assisted modulation (PSAM) [53] and the transparent tone in band [54], [55] (TTIB) modulation have been proposed in order to facilitate multilevel signaling over mobile channels. A surge of QAM research activities is hallmarked by [34]–[46]. Amplifier linearity problems have been addressed in [51], [52], although at low transmitted power levels the class-A amplifier's excess power requirement is not significant. While in indoors scenarios the interference mitigating effects of the partitions are useful, in outdoors cells the channel segregation algorithm proposed in [47] can be used to segregate more interfered time slots for 4-QAM transmissions, while low-interference slots for 16-QAM communications. Instead of tolerance-sensitive linear-phase Nyquist filtering nonlinear filtering (NLF) joining time-domain signal transitions with a smooth curve can be employed [24]. In case of coherent detection, typically 3-dB lower channel SNR is sufficient over AWGN channels than in case of lower-complexity noncoherent differential modems. Hence, in the proposed system, we opted for using diversity- and pilot-aided rectangular 4- and 16-QAM schemes [24], [53].

[3]The subjective video quality of these schemes can be compared using the down-loadable demonstration package under the www address http://www-mobile.ecs.soton.ac.uk.

References [19], [23], and [24] suggested that the rectangular 16-QAM constellation possesses two independent 2-b subchannels having different BER, which naturally lend themselves to unequal protection coded modulation. The BER of the lower integrity C2 subchannel was found a factor two–three times higher than that of the C1 subchannel. These BER differences can be adjusted using appropriate FEC codes to match arbitrary bit sensitivity requirements.

B. Forward Error Correction

For convolutional codes to operate at their highest achievable coding gain it is necessary to invoke soft-decision information [48]. Furthermore, in order to be able to reliably detect decoding errors they are typically combined with an external error detecting block code, as in the GSM system [57]. Powerful block codes possess an inherently reliable error detection capability [48], which is a useful means of monitoring the channel's quality in order to activate automatic repeat requests (ARQ), handovers or error concealment. In the proposed systems we have opted for binary BCH codes.

In Systems 1, 3, and 6, we used the $R = 71/127 \approx 0.56$-rate BCH(127,71,9) code in both 16-QAM subchannels, which, in conjunction with the 11.36-kb/s video Codec 1, resulted in $1136 \times 127/71 = 2032$ b/frame and a bit rate of 20.32 kb/s at an image frame rate of 10 frames/s. These system features are summarized in Table IV. In contrast, in Systems 2, 4, and 5, we employed the 8-kb/s video Codec 2, which achieved this lower rate essentially due to invoking RL coding in order to compress the motion- and VQ-activity tables. Since erroneous RL-coded bits corrupt the entire video frame, their protection is crucial. Hence, in Systems 2, 4, and 5, we decided to use the stronger BCH(127,50,13) code, which after FEC coding yielded the same 20.32-kb/s bit rate, as the remaining systems. This will allow us to assess in Section VI, whether it is a worthwhile complexity investment in terms of increased system robustness to use vulnerable RL coding in order to reduce the bit rate and then accommodate a stronger FEC code at the previous 20.32-kb/s overall bit rate.

Before FEC coding, the output of Codec 1 was split in two equal sensitivity classes, Class One and Two. The bit stream of Codec 1 was split in two sensitivity classes according to our findings in Fig. 13. 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. In case of Codec 2, all RL encoded bits were assigned to Class 1 and then the remaining bits were assigned according to their sensitivities. Let us now consider the architecture of System 1.

C. Architecture of System 1

Transmission Format: The transmission packets of System 1 are constituted by a Class One BCH(127,71,9) codeword transmitted over the C1 16-QAM subchannel and a Class Two BCH(127,71,9) codeword conveyed by the C2 subchannel and a stronger BCH(127,50,13) codeword for the packet header. The generated 381-b packets are represented by 96 16-QAM symbols. Then, 11 pilot symbols are inserted according to a pilot spacing of $P = 10$ and four ramp symbols are concatenated

TABLE IV
SUMMARY OF SYSTEM FEATURES

Feature	System 1	System 2	System 3	System 4	System 5	System 6
Video Codec	Codec 1	Codec 2	Codec 1	Codec 2	Codec 2	Codec 1
Video rate (kbps)	11.36	8	11.36	8	8	11.36
Frame Rate (fr/s)	10	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)	BCH(127,71,9) BCH(127,71,9) BCH(127,50,13)
FEC-coded Rate (kbps)	20.32	20.32	20.32	20.32	20.32	20.32
Modem	4/16-PSAQAM	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	Cl. One
User Signal. Rate (kBd)	18 or 9	9	18 or 9	18 or 9	18 or 9	9
System Signal. Rate (kBd)	144	144	144	144	144	144
System Bandwidth (kHz)	200	200	200	200	200	200
No. of Users	8-16	(16-2)=14	6-14	6-14	8-16	(16-2)=14
Eff. User Bandwidth (kHz)	25 or 12.5	14.3	33.3 or 14.3	33.3 or 14.3	33.3 or 14.3	14.3
Min. AWGN SNR (dB) 4/16QAM	5/11	11	4.5/10.5	6/11	8/12	12
Min. Rayleigh SNR (dB) 4/16QAM	10/22	15	9/18	9/17	13/19	17

over which smooth power amplifier ramping is carried out in order to mitigate spectral spillage into adjacent frequency bands. Eight 111-symbol packets represent a whole image frame, and hence, the signaling rate becomes 111 symb/12.5 ms $\approx$ 9 kBd. When using a 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 signaling 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 [57].

If the channel conditions degrade and 16-QAM communications cannot be supported, the BS instructs the system to switch to 4 QAM, which requires twice as many time slots. The 381-b packets are now conveyed by 191 4-QAM symbols and after adding 20 pilot symbols and four ramp symbols the packet-length becomes 225 symb/12.5 ms resulting in a signaling rate of 18 kBd. Hence, the number of videophone users supported by System 1 is reduced to eight, as in the full-rate GSM speech channel. The system also facilitates mixed-mode operation, where 4-QAM users must reserve two slots in each 12.5-ms TDMA frame toward the fringes of the cell, while in the central section of the cell 16-QAM users will only require one slot per frame in order to maximize the number of users supported. Assuming an equal proportion of four and 16-QAM users the average number of users per carrier becomes 12. The equivalent user bandwidth of the 4-QAM PS's is 200 kHz/8 = 25 kHz, while that of the 16-QAM users is 200 kHz/16 = 12.5 kHz. The characteristics of the whole range of our candidate systems are highlighted in Table IV.

D. Architecture of System 2

As already alluded to in Section V-B, the philosophy behind System 2 was to contrive a scheme, which allowed us to assess the value of Codec 2 in a systems context. Specifically, whether it pays dividends in robustness terms to invest complexity and hence reduce the 11.36-kb/s rate of Codec 1 to 8 kb/s and then accommodate a more powerful, but more complex BCH(127,50,13) code within the same 20.32-kb/s bit-rate budget. The sensitive RL and BCH(127,50,13) coded Class One bits are then conveyed over the lower BER C1 16-QAM subchannel. Furthermore, in order to ensure the error-free reception of the RL-coded Class One bits, ARQ is invoked, if the C1 BCH(127,50,13) code signals their corruption. Similarly to System 1, System 2 can also accommodate 16 time slots for the nine kBd users, but if we want to improve the integrity of the RL-coded video bits by ARQ, some slots will have to be reserved for the ARQ packets. This proportionately reduces the number of users supported, as the channel conditions degrade and requires more ARQ attempts. The underlying trade-offs will be analyzed in Section VI.

Automatic Repeat Request: ARQ techniques have been successfully used in data communications [27]–[30] 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 acknowledgment 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 acknowledgment, and the short TDMA frame length ensures a low packet delay, hence, ARQ can be invoked.

System 2 was contrived such that for the first transmission attempt (TX1) we use contention-free TDMA and 16 QAM to deliver the video packet. If the Class One bits transmitted over the C1 16-QAM subchannel are corrupted and an ARQ-request occurs, only the Class One bits are retransmitted using the "halved-capacity," but more robust four-QAM mode of operation. These packets have to contend for a number of earmarked time slots similarly to packet reservation multiple access (PRMA) [24]. If, however, there are only C2-bit errors in the packet, it is not retransmitted. Since the Class One bits are delivered by the lower BER C1 16-QAM subchannel, often they may be unimpaired, when there are Class Two errors. In order to strike a compromise between the minimum required channel SNR and the maximum number of users supported we limited the number of transmission attempts to three. Hence, two slots per frame must be reserved for ARQ. For the sake of 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 of this scheme 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 IV.

E. Architecture of Systems 3–6

With the aim of exploring the whole range of design options we have created four further systems, namely Systems 3–6.

System 3 is similar to System 1, since it is constructed of the 11.36-kb/s Codec 1, BCH(127,71,9) FEC codecs and reconfigurable 16-QAM/4-QAM modems, but it relies on ARQ assistance. However, in contrast to System 2, where only the Class One bits were retransmitted using 4 QAM, here both the Class One and Two bits are retransmitted using the same modulation scheme, as during the first attempt. System 3 also limits the maximum number of transmission attempts to three in case of C1 BCH(127,71,9) decoding errors and therefore reduces the number of users by two, but improves the transmission integrity. Explicitly, in the 16- and 4-QAM modes, 14 and six users can be supported, respectively.

System 4 is similar to System 2, since it employs the 8-kb/s Codec 2 in conjunction with the robust BCH(127,50,13) FEC codecs and the 16-QAM/4-QAM modems, but instead of Class One retransmission it invokes full packet retransmissions, if the Class One bits are corrupted. The number of video users is again reduced to 14 and 6 in the 16- and 4-QAM modes, respectively. *System 5* is a derivative of System 4, which dispenses with ARQ and, hence, can serve 16/8 users in the 16-QAM/4-QAM modes of operation, respectively. Last, *System 6* is related to System 1, since it is constructed of the same components, but allows ARQ's in the same fashion, as System 2, where in case of Class One errors only these bits are retransmitted using the previously introduced 16-QAM/4-QAM/4-QAM scheme. Therefore, System 6 can support 14 users.

VI. System Performance

A. Simulation Environment

Before commenting on the system's performance let us consider the simulation environment and the system parameters. The 11.36 and 8-kb/s VQ video codecs, the previously mentioned BCH codecs and pilot-assisted, diversity and ARQ-aided four and 16-QAM modems were simulated, including the AWGN and Rayleigh channel. The receiver carried out the inverse functions of the transmitter. Simple switched diversity was implemented. The criterion for deciding upon which diversity channel to decode was the minimum phase shift of the pilot symbols, rather than the maximum received signal power, since this has resulted in slightly better BER performance. As already alluded to in Section V-D, the ARQ scheme, which was invoked in various situations in the different systems of Table IV, was controlled by the error detection capability of the BCH codes used, which was extremely reliable in case of the powerful codes used and, hence, allowed us to dispense with cyclic redundancy coding (CRC).

In interactive communications, ARQ techniques have not been popular due to their delay and the required feed-back channel for message acknowledgment. In intelligent low-delay picocells, however, it is realistic to assume a near-instantaneous full duplex control link between the BS and PS, for example, for ARQ acknowledgment, which in our case were always error-freely received. In order to compromise between traffic capacity and robustness only two slots per frame were reserved for ARQ, limiting the number of transmission attempts to three, where we allocated these two time slots to that particular user, whose packet was first corrupted within the TDMA frame. Since in the ARQ-assisted arrangements retransmission took place within the last two slots of the same TDMA frame, no extra delay was inflicted. As regards to the video packet of the last user in the frame, transmitted during the third slot from the end of the frame it is irrealistic however to expect a near-instantaneous acknowledgment even in low-delay picocells, implying the lack of ARQ assistance for this user. The specific criterion for invoking ARQ depended on which system of Table IV was used.

Our results refer to the previously mentioned signaling rate of 144 kBd, while the propagation frequency and the vehicular speed were 1.8 GHz and 30 mph, respectively. In case of lower walking speeds the signal's envelope fades more slowly, therefore the pilots provide a better fading estimate and the PSAM modem has a lower BER. On the other hand, the fixed-length BCH codecs have to face a somewhat more pessimistic scenario, since in some codewords despite interleaving there may be a high number of errors, and the error correction capability may be overloaded. In contrast, in some codewords there may be only a low number of errors, which does not fully exploit the BCH codec's power. Throughout these experiments noise-limited, rather than interference-limited operation was assumed, which is a realistic assumption in partitioned indoors picocells or in the six-cell microcellular cluster used for example in the Manhatten model of [24, ch. 17]. If, however, the decoded BCH frame error rate due to either interference or noise, which is monitored by the BS becomes high for 16-QAM communications, the BS may instruct the system to switch to 4 QAM, thereby reducing the number of users supported. The consideration of these reconfiguration algorithms is beyond the scope of this treatise. Having summarized the simulation conditions, let us now concentrate on the performance aspects of the candidate systems.

B. Performance of Systems 1 and 3

Recall from Table IV that a common feature of Systems 1, 3, and 6 was that these all used the 11.36-kb/s video codec associated with the BCH(127,71,9) FEC codec, although Systems 3 and 6 used additionally ARQ in order to enhance their robustness at the price of reserving two more time slots and hence reducing the number of users supported. The PSNR versus channel SNR (ChSNR) performance of System 1 is portrayed in Figs. 17 and 18 for its 4-QAM/16-QAM 18/9-kBd modes of operation, respectively. Observe in these figures that the system's performance was evaluated for the best-case AWGN channel and the worst case Rayleigh (R) channel both with diversity (D) and with no diversity (ND), using one transmission attempt (TX1). However, in order to be able to assess the benefits of ARQ in terms of required ChSNR reduction Figs. 17 and 18 also display the corresponding results for System 3, using three transmission attempts (TX3).

Each system was deemed to deliver nearly un-impaired video, if its PSNR was reduced by less than 1 dB due to channel impairments. These corner SNR values are tabulated for ease of reference in Table IV for each system studied. Considering System 1 first, in its 4-QAM/18-kBd mode of operation over AWGN channels a channel SNR of about 5

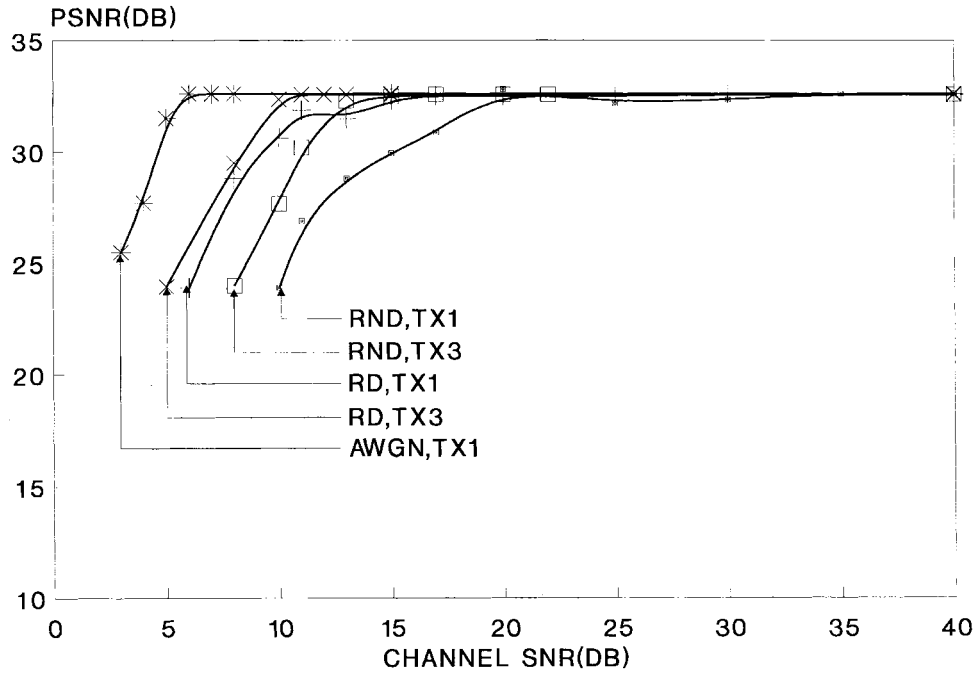

Fig. 17. PSNR versus channel SNR performance of Systems 1 and 3 in their 4-QAM 18-kBd mode of operation over various channels.

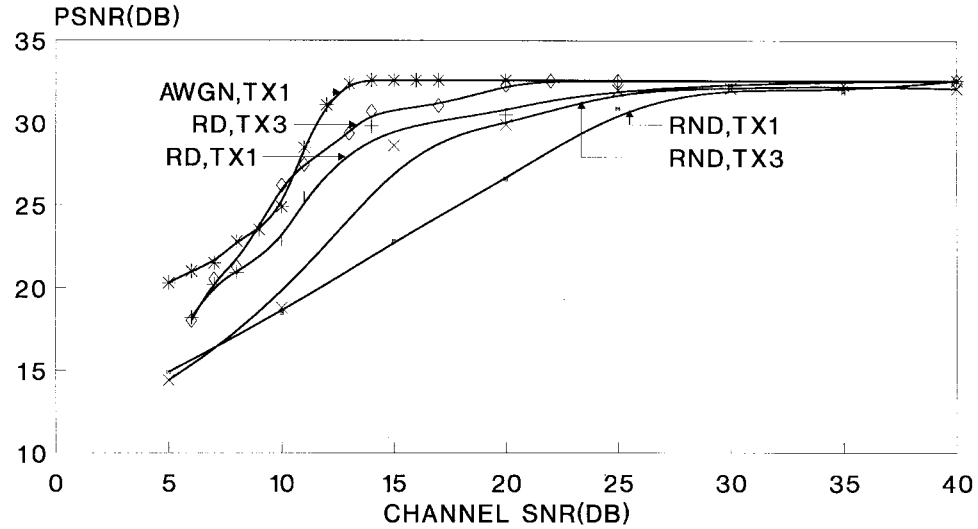

Fig. 18. PSNR versus channel SNR performance of Systems 1 and 3 in their 9-kBd mode of operation over various channels.

dB was required, which had to be increased to around 10 and 18 dB for the RD, TX1, and RND, TX1 scenarios. When opting for the doubled bandwidth efficiency, but less robust 16-QAM/9-kBd mode of System 1, the corresponding ChSNR values were 12, 23, and 27 dB, respectively. This trade-off becomes explicit in Fig. 19.

Comparing the above values in Fig. 17 to the corresponding ChSNR's of System 3 reveals an approximately 6-dB reduced corner SNR when using 4 QAM with no diversity, which significantly eroded in the inherently more robust diversity-assisted scenario. In the 16-QAM/9-kBd mode, Fig. 18 shows approximately 3-dB ARQ gain with no diversity, which is increased to about 4–5 dB when using the diversity-aided arrangement.

C. Performance of Systems 4 and 5

Systems 4 and 5 combine the lower rate, but more error sensitive RL-coded 8-Codec 2 with the more robust BCH(127,50,13) codec. These systems are more complex to implement, and hence, we can assess, whether the system benefited in robustness terms. The corresponding PSNR versus ChSNR curves are plotted in Figs. 20 and 21 for the 4-QAM/18-kBd and the 16-QAM/9-kBd modes of operation, respectively. Considering the 4-QAM/18-kBd mode of System 4 first, where a maximum of 3 transmission attempts are used, if any of the C1 or C2 bits are corrupted, an AWGN performance curve similar to that of System 3 is observed. Over Rayleigh channels with diversity about 9 dB, ChSNR

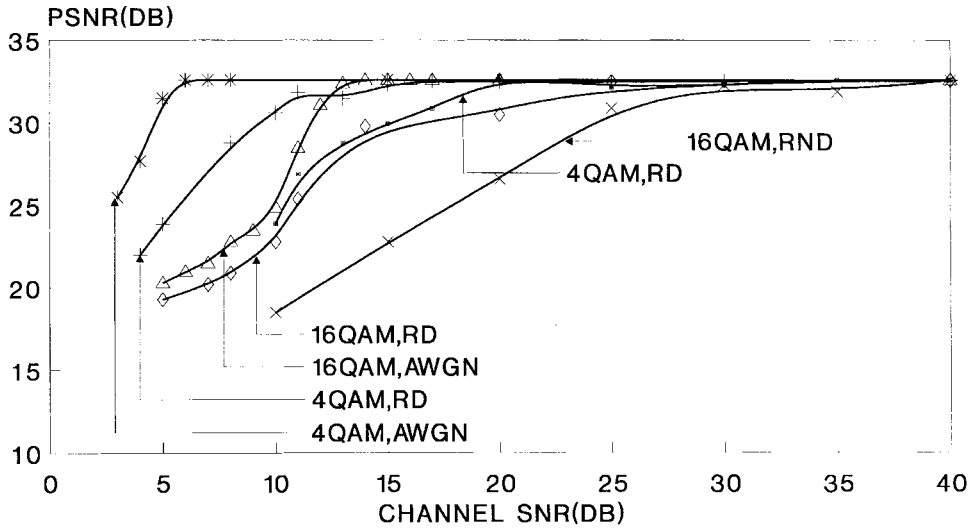

Fig. 19. PSNR versus channel SNR performance of System 1 in its 4-QAM/16-QAM 18/9-kBd modes of operation over various channels.

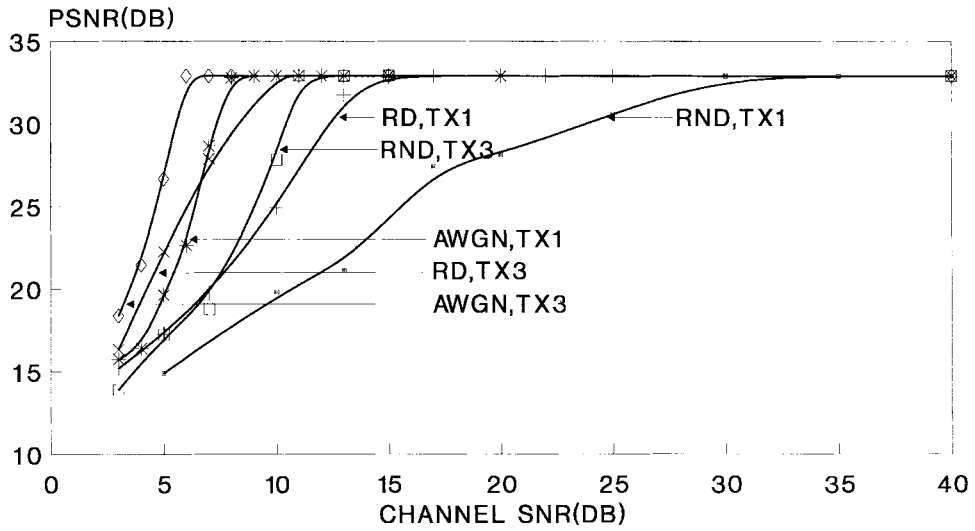

Fig. 20. PSNR versus channel SNR performance of the 4-QAM/18-kBd modes of Systems 4 and 5 over various channels.

is necessitated for near-unimpaired performance, increasing to around 11 dB with no diversity. In the 16-QAM mode, the AWGN performance is similar to that of System 3. The ChSNR over Rayleigh channels with diversity must be in excess of about 17 dB, increasing to around 24 dB without diversity. As expected, System 5 is typically less robust than the ARQ-assisted System 4. It is interesting to observe that the RL-coded Systems 4 and 5, despite more robust FEC coding, tend to be less robust than the corresponding schemes without RL-coding, such as System 1 and 3.

D. Performance of Systems 2 and 6

The situation is slightly different, when only the Class One bits are retransmitted in case of erroneous decoding. These Systems were contrived to constitute a compromise between robustness and high-user capacity. For the RL-coded System 2 it is vital to invoke ARQ, whereas for System 6 it is less critical. However, System 6 uses the weaker BCH(127,71,9) code. The PSNR results of Fig. 22 reveal that due to the stronger FEC code System 2 is more robust than System 6. However, for all studied scenario the degradation of System 2 was more rapid with decreasing ChSNR values than that of System 6. The minimum required ChSNR values are given in Table IV.

As expected, in all systems, irrespective of the mode of operation invoked, best performance was always achieved over AWGN channels, followed by the diversity-assisted Rayleigh (RD) curves, while the least robust curves represented the system's performance with ND. The ARQ attempts did improve the performance of all systems, but the lowest ARQ gains were offered over AWGN channels, since due to the random error distribution in every new attempt the video packet was faced with similar channel conditions, as during its first transmission. The highest ARQ gains were typically maintained over Rayleigh channels with ND, since during retransmission the signal envelope had a fair chance of emerging from a deep fade, increasing the decoding success probability.

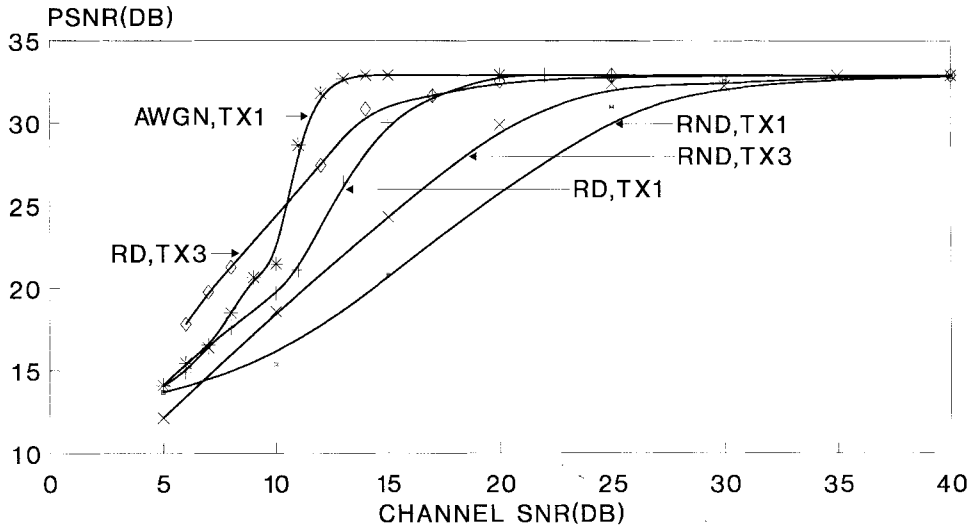

Fig. 21. PSNR versus channel SNR performance of the 16-QAM/9-kBd mode of Systems 4 and 5 over various channels.

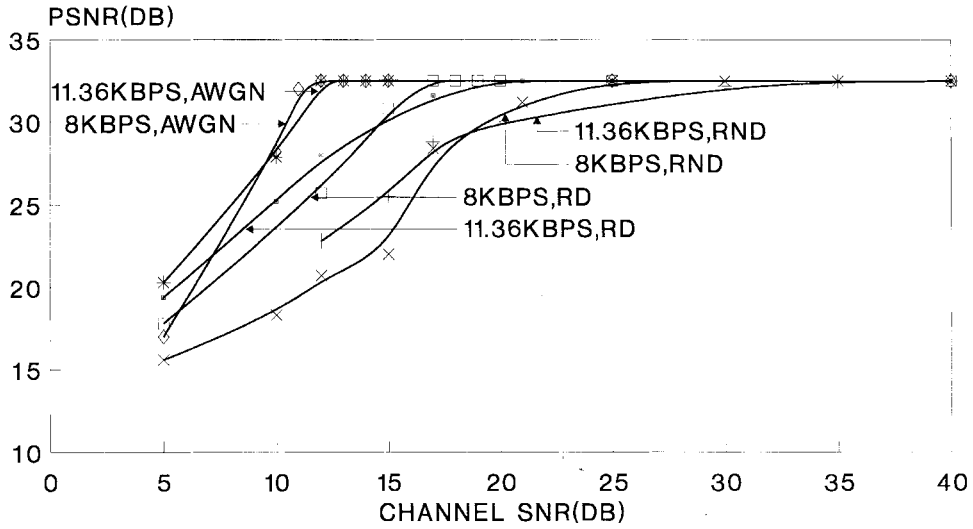

Fig. 22. PSNR versus channel SNR performance of Systems 2 and 6.

VII. Summary and Conclusions

Following the comparative study of a variety of VQ-based constant-rate video codecs, an 8- and an 11.36-kb/s scheme was earmarked for further study in the context of various systems. It was demonstrated that the proposed VQ codecs slightly outperformed the identical-rate DCT and QT codecs. The various studied reconfigurable mobile videophone arrangements are characterized by Table IV. The video codecs proposed are programmable to any arbitrary transmission rate in order host the videophone signal by conventional mobile radio speech channels, such as the Pan-European GSM system, the IS-54, or IS-95 systems as well as the Japanese digital cellular system.

In Systems 1, 3, and 6 the 11.36-kb/s Codec 1 was invoked. The BCH(127,71,9) coded channel rate was 20.32 kb/s. Upon reconfiguring the 16- and 4-QAM modems depending on the BCH decoded frame error rate due to the prevailing noise and interference conditions experienced, signaling rates of 9 and 18 kBd, respectively, can be maintained. Hence, depending on the noise and interference conditions, between eight and 16 videophone users can be accommodated in the GSM bandwidth of 200 kHz, which implies minimum and maximum user bandwidths of 12.5 and 25 kHz, respectively.

In Systems 2, 4, and 5 the RL-coded 8-kb/s Codec 2 performed similarly to Codec 1 in terms of PSNR, but it was more error sensitive. The lower rate allowed us, and the lower robustness required us to use the stronger BCH(127,50,13) FEC codecs. Further system features are summarized in Table IV.

As expected, the 4-QAM/18-kBd mode of operation of all systems is more robust than the 16-QAM/9-kBd mode, but it can only support half the number of users. Explicitly, without ARQ, eight or 16 users can be supported in the 4-QAM/18-kBd or 16-QAM/9-kBd modes. When ARQ is used, these numbers are reduced by two in order to reserve slots for retransmitted packets. A compromise scheme is constituted by Systems 2 and 6, which retransmit only the corrupted Class One bits in case of decoding errors and halve the number of bits per modulation symbol, facilitating higher

integrity communications for 14 users. Both Systems 2 and 6 slightly outperform in terms of robustness the 16-QAM modes of Systems 1 and 5, but not their 4-QAM mode, since the initially erroneously received 16-QAM Class Two bits are not retransmitted and hence remain impaired. Systems 2, 4, and 5, which are based on the more sensitive Codec 2 and the more robust BCH(127,50,13) scheme, tend to have a slightly more robust performance with decreasing channel SNR's than their corresponding counterparts using the Codec 1 and BCH(127,71,9) combination, namely Systems 6, 3, and 1. However, once the error correction capability of the BCH(127,50,13) codec is overloaded, the corresponding PSNR curves decay more steeply, which is explained by the violently precipitated RL-coded bit errors. This becomes quite explicit in Fig. 22. Clearly, if best SNR performance is at premium, the Codec 2 and BCH(127,50,13) combination is preferable, while in terms of low complexity the Codec 1 and BCH(127,71,9) arrangement is more attractive.

Careful inspection of Table IV provides the system designer with a vast plethora of design options, while documenting their expected performance. Many of the listed system configuration modes can be invoked on a time-variant basis and, hence, do not have to be preselected. In fact, in a true "software radio" [13] these modes will become adaptively selectable in order to comply with the momentary system optimization criteria.

The most interesting aspects are highlighted by contrasting the minimum required channel SNR values with the systems' user capacity, in particular, in the purely 4-QAM- and entirely 16-QAM-based scenarios. System 2 is an interesting scheme, using 4 QAM only during the Class One ARQ attempts, thereby ensuring a high-user capacity, although the 11 and 15 dB minimum channel SNR requirements are only valid for one of the 14 users, whose operation is 4-QAM/ARQ assisted. The remaining 13 users experience a slightly extended error-free operating SNR range in comparison to the 16-QAM System 1 mode due to the employment of the more robust BCH(127,50,13) Class One codec. The comparison of System 1 to the otherwise identical, but ARQ-aided System 3 suggests that over AWGN channels only very limited SNR gains can be attained, but over Rayleigh channels up to 4-dB SNR reduction can be maintained in the 16-QAM mode. In general, the less robust the modem scheme, the higher the diversity and AGC gains, in particular over Rayleigh channels. Hence, we concluded that both the diversity and ARQ gains are substantial and the slightly reduced user capacity is a price worth paying for the increased system robustness. Slightly surprisingly perhaps, but it is often beneficial to opt for the lower rate, higher sensitivity Codec 2, if the additional system complexity is acceptable, since due to the stronger accommodated BCH(127,50,13) codec an extended operating range is maintained. All in all, in terms of performance, System 4 constitutes the best arrangement with a capability of supporting between six and 14 users at minimum channel SNR's between nine and 17 dB. Here, we refrain from describing all potentially feasible systems, this is left for the interested reader to explore.

In summary, the proposed systems are directly suitable for mobile videophony by replacing a speech channel of conventional second generation wireless systems, such as the GSM, IS-54, IS-95, and the Japanese systems. Our future work is aimed at improving the complexity/quality characteristics of these systems using parametrically enhanced object oriented analysis by synthesis coding.

ACKNOWLEDGMENT

The helpful critique of the anonymous reviewers is gratefully acknowledged.

REFERENCES

[1] K. H. Tzou, H. G. Musmann, and K. Aizawa, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, pp. 213–357, June 1994.
[2] M. F. Chowdhury *et al.*, "A switched model-based coder for video signals," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, pp. 228–235, June 1994.
[3] G. Bozdagi *et al.*, "3-D motion estimation and wireframe adaptation including photometric effects for model-based coding of facial image sequences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, pp. 246–256, June 1994.
[4] C. S. Choi *et al.*, "Analysis and synthesis of facial image sequences in model-based image coding" *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, pp. 257–275, June 1994.
[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 unequal error protected px8 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*, Stockholm, Sweden, June 8–10, 1994, 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] L. Hanzo and J. Streit, "Adaptive low-rate wireless videophone systems," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 305–319, Aug. 1995.
[10] J. Streit and L. Hanzo, "Quadtree-based parametric wireless videophone systems," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 2, pp. 225–237, 1996.
[11] 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, pp. 15–27, Feb 1993.
[12] *IEEE Personal Comm.—The Mag. Nomadic Commun. Comp.*, Special Issue on the European Path Toward UMTS, vol. 2, pp. 12–63, Feb. 1995.
[13] J. Mitola, "Software radios," *IEEE Commun. Mag.*, vol. 33, no. 5, pp. 26–69, 1995.
[14] *Advanced Commun. Tech. Services (ACTS)*, Workplan, DGXIII-B-RA946043-WP, Commission of the European Community, Brussels, Belgium, 1994.
[15] J. W. Woods, Ed., *Subband Image Coding.* Boston, MA: Kluwer, 1991.
[16] H. Gharavi, "Subband coding of video signal," in *Subband Image Coding*, J. W. Woods, Ed. Boston, MA: Kluwer, 1991, ch. 6, pp. 229–271.
[17] K. N. Ngan and W. L. Chooi, "Very low bit rate video coding using 3-D subband approach," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, pp. 309–316, June 1994.
[18] L. Hanzo and J. P. Woodard, "An intelligent cordless voice terminal for indoors communications," *IEEE Trans. Veh. Technol.*, vol. 44, no. 4, pp. 735–749, 1995.
[19] L. Hanzo, R. Steele, and P. M. Fortune, "A subband coding, BCH coding and 16-QAM system for mobile radio communication," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 327–340, Nov. 1990.
[20] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379–423/623–656, June/Oct. 1948.
[21] J. Hagenauer, "Quellengesteuerte Kanalcodierung fuer Sprach- und Tonuebertragung im Mobilfunk, Aachener Kolloquium: Signaltheorie," Mar. 23–25, 1994, Mobile Kommunikationssysteme, pp. 67–76.
[22] A. J. Viterbi, "Wireless digital communication: A view based on three lessons learned," *IEEE Commun. Mag.*, pp. 33–36, Sept. 1991.
[23] L. F. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, vol. 41, pp. 1439–1450, Oct. 1993.

[24] W. T. Webb and L. Hanzo, *Modern Quadrature Amplitude Modulation: Principles and Applications for Wireless Communications*. New York: IEEE, 1994.
[25] L. Hanzo, R. Stedman, R. Steele, and J. C. S. Cheung, "A portable multimedia communicator scheme," *Chapter in Multimedia Technologies and Future Applications*, R. I. Damper, *et al.*, Eds. New York: Pentech, 1994, pp. 31–54.
[26] A. K. Jain, *Fundamentals of Digital Image Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1989.
[27] 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./Mar./Apr. 1994.
[28] L. de Alfaro and A. R. Meo, "Codes for second- and third-order GH-ARQ schemes," *IEEE Trans. Commun.*, pp. 988–910.
[29] T.-H. Lee, "Throughput performance of a class of continuous ARQ strategies for burst-error channels," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 380–386, Nov. 1992.
[30] S. Lin, D. J. Costello, and M. J. Miller, "Automatic-repeat-request error-control schemes," *IEEE Commun. Mag.*, pp. 5–17, Dec. 1984.
[31] "Speech and image coding," Special Issue of the *IEEE J. Select. Areas Commun.*, vol. 10, pp. 793–976, June 1992.
[32] R. A. Salami and L. Hanzo, *et al.*, "Speech coding," in *Mobile Radio Communications*, R. Steele, Ed. New York: Pentech, 1992, ch. 3, pp. 186–346.
[33] G. D. Forney *et al.*, "Efficient modulation for band-limited channels," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 632–647, Sept. 1984.
[34] K. Feher, "Modems for emerging digital cellular mobile systems," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 355–365, May 1991.
[35] 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.
[36] R. J. Castle and J. P. McGeehan, "A multilevel differential modem for narrowband fading channels," in *Proc. VTC'92*, Denver, CO, pp. 104–109.
[37] 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.
[38] Y. Kamio and S. Sampei, "Performance of reduced complexity DFE using bidirectional equalizing in land mobile communications," in *Proc. VTC'92*, Denver, CO, pp. 372–376.
[39] 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.
[40] E. Malkamaki, "Binary and multilevel offset QAM, spectrum efficient modulation schemes for personal communication," in *Proc. VTC'92*, Denver, CO, pp. 325–378.
[41] Z. Wan and K. Feher, "Improved efficiency CDMA by constant envelope SQAM," in *Proc. VTC'92*, Denver, CO, pp. 51–55.
[42] H. Sasaoka, "Block coded 16-QAM/TDMA cellular radio system using cyclical slow frequency hopping," in *Proc. VTC'92*, Denver, CO, pp. 405–408.
[43] W. T. Webb, L. Hanzo, and R. Steele, "Bandwidth-efficient QAM schemes for Rayleigh-fading channels," *Proc. Inst. Elect. Eng.*, vol. 138, no. 3, pp. 169–175, June 1991.
[44] T. Sunaga and S. Sampei, "Performance of multilevel QAM with postdetection maximal ratio combining space diversity for digital land-mobile radio communications," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 294–302, Aug. 1993.
[45] P. K. Ho, J. K. Cavers, and J. L. Varaldi, "The effects of constellation density on trellis-coded modulation in fading channels," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 318–326, Aug. 1993.
[46] J. K. Cavers and J. Varaldi, "Cochannel interference and pilot symbol assited modulation," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 407–414, Nov. 1993.
[47] M. Frullone *et al.*, "Investigation on dynamic channel allocation strategies suitable for PRMA schemes," *1993 IEEE Int. Symp. Circuits Syst.*, Chicago, IL, May 1993, pp. 2216–2219.
[48] K. H. H. Wong and L. Hanzo, "Channel coding," in *Mobile Radio Communications*, R. Steele, Ed. New York: IEEE Press, 1992, ch. 4, pp. 347–488.
[49] E. Biglieri and M. Luise, "Coded modulation and bandwidth-efficient transmission," in *Proc. 5th Tirrenia Int. Wkshp., Sept. 8–12, 1991*. Amsterdam, The Netherlands: Elsevier, 1992.
[50] *IEEE Commun. Mag.*, Special Issue on Coded Modulation, vol. 29, no. 12, Dec. 1991.
[51] A. S. Wright and W. G. Durtler, "Experimental performance of an adaptive digital linearized power amplifier," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 395–400, Nov. 1992.
[52] M. Faulkner and T. Mattson, "Spectral sensitivity of power amplifiers to quadrature modulator misalignment," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 516–525, Nov. 1992.
[53] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.
[54] A. Bateman and J. P. McGeehan, "Feedforward transparent tone in band for rapid fading protection in multipath fading," in *Inst. Elec. Eng. Int. Conf. Commun.*, 1986, vol. 68, pp. 9–13.
[55] A. Bateman, "Feedforward transparent tone in band: Its implementation and applications," in *IEEE Trans. Veh. Technol.*, vol. 39, pp. 235–243, Aug. 1990.
[56] 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.
[57] L. Hanzo and J. Stefanov, "The Pan-European digital cellular mobile radio system-known as GSM," in *Mobile Radio Communications*, R. Steele Ed. New York: Pentech, 1992, ch. 8, pp. 677–773.
[58] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston, MA: Kluwer, 1992.
[59] B. Ramamurthi and A. Gersho, "Classified vector quantization of images," *IEEE Trans. Commun.*, vol. 31, pp. 1105–1115, Nov. 1986.
[60] L. Torres and J. Huguet, "An improvement on codebook search for vector quantization," *IEEE Trans. Commun.*, vol. 42, pp. 208–210, Feb. 1994.

Jürgen Streit was born in Cologne, Germany, in 1968. He received the Dipl.-Ing. degree in electronic engineering from the Aachen University of Technology, Germany, in 1993. He is currently working toward the Ph.D. degree in image coding.

Since 1992 he has been with the Department of Electronics and Computer Science, University of Southampton, U.K., working with the mobile multimedia communications research group.

Lajos Hanzo (M'91–SM'92) received the Bachelor's degree in electronics in 1976 and the Ph.D. degree in 1983, both from the Technical University of Budapest, Budapest, Hungary.

During his 20-year career in telecommunications he has held various research and academic posts in Hungary, Germany, and the U.K. 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. He coauthored two books on mobile radio communications, published more than 140 research papers, and was awarded a number of distinctions. Currently, he is managing a range of research projects in the field of wireless multimedia communications under the auspices of the Engineering and Physical Sciences Research Council (EPSRC) U.K., and the European Advanced Communications Technologies and Services (ACTS) Programme.

Dr. Hanzo is a Member of the IEE.