Transmission of Subband-Coded Images via Mobile Channels

Robert Stedman, *Student Member, IEEE*, Hamid Gharavi, *Fellow, IEEE*, Lajos Hanzo, *Senior Member, IEEE*, and Raymond Steele, *Senior Member, IEEE*

Abstract—**The feasibility of a mobile video handset is investigated via Rayleigh-fading channels, where transmissions must be confined to the channel's coherence bandwidth to avoid the deployment of complex high-power-consumption channel equalizers. This necessitates the utilization of a low-bit-rate image codec error protected by embedded low-complexity Bose–Chaudhuri–Hocquenghem (BCH) codecs and spectrally efficient 16-level quadrature amplitude modulation (16-QAM). Motion-compensated nonuniform 7-band subband coding (SBC) with subband-specific scanning, adaptive quantization (Q), run-length coding (RLC), and adaptive buffering to equalize bit-rate fluctuations offer good objective and subjective image quality at moderate complexity and a bit rate of 55 kbit/s. Using a twin-class embedded-BCH error protection as well as pilot symbol and diversity-assisted 16-QAM, our 22-kBd candidate system yields unimpaired image quality for average channel signal-to-noise ratios (SNR) in excess of about 16–18 dB, when the mobile speed is 4 mi/h, an SNR value readily maintained in personal-communications mobile systems characterized by small cells.**

I. Introduction

It is envisaged that the future will see the emergence of portable, hand-held video telephones for businesses and, ultimately, for personal use. Both the speech and image codecs have to achieve high compression ratios to guarantee acceptable spectral efficiency. However, high-compression source codecs usually have high bit-error sensitivities, which necessitates the careful deployment of error-correction codes, particularly if the transmissions are via hostile fading channels. For reasons of lack of robustness against channel errors, multilevel quadrature amplitude modulation (QAM) schemes have been considered unsuited for fading mobile channels. However, our recent endeavors toward the joint optimization of source and channel codecs as well as 16-QAM modems were rewarded by quite robust speech-transmission schemes in the friendly microcellular propagation environment [1], [2]. Encouraged by our findings, we embarked on the investigation of an embedded video subband codec (SBC) [3], a Bose–Chaudhuri–Hochquenghem (BCH) error-correction codec, and a 16-QAM modem over microcellular mobile channels. The high bandwidth efficiency of the 16-QAM modem is crucial to reduce the bandwidth occupancy of the video signal to render the channel nondispersive, while the increased bit-error sensitivity is counteracted by the carefully matched BCH error-protection scheme, as well as by the pilot-symbol-assisted [13] fade-tracking AGC and second order diversity arrangement deployed.

This article examines some of the design factors and tradeoffs that must be considered when designing such a video link. We begin with a description of the proposed motion-compensated (MC) video codec and justify the choice of its parameters. Then we select the appropriate BCH codes to embed the error-protected SBC-coded video bits into the modulated signal to ensure higher protection for more vulnerable image bits. Lastly, the overall transmission scheme is evaluated in terms of objective and subjective image quality and robustness against channel errors.

Manuscript received January 9, 1992; revised June 8, 1992. Paper was recommended by Associate Editor Yasuhiko Yasuda.

R. Stedman, L. Hanzo, and R. Steele are with the Department of Electronics and Computer Science, University of Southampton, Southampton, SO9 5NH, UK.

H. Gharavi is with Bellcore, Red Bank, NJ 07701.

IEEE Log Number 9203957.

II. The Subband Image Codec

A. Outline of the Subband Codec

Motion-compensated SBC image codecs of the kind shown in Fig. 1 have been studied by Gharavi [3], [8], [16]. The locally reconstructed previous frame is subtracted from the present source frame to be encoded in the motion-compensated prediction loop. In this way the prediction-error frame is generated along with an array of motion vectors, which have to be encoded and multiplexed for transmission via the mobile-radio channel. As seen in Fig. 1, the motion vectors are computed and encoded in the block MV, while the error frame undergoes more complex signal processing to squeeze out more redundancy and to allocate coding capacity where it is most needed to maintain high image quality at low bit rate. Therefore, the error frame is split into 10 subbands $B0,\cdots, B9$ by the two-dimensional (2-D) quadrature mirror filter (QMF) [9]. Then the subjectively more important low-frequency subbands are allocated more bits than their less important high-frequency counterparts. In very low-bit-rate applications no bits are allocated to the extreme high bands at the cost of some image-quality degradation.

The 2-D subbands are then scanned in the SCAN block to a one-dimensional (1-D) stream for encoding. As seen in Fig. 1, run-length coding (RLC) with variable-length

1051-8215/93$03.00

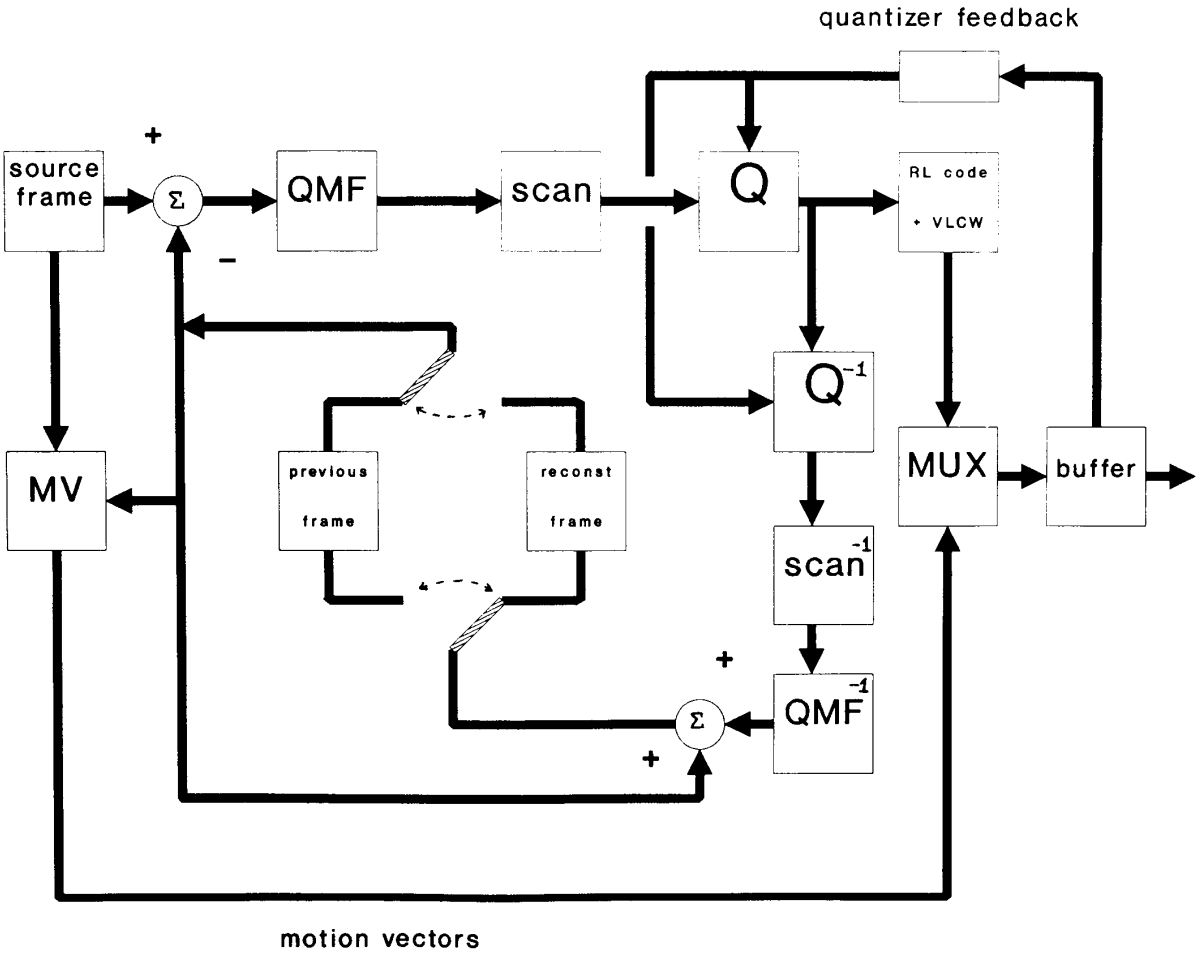


Fig. 1. Full subband coder.

codewords (VLCW) will be deployed after the quantization (Q) of the 1-D-scanned signal, where the subband-scanning algorithms are adapted to the subband features to minimize the bit-rate by generating long runs. However, the time-variant motion activity and the bandwidth-efficient VLCW's result in a variable bit rate, which is inconvenient for our proposed radio transmitter. Hence, buffering with adaptive feedback control is invoked. When the "buffer fullness," defined as the proportion of the buffer already filled with bits, is approaching overflow, the quantizer feedback modifies the quantizer parameters to generate less bits in subsequent coding steps. The exact control algorithm will be highlighted as further details are unraveled. Here, suffice it to say that the quantizer parameters t and d are adaptively increased or decreased to maintain the targeted "buffer fullness" and constant bit rate.

Finally, the coded data are multiplexed with the motion vectors and conveyed to the radio transmitter entailing the forward-error-correction (FEC) channel codec and the 16-QAM modem. Also, the quantized subband signals are inverse quantized, scanned back in 2-D, and the full-band-motion error frame reconstructed in the QMF^{-1} block. This locally reconstructed error frame is added to the previously reconstructed image frame to generate the current one to be utilized in the next motion-compensation step.

Having portrayed the SBC video codec in global terms, we now derive its essential parameters.

B. Interframe Prediction

The interframe prediction technique deployed in the coder is based on a commonly used block-motion-compensation algorithm. The source picture is divided up into a regular grid of typically 16×16 pel blocks. Each source-picture block is slid along neighboring blocks in the previous frame to find the most highly correlated "best match" in the search area. The motion vector for that block is defined as the spatial offset between the current source block and the best matching block from the previous frame. Assuming that $s_n(x, y)$ is the nth frame in the source sequence with $x = 0, 1, \cdots, X_{\max} - 1$, $y = 0, 1, \cdots, Y_{\max} - 1$ and $b_n^{k,l}(i, j)$ being the block at position (k, l) in frame n, where $i, j = 0, 1, \cdots, B_{\max} - 1$ with

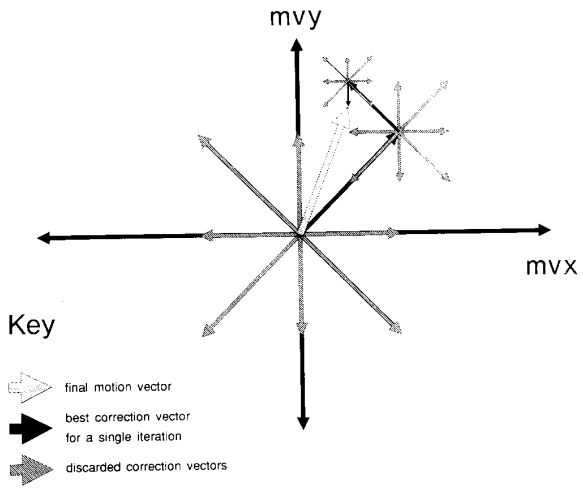


Fig. 2. Suboptimal motion-vector search.

$B_{\max}$ being the block size, we have

$$b_n^{kl}(i,j) = s_n(kB_{\max} + i, lB_{\max} + j). \quad (1)$$

A block-matching measure e for a given motion vector (mv_x, mv_y) is defined as

$$e_n^{kl}(mv_x, mv_y) \sim \sum_{i=0}^{B_{\max}-1} \sum_{j=0}^{B_{\max}-1} d\left\{b_n^{kl}(i,j) - b_{n-1}^{kl}(i + mv_x, j + mv_y)\right\} \quad (2)$$

where the distortion function d is taken as $d(x) = x^2$. The motion vector which gives the best match is defined as that minimizing e for a given set of trial motion vectors.

As frame-to-frame motion translation in video phone sequences is slow, only motion vectors within a small range, for example $(\pm 7, \pm 7)$ need to be considered. To avoid the computational demands of computing e for all motion vectors within the chosen search range, the optimum motion vector is estimated using an algorithm proposed by Jain and Jain [10].

The algorithm initially estimates the block motion vector as $(0,0)$. A set of eight correction vectors arranged on a grid, as seen in Fig. 2, is in turn added to the estimated motion vector, and a new value of e is computed each time. The correction vector that yields the lowest value of e so far computed is added to the estimated motion vector to obtain a more refined estimate. The size of the grid of correction vectors is then halved. The process repeats until the final grid size is 3×3, at which point the algorithm terminates with an estimate of the best motion vector. When comparing this suboptimum algorithm with the optimal full search, the motion-compensated error energy is some 10% higher at a five-times-shorter search time, a definitely worthwhile tradeoff. In our further experiments we favor the suboptimum approach.

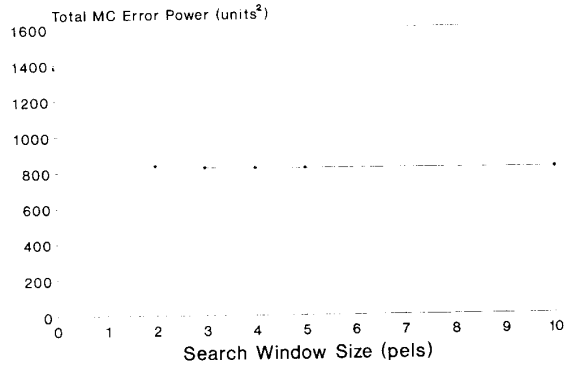


Fig. 3. Motion-compensated error versus search-window size.

The motion vectors are simply encoded with a fixed-length code, where an 8-bit code is sufficient for the range $(\pm 7, \pm 7)$. The bit rate economy can be improved if motion vectors are VLCW-coded. The search area of $(\pm 7, \pm 7)$ and block-size value of 16×16 pels represent a good compromise in terms of transmission bit rate for motion vectors, search time, and motion-compensated error energy. This is demonstrated for a typical video-conferencing sequence by Fig. 3, where the motion-compensated error energy is seen to become virtually nondecreasing, as the search-window size is increased beyond 7.

An interesting worse-case video-codec performance bound is arrived at when the MC error frame is not transmitted at all, i.e., it is assumed to be zero, and motion vectors are used merely to reconstruct the image both locally and at the receiver. Initially the codec has an astonishingly high performance for the low bit-rate required, particularly for slowly moving objects. But as expected, with degrading motion estimates due to zero MC error feedback, the peak SNR exhibits a gradual decay as shown in Fig. 4, which is more dramatic for the faster changing "Salesman" sequence than for "Miss America." When transmitting the encoded MC error signal at low bit rates, initially only limited peak SNR improvement is achieved, but with the increased time lapse the MC error feedback becomes more crucial for the codec's operation, yielding substantial peak SNR gains.

C. Quadrature Mirror Filters

The principle of 1-D QMF's is well understood [9] and filter designs are readily available in the literature [11]. A similar concept can also be applied to extend a 1-D QMF to a 2-D separable QMF designed for image applications. This is performed by applying a 1-D QMF along the rows and then down the columns, which results in a basic four-band decomposition. The decomposition of the image to more than four bands can be achieved by repeating the above process for each band in a tree structural manner to yield $16, 64, \ldots$ bands [5]. Better bit-rate economy is achieved when using nonuniform band splitting, where only the low horizontal and vertical spatial frequency

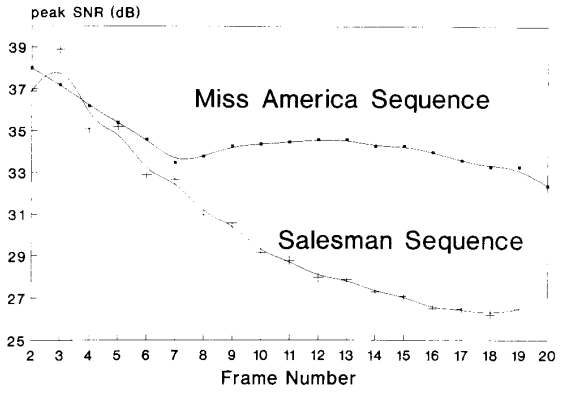


Fig. 4. Worst-case SBC SNR performance using motion vectors only.

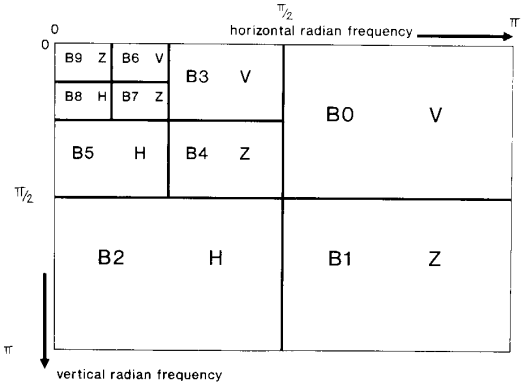


Fig. 5. SBC band-splitting scheme.

bands are further subdivided to enhance their reproduction [6], [7] yielding 7, 10, ... bands.

The 10-band scheme utilized in our systems is depicted in Fig. 5. The area associated with each band is proportional to its bandwidth and the number of samples in it, having undergone a number of decimation steps. For example, $B9$ represents the most important lowest horizontal and vertical frequency subband, which has to be quantized most accurately, while $B1$ is the least important highest horizontal and vertical spatial-frequency band. This explains the different scanning algorithms of the various subbands in order to exploit the spatial correlation more effectively. Clearly, when the band has a low horizontal content, it must be scanned horizontally (H), while low vertical frequency bands must be scanned vertically (V). Should both the H and V contents be of high frequency, zigzag (Z) scanning is used, as reflected by Fig. 6. Band $B9$ is an exception containing low H and V frequencies, yet it is also scanned in a Z pattern.

D. Quantization in Subbands

Quantization of appropriately scanned QMF outputs could be approached in an optimal least squares sense by matching Max–Lloyd quantizers to the subband probabil-

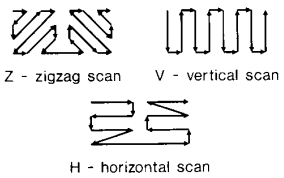
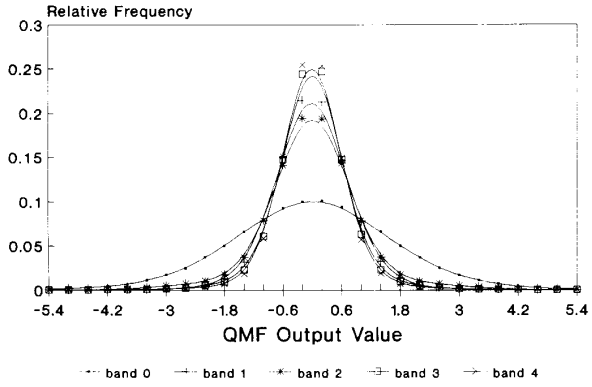
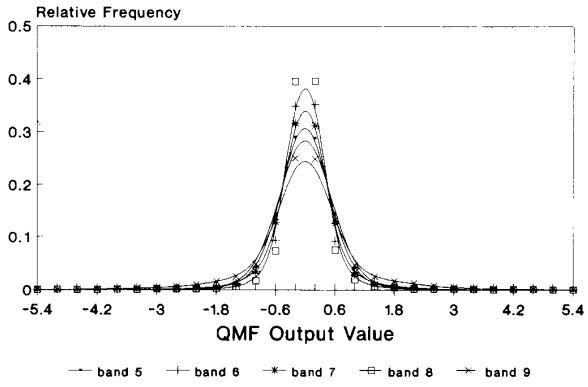


Fig. 6. Various subband scanning schemes.



(a)



(b)

Fig. 7. Typical subband PDF's.

ity density functions (PDF's) exemplified in Fig. 7 for the "Miss America" sequence. The highly peaked PDF's would require the subband quantizers to allocate many quantization levels around zero, yielding more coarse quantization in the PDF tails. Closer investigation of these PDF's reveals that, particularly in the high-frequency bands, they peak around zero due to low-level noise (e.g., camera noise) and their accurate representation by a Max–Lloyd quantizer is in fact detrimental to the overall subjective image quality, because it degrades the representation of the more important PDF tails.

To circumvent this problem, Gharavi and Tabatabai [6], [7] proposed a highly efficient subband quantizer with the

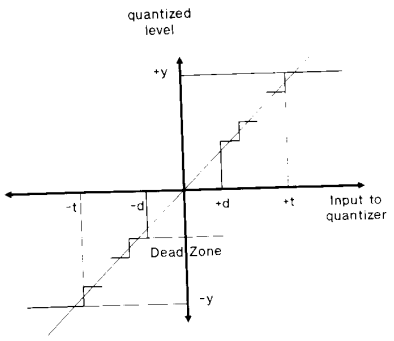


Fig. 8. Quantizer characteristic.

TABLE I
MAPPING RUN LENGTHS AND COLORS TO B1 CODE WORDS (C REFERS TO THE RUN COLOR)

Run length	B1 code
1	C0
2	C1
3	C0C0
4	C0C1
5	C1C0
EOL	C1C1
6	C0C0C0
7	C0C0C1
8	C0C1C0
.	.
.	.
.	.

features demonstrated in Fig. 8. The dead zone d eliminates camera noise, while low-probability extreme peaks are clipped at a maximum value of t. The input values in the active range are then uniformly quantized to L levels. The total number of levels is hence $(L + 3)$. For the lowest frequency band, however, where low-level camera noise is not dominant, the classic Max–Lloyd approach gives slightly better results, but for reasons of uniformity in our system we favored a quantizer with dead zone in this band as well.

The quantizer parameters $[d, t, L]_i$ have to be matched to the typical subband PDF's seen in Fig. 7. A number of approaches have been tested to contrive a range of codecs with various complexity/bit rate/quality tradeoffs. Here we highlight our approach that led to the 55-kbit/s low-bit-rate codec used in our proposed PCN video handset. Bearing in mind that the subsequent VLCW-based RLC produces a variable bit rate, buffering with adaptive dead-zone control is deployed, where the initial dead-zone values for each subband have to be determined individually.

1) Run Length coding: In this subsection we use the nomenclature "black" for zero-valued samples falling in the dead-zone domain and "white" for nonzero quantized samples of the active quantizer domain. Both the run colors black and white and the run length are coded in one variable-length code word. If a white run is coded, then the actual nonzero quantized levels are transmitted as fixed-length binary codes immediately following the white-run code.

The run-length code words are based on the so-called B1 codes that have been successfully used to encode subband images [3], [6], [7]. Table I shows how each run length and color maps to a unique B1 code, where C represents the run color, black or white. As a simple example, the quantized stream $\ldots 0, 0, 0, 0, 0, -3, 2, -1, 0, 0, 1 \ldots$ would be coded as [blackrun = 5 code] [whiterun = 3 code][fixed code -3][fixed code 2][fixed code -1][blackrun = 2], yielding the codes $\ldots$[0100] [1010][-3][2][-1][01]$\ldots$, where $[-3], [2], [-1]$ correspond to the nonzero quantized samples and are coded using fixed-word-length codes. Since only nonzero quantized pels are transmitted, it is therefore possible to encode $(L + 3)$ quantized levels by $\log_2(L + 3)$ bits. To improve codec robustness Table I also contains an end-of-line (EOL) code word. After coding one complete line (of length L pels), an EOL code is inserted, giving the output bit stream some redundancy, which is useful when decoding a corrupted bit stream.

The run-length decoding algorithm is straightforward. Black and white run lengths are read from the input bit stream. The run lengths are used to reconstruct a scan line of the coded picture. After each line is coded, an EOL code is inserted into the bit stream; thus, once a full line has been decoded an EOL is read. The redundancy introduced by the EOL code, coupled with the fact that the sum of run lengths on a given scan line is constant, gives the coding technique some robustness against errors. As a line is being decoded, the decoder checks whether the reconstructed line length exceeds the scan line length. If it does, then there must be at least one error in the bit stream. The error(s) are located somewhere in the bit stream read for this scan line. To combat single errors, the decoder systematically inverts each bit in the bit stream and redecodes the line until either the original line is recovered or all single errors were tried without success.

This decoding scheme cannot correct more than one error per scan line and does have some problems with single errors. If errors are found in a given line, the decoder systematically inverts bits in the bit stream and redecodes the line. Occasionally, the line will decode as error free when the wrong bit is inverted. Another problem which cannot be overcome crops up if there is an error in a nonzero value, since the error does not change any of the run lengths and so appears to the decoder as an error-free line. Clearly, the decoding algorithm performs well for a maximum of one error per scan line, if the bit in error is a run-length bit.

2) Output Bit Rate Control: When aiming for a constant SBC bit rate, in terms of complexity it is convenient to assume that each subband is encoded into a constant number of bits and its output rate is monitored on a line-by-line basis. The instantaneous bit rate of the specific line being encoded is compared against its targeted rate, and its dead zone d_i is appropriately increased, decreased, or maintained. The saturation clipping thresh-

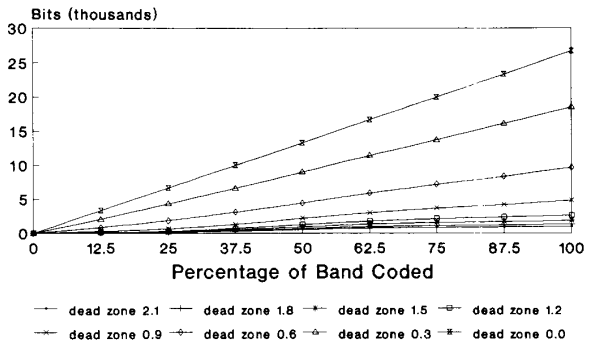


Fig. 9. Number of run-length coded bits versus proportion of pixels encoded for $B5$ using various dead-zone values.

olds $t_i, i = 0,\cdots,9$ are kept constant for each subband to moderate codec complexity. The choice of $t_i, i = 0,\cdots,9$ is based on the 1% quantiles of the subband PDF's seen in Fig. 7 to minimize the effects of quantizer saturation. Initial values for d_i will be derived after highlighting the dead-zone control algorithm.

With the aim of finding an appropriate dead zone and bit-rate-control algorithm, we evaluated in each subband the number of run-length coded bits as a function of the proportion of encoded pels for various dead-zone values. As a representative example, our results are shown for $B5$ in Fig. 9, where, as expected, there is a near-linear relationship for all the subbands. This prompted us to use a simple linear model to estimate the total number of B_{Total} encoded bits in any subband on the basis of encoding $x\%$ of its pels into B_{Model} number of bits, yielding

$$B_{\text{Model}}(x) = B_{\text{Total}}(d)\frac{x}{100} \tag{3}$$

where $B_{\text{Total}}(d)$ is the targeted total number of bits produced by quantizing and run-length coding the given band for a given dead-zone value d. This model can be used for all bands; however, the values of dead zone d and $B_{\text{Total}}(d)$ will be different for each of the 10 bands $B0,\cdots, B9$.

After coding each line of length L pels, the coder compares the actual number of coded bits produced $B_{\text{Actual}}(x)$ against the ideal number predicted by the model $B_{\text{Model}}(x)$ to arrive at a total of B_{Total}. Then a value for the excess number of coded bits $B_{\text{Excess}} = B_{\text{Actual}} - B_{\text{Model}}$ is computed. If B_{Excess} is positive, too many bits are being produced by the coder and hence the dead zone should be increased to lower the output bit rate for the next line, whereas d is decreased if B_{Excess} is negative. The dead zone d_{New} is then recomputed as a function of its current value D_{Current} and the current B_{Excess} value as

$$d_{\text{New}} = d_{\text{Current}}\left[1 + \frac{B_{\text{Excess}}(x)}{B_{\text{Model}}(x)}C\right] \tag{4}$$

where the constant $0.1 < C < 1$ is used to control the steepness of the algorithm. Again, for all subbands different d_{New}, B_{Total}, and C values have to be used.

TABLE II
QUANTIZER INITIAL VALUES FOR EACH BAND

Band	Range $\pm t$	Dead zone $\pm d$	Bits/White Value
0	N/A	N/A	0
1	N/A	N/A	0
2	N/A	N/A	0
3	6.0	4.4	1
4	5.0	4.7	1
5	4.8	3.3	1
6	4.8	2.1	1
7	4.8	2.0	1
8	4.8	1.6	1
9	5.8	2.6	1

As expected, the subband SNR's and hence the overall image SNR, as well as the average total bit rate, can be conveniently controlled by the parameters $[d, t, L]_i$, whereby t_i is fixed to the values listed in Table II to yield 1% clipping probability. According to our experiments, at bit rates of approximately 55 kbit/s it is more beneficial to allocate all the coding capacity to bands $B3,\cdots, B9$, while ignoring the less important high-frequency bands $B0,\cdots, B2$, assuming an image representation of 320×256 pels at 10 frames/s. To achieve better quality at higher bit rates, however, all the bands have to be appropriately quantized. In our experiments we observe that the bit rate can be dramatically reduced at graceful subband SNR degradation, allowing adequate subband representation at low bitrates. This tendency is noted for all subbands in terms of both objective and subjective image quality. For example, after a certain point (i.e., at about $d = 3$ for $B5$), for any further increase in the dead zone the subband bit number decreases very slowly, which makes further d increments unattractive. Clearly, this characteristic point at approximately $d = 3$ gives the best initial value in terms of quality and bit rate. Similar figures for all the other bands yielded the initial dead-zone figures summarized in Table II, where the number of bits used to encode white values of the active-quantizer domain is seen to be one for all the active subbands to achieve the low bit rate targeted.

Our final SBC codec performance indicates that for the "Miss America" sequence a peak SNR of about 38 dB is achieved. This can be associated with pleasant subjective communications quality at a bit rate of 55 kbit/s. Occasional coding impairments can be detected by rigorous scrutiny in the regions of intensive temporal activity, which is attributable to the 1% saturation clipping probability of the quantizer. Having portrayed the interframe prediction, QMF band splitting, scanning, quantization, and RLC, we now focus on our analytical model.

E. Analytical Subband Interframe Model

Let us assume that each pel of an image is undergoing linear translation and appears at other locations at other time instants. If $s(k, l, t)$ denotes the intensity of the pel at location (k, l) at time t and $s(k + dk, l + dl, t - T)$ is the pel intensity at the previous location at time $t - T$,

then the motion trajectory is given by

$$s(k,l,t) = s(k + dk, l + dl, t - T) \tag{5}$$

where dk and dl are the horizontal and vertical components of the motion trajectory during a frame period T.

As shown in Fig. 1, the input pel is first subtracted from the motion-compensated (MC) estimate to form the prediction-error signal. Thus

$$e(k,l,t) = s(k,l,t) - \hat{s}(k,l,t). \tag{6}$$

Assuming that the motion trajectory is estimated piecewise, the MC interframe estimate $s(k,l,t)$ is given by

$$\begin{aligned}\hat{s}(k,l,t) &= s\left[k + \overline{dk}, l + \overline{dl}, t - T\right] \\ &= s\left[k - \left(dk - \overline{dk}\right), l - \left(dl - \overline{dl}\right), t\right]\end{aligned} \tag{7}$$

where $\overline{dk}$ and $\overline{dl}$ are the components of the motion-displacement estimate. Thus, $(dk - \overline{dk})$ and $(dl - \overline{dl})$ can be considered as the components of the motion-estimation error. In the above equation, the quantization noise added to the previous coding frame has been ignored for the sake of simplicity.

From (6) and (7) and ignoring the quantization noise we have

$$e(k,l,t) = s(k,l,t) - s[k - (dk - \overline{dk}), l(dl - \overline{dl}), t]. \tag{8}$$

Taking the Fourier transform of (8) in the spatial coordinate

$$E(\omega_0, \omega_1, t) = S(\omega_0, \omega_1, t) \cdot [1 - e^{j(\omega_0 \bar{t}_k + \omega_1 \bar{t}_l)}] \tag{9}$$

where

$$\begin{aligned}\bar{t}_k &= (dk - \overline{dk}) \cdot T_J \\ \bar{t}_l &= (dl - \overline{dl}) \cdot T_L\end{aligned} \tag{10}$$

and T_J and T_L correspond to pel and line periods, respectively.

Assuming uniform QMF band splitting inside the motion-compensated loop in Fig. 1, the MC error signal is split into $N = n^2$ bands, where n is the 1-D decimation factor and can assume values of the power of two, and we have

$$\begin{aligned}E_{ij}(\omega_0, \omega_1, t) = \frac{1}{N} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} H_{ij}\left(\frac{\omega_0 + 2x\pi}{n}, \frac{\omega_1 + 2y\pi}{n}\right) \\ \cdot E\left(\frac{\omega_0 + 2x\pi}{n}, \frac{\omega_1 + 2y\pi}{n}\right)\end{aligned} \tag{11}$$

where $H_{ij}(.,.)i, j = 0,\cdots, n-1$ is the Fourier transform of the separable 2-D QMF impulse response representing N-band uniform decomposition. The decomposed error bands $E_{ij}(.,.)$ are each quantized independently. Assuming that the quantization noise is purely additive, the quantized subband signals are of the form

$$L_{ij}(\omega_0, \omega_1, t) = E_{ij}(\omega_0, \omega_1, t) + Q_{ij}(\omega_0, \omega_1, t). \tag{12}$$

Substituting $E_{ij}(\omega_0, \omega_1, t)$ form (11) in (12) yields

$$\begin{aligned}L_{ij}(\omega_0, \omega_1, t) = \frac{1}{N} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} H_{ij}\left(\frac{\omega_0 + 2x\pi}{n}, \frac{\omega_1 + 2y\pi}{n}\right) \\ \cdot S\left(\frac{\omega_0 + 2x\pi}{n}, \frac{\omega_1 + 2y\pi}{n}\right) \\ \cdot [1 - e^{j(\omega_0 \bar{t}_{k/n} + \omega_1 \bar{t}_{l/n})}] \\ + Q_{ij}(\omega_0, \omega_1, t).\end{aligned} \tag{13}$$

The quantized bands are RL coded and then multiplexed into two sensitivity classes (i.e., subchannels). The first subchannel includes the lower frequency bands, i.e., L_{ij}, $i = 0$, $p - 1$ and $j = 0$, $q - 1$ where $p, q < n$. The second subchannel covers the remaining bands, i.e., $i = p$, $n - 1$ and $j = q, n - 1$. As will be discussed later, this arrangement has been made to transmit the coded video via two separate subchannels of the 16-level QAM, having two different integrities (see Section V).

In the reconstruction feedback loop the quantized subbands, after remapping, are sent to the interpolation QMF filters. The reconstructed output is given by

$$\begin{aligned}\bar{L}(\omega_0, \omega_1, t) = \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} S\left(\omega_0 + \frac{2x\pi}{n}, \omega_1 + \frac{2y\pi}{n}\right) \\ \cdot [1 - e^{j(\omega_0 \bar{t}_k/n + \omega_1 \bar{t}_l/n)}] \\ \times \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{i+j} H_{ij}(\omega_0, \omega_1,) \\ \cdot H_{ij}\left(\omega_0 + \frac{2x\pi}{n}, \omega_1 + \frac{2y\pi}{n}\right) \\ + N \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{i+j} \\ \cdot H_{ij}(\omega_0, \omega_1) Q_{ij}(n\omega_0, n\omega_1, t).\end{aligned} \tag{14}$$

For a perfect reconstruction-separable 2-D QMF, (14) can be simplified as

$$\begin{aligned}\bar{L}(\omega_0, \omega_1, t) = S(\omega_0, \omega_1, t)[1 - e^{j(\omega_0 \bar{t}_k + \omega_1 \bar{t}_l)}] \\ + N \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (-1)^{i+j} \\ \cdot H_{ij}(\omega_0, \omega_1) Q_{ij}(n\omega_0, n\omega_1, t).\end{aligned} \tag{15}$$

In the MC feedback loop the reconstructed signal is added to the MC estimate to obtain the decoded signal as

$$\hat{S}(\omega_0, \omega_1, t) = S(\omega_0, \omega_1, t) + N \sum_i \sum_j (-1)^{i\ j} \cdot H_{ij}(\omega_0, \omega_1) Q_{ij}(n\omega_0, n\omega_1, t). \quad (16)$$

This signal is the same as the decoded signal at the receiver side in the absence of channel noise.

III. 16-QAM Modem Scheme

Since the low-bit-rate SBC image codec transmits the motion-compensated difference between adjacent image frames, it is vital to protect the information by FEC to prevent catastrophic error precipitation. In pedestrian microcells of the emerging personal-communication networks (PCN), relatively high average channel SNR's associated with benign Rician fading are fairly typical, but in deep fades of the worst-case Rayleigh channel used in our experiments the deployment of an FEC scheme is clearly crucial. Fortunately, due to the favorable SNR (typically $\geq$ 25 dB) and high channel coherence bandwidth ($\geq$ 100 kHz) experienced in microcells, no channel equalization is necessary. Our positive experiences with subband-coded [1] and code-excited linear-predictive (CELP) coded speech [2], using 16-QAM in the PCN environment, prompted us to attempt the development of a 16-QAM system for image transmissions. In fact, the reduced signalling rate of the nonbinary modulation scheme is vital to ensure that the narrow-band channel conditions hold for the low-bit-rate image codec, which removes the need for a channel equalizer and hence maintains low complexity.

A variety of QAM schemes has been analyzed in [15]. In a companion paper [1] we have shown that the first two bits of a 4-bit QAM symbol have dramatically different BER's from that of the second pair of bits, which we refer to as class 1 (C1) and class 2 (C2) subchannels, respectively. The analytical formulas for these BER's, given an average SNR value Γ, over a Raleigh-distributed fading channel are given by [1]

$$P_{C1}(\Gamma) = \frac{1}{2\Gamma} \int_0^\infty \left[Q\left(\sqrt{\frac{\gamma}{5}}\right) + Q\left(3\sqrt{\frac{\gamma}{5}}\right) e^{-\gamma/\Gamma} \right] d\gamma \quad (17)$$

$$P_{C2}(\Gamma) = \frac{1}{2\Gamma} \int_0^\infty \left[Q\left(\left(\frac{2\bar{\alpha}}{\alpha} - 1\right)\sqrt{\frac{\gamma}{5}}\right) + Q\left(\left(\frac{3\bar{\alpha}}{\alpha} - 2\right)\sqrt{\frac{\gamma}{5}}\right) e^{-\gamma/\Gamma} \right] d\gamma \quad (18)$$

where α represents the Rayleigh-distributed fading envelope, and $\bar{\alpha}$ is the expected value of α. When evaluated for practical SNR values encountered in the PCN environment, the C1 BER turns out to be sufficiently low to be further reduced by FEC techniques, while the C2 BER is excessively high for image transmission. Therefore, some grade of fade tracking ability has to be incorporated into the system. For the subband-coded speech transmissions proposed in [1], an average-locking automatic gain control (AGC) was deployed due to its adequate performance for randomly distributed input data statistics and low complexity. Furthermore, no side information was required for its operation. Unfortunately, for subband and run-length coded images, the source statistics are not random, and hence the AGC performs worse than for speech signals, although the transmission integrity required is higher. Bit scrambling would improve the average-locking AGC's performance, but better results are achieved by using the pilot-symbol-assisted [13] fade-tracking AGC developed for the 64-QAM speech-transmission system proposed in [2].

The pilot-symbol-assisted fade-tracking AGC transmits a known channel sounding phasor at regular periods (P) to estimate the gain factor required to compensate for the fading signal envelope, which is then linearly interpolated between two adjacent sounding phasors. As expected, the AGC performance depends on the length of the sounding period P, as well as on the fading rate. We simulated the 16-QAM modem using the pilot-assisted fade-tracking AGC, and evaluated the C1 and C2 subchannel BER performance for a range of P values using no diversity as well as second-order switched diversity. Our results are depicted in Figs. 10 and 11 for the C1 and C2 subchannels, respectively, using P = 10, 20, and 40 over a pedestrian Rayleigh-fading channel sampled at 22 ksamples/s, which is the signalling rate of our FEC-coded image source. The wave-propagation frequency is 1.8 GHz, the frequency of the emerging PCN, and the pedestrian mobile speed is 4 mi/h. Note that, at this high propagation frequency, the deployment of switched diversity becomes realistic and is used in conjunction with second-order diversity in, for example, the Japanese mobile-radio system [12].

Observe from Fig. 10 that, in the case of no diversity, the C1 BER performance is nearly independent of the sounding period P, reading 1% at SNR = 20 dB and 0.1% at SNR = 30 dB. When deploying second-order diversity, the channel sounding period P plays a more crucial role in controlling the C1 BER, although the improvement from P = 20 to P = 10 is lower than between P = 40 and P = 20. At P = 10 and SNR = 20 dB there is an order-of-magnitude C1 BER reduction to 0.1%, which is further improved to one and a half orders of BER reduction at 30 dB when compared to the no diversity scenario. For the C2 BER, the sounding period P plays a more prominent role in combating the channel fading, both with and without diversity. As seen in Fig. 11 for P = 10 and SNR = 20 dB, similar to the C1 subchannel, a BER improvement of an order of magnitude is yielded when second-order diversity is utilized. As the BER requirements of our SBC video codec are quite stringent, we opt for P = 10 and second-order diveristy in our prototype systems. It is important to note that for this

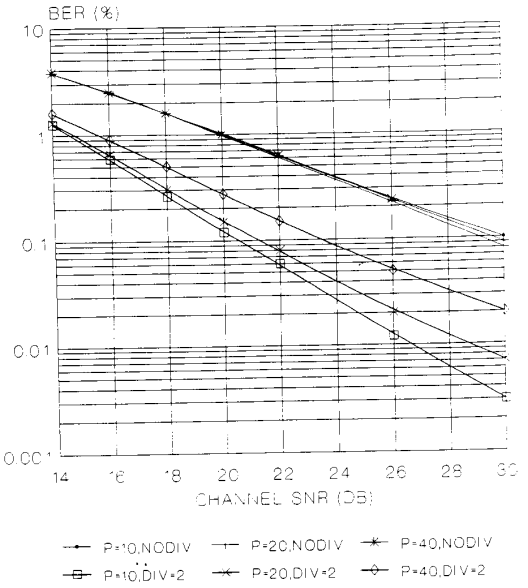


Fig. 10. Unprotected 16-QAM C1 BER versus channel SNR.

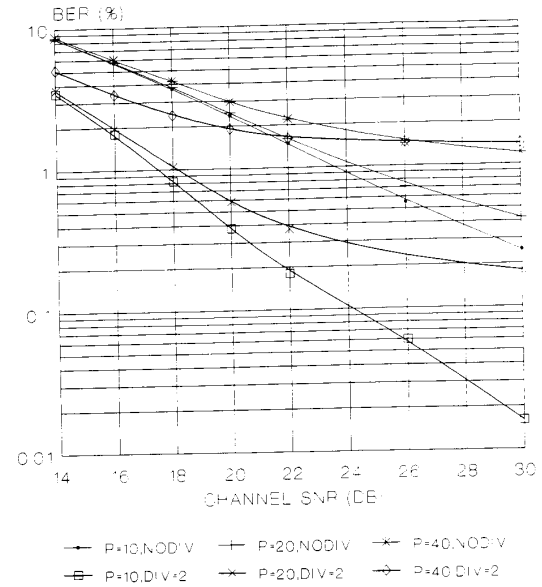


Fig. 11. Unprotected 16-QAM C2 BER versus channel SNR.

specific 16-QAM modem and AGC, the BER performance ramifications of doubling the vehicular speed is similar to doubling the channel sounding period P and hence can be readily studied in Figs. 10 and 11. However, when the FEC scheme's interplay is taken into account as well, increasing the mobile speed appears less detrimental to system performance due to a more random error distribution enjoyed by the interleaving/FEC arrangement.

IV. Selection of FEC Codes

Over bursty mobile channels, both convolutional and block codes can successfully be deployed with the proviso of appropriate code design [4]. However, convolutional codes bring the full benefit of maximum likelihood Viterbi decoding only with the prerequisite of soft decisions in the demodulation process. Unfortunately, the decoding complexity increases exponentially, both with the number of modulation states and with the code constraint length, which becomes prohibitive for the hand-held multimedia-communications adapters to be deployed in PCN's when using multilevel QAM constellations. Furthermore, they are incapable of detecting, whenever the decoder diverges from the error-free trellis path, thus yielding long strings of decoding errors before re-emerging with the error-free path. This is a severe disadvantage as compared to block codes, which have reliable error detections with a host of useful applications.

In contrast, block codes are powerful without soft-decision decoding. The class of nonbinary BCH codes, called Reed–Solomon (RS) codes, exhibits maximum minimum-distance properties, associated with reliable error-detection capability, which becomes extremely useful in controlling handovers or invoking image-post-processing algorithms. The error-correcting power of the optimum but complex RS codes is reasonably matched by that of computationally less-complex binary BCH codes, accompanied by appropriate interleavers, to transform the bursty error statistics of the Rayleigh channel into Gaussian-like statistics required by the shorter BCH codes. The family of 63-bit-long binary BCH codes with t = 2-, 3-, 4-, and 5-bit error-correcting capabilities constitutes a good compromise for our embedded SBC/BCH/16-QAM image system in terms of error-correcting power and complexity, and we opt for their deployment.

The error-protected C1 and C2 QAM superchannel BER performances versus channel SNR over the 1.8-GHz Rayleigh-fading channel are depicted in Figs. 12 and 13 using the BCH(63, 51, 2), BCH(63, 45, 3), BCH(63, 39, 4), and BCH(63, 36, 5) codes, as well as second-order diversity. With these BER curves in mind, we can now select the appropriate FEC codes for our image-transmission scheme.

V. Embedded SBC/BCH/16-QAM Schemes

Using the SBC image codec, BCH FEC codecs, and 16-QAM modem schemes described, a variety of embedded image-transmission systems can be contrived [14], which share the structure depicted in the block diagram in Fig. 14. Our basic philosophy is, rather than try to equalize the differing BER's of the C1 and C2 subchannels using FEC codes with matching correcting powers, we actually exploit their differences to reduce the overall system complexity. Clearly, we divide the SBC-coded image stream in the SBC MAP block into two substreams having different integrity requirements. While a full-scale bit-sensitivity analysis is feasible, for example, for speech

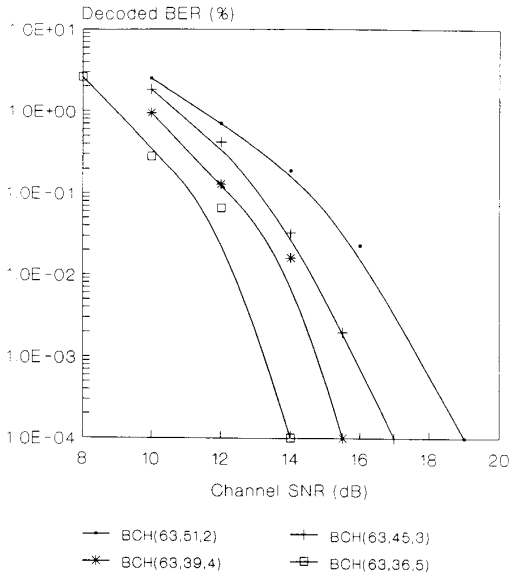


Fig. 12. FEC-protected 16-QAM C1 BER versus channel SNR.

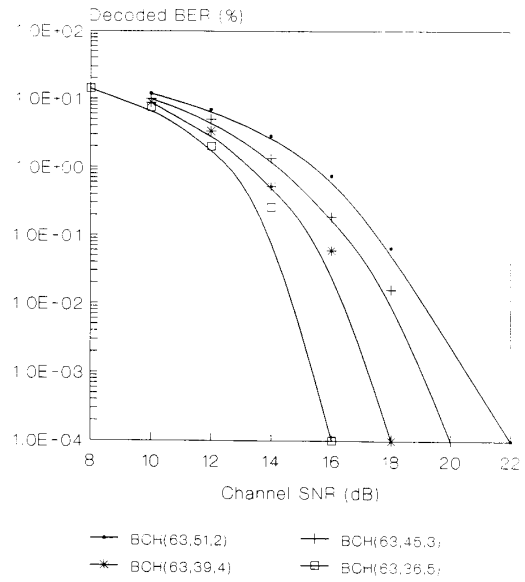


Fig. 13. FEC-protected 16-QAM C2 BER versus channel SNR.

source codecs [1], [2], where the frame length usually does not exceed 200 bits, in our 5481-bit-long image frames this is neither practical nor necessary. An appropriate twin-class bit-sensitivity classification protects the vitally important motion vector's bits and some of the sensitive low-frequency bands from channel errors by transmitting them via a higher integrity superchannel rather than the less sensitive high-frequency bands. Therefore, we arrange for the more significant bits (MSB's) to travel via the better FEC-protected superchannel, and the least significant bits (LSB's) to be transmitted through the worse

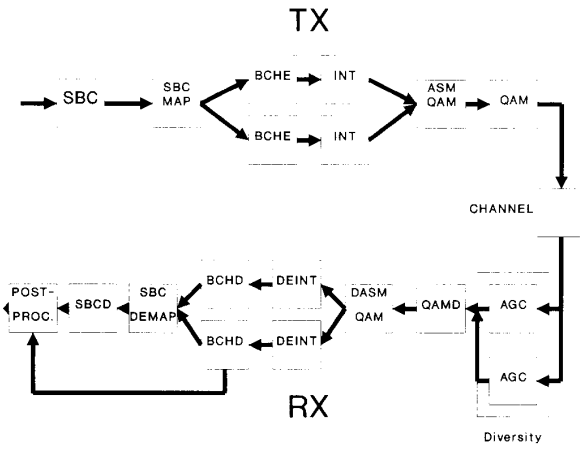


Fig. 14. SBC/BCH/QAM video-system schematic.

superchannel. Due to the high source-message-integrity requirement, we propose to utilize second-order diversity, since it reduces the BER's of both subchannels by an order of magnitude at an SNR of 20 dB when using a channel sounding period of $P = 10$. The quality (BER) of the subchannels is controlled by the careful selection of the BCH codes used. The blocks BCHE and INT are optional for the C1 subchannel, since its BER might be sufficiently low for the transmission of the SBC LSB's constituted by the high-frequency bands if $P = 10$ and second-order diversity is used, as evidenced by Fig. 10. The higher C2 BER always has to be lowered by some grade of error protection. The appropriate FEC-coded and interleaved C1 and C2 image streams are assembled for transmission in the block ASM QAM and mapped onto the QAM subchannels. The QAM signal is passed through the nondispersive microcellular Rayleigh-fading channel, and in the receiver, diversity reception follows, which necessitates two separate fade-tracking AGC's. After QAM demodulation in QAMD and bit disassembling in DASM QAM, the subchannels are deinterleaved (DEINT), BCH decoded (BCHD), and demapped to their original positions in the SBC image frame. At this stage the error-detecting capability of the C1 BCH decoder can be exploited to activate an image postprocessing algorithm whenever the frame is deemed to be corrupted.

For the target-channel SNR of about 20 dB, the unprotected image LSB's transmitted via the better C1 QAM subchannel have a BER of around 0.1%, as seen in Fig. 10. For the transmission of the sensitive-image MSB's we create a thoroughly FEC-protected superchannel by deploying the powerful BCH(63, 36, 5) code in the worse C2 QAM subchannel. This then, combined with second-order diversity, effectively removes all residual channel errors in this subchannel at the target SNR of 20 dB.

When using error-correction coding, the unprotected C1 QAM subchannel has a higher effective channel capacity than the C2 subchannel, because in the C2 subchannel, 27 bits in each 63-bits BCH frame are redun-

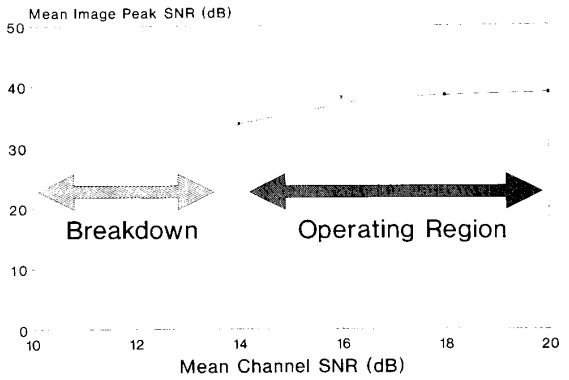


Fig. 15. Overall SBC/BCH/QAM video SNR performance versus channel SNR via Rayleigh channel.

dant parity bits. Therefore, the image bits generated have to be subdivided into more significant bits (MSB) and less significant bits (LSB), so that the number of bits after BCH encoding traveling via the C2 subchannel becomes equal to that of the unprotected C1 subchannel. This reduces the number of more significant bits transmitted via the protected C2 subchannel, relegating some of them to the higher BER unprotected C1 subchannel. The received images become unimpaired for average-channel SNR values in excess of 20 dB via Rayleigh-fading channels. However, for lower values the unprotected-image LSB's in the C1 subchannel have a detrimental effect on the received images, although the protected C2 bits are error free down to average-channel SNR values of 16–18 dB, as seen in Fig. 13.

This discovery prompted us to also utilize some grade of error correction in the C1 subchannel, where deploying the BCH(63, 51, 2) code removes most of the channel errors from the group of image LSB's as well, as shown in Fig. 12. When using the BCH(63, 51, 2) code in the C1 subchannel, the decoded BER is higher than that of the BCH(63, 36, 5)-coded—originally poorer—C2 QAM subchannel but sufficiently robust for the transmission of the image LSB's, as demonstrated by Figs. 12 and 13. The maximum total number of uncoded bits per frame in our final proposed system is 5481 yielding, incidentally, 63 BCH codewords in both subchannels, which corresponds to a total of 2268 image MSB's to be BCH(63, 36, 5) coded and 3213 image LSB's to be BCH(63, 51, 2) coded. The number of FEC-coded bits in both superchannels becomes 3969, yielding a total of 7938 FEC-protected image bits. Therefore, the overall bit rate is 79.380 kbit/s or approximately 20-kBd. After adding the channel sounding capacity, the baud rate becomes 22-kBd, a value readily accommodated within the coherence bandwidth of microcellular PCN systems.

The overall SBC/BCH/QAM system performance is shown in Fig. 15, giving practically error-free performance down to an average-channel SNR value of 16–18 dB, where communication rapidly breaks down with the image SNR becoming an unreliable quality measure. The mean image peak SNR was computed as the average of the peak image SNR using the "Salesman" and "Miss America" sequences, but was plotted only for the error-free operating channel SNR > 16–18 dB region of our system, as symbolized in the figure, below which the image quality becomes unacceptable. This underlines the need for an efficient and rapid handover algorithm that can be conveniently controlled by the BCH(63, 36, 5) codec's error-detection capability. When system complexity is at a premium, slightly less robust performance is achieved at a lower bandwidth requirement using no FEC in the C1 QAM subchannel. When channel capacity, (bit rate) is at a premium, using the so-called quarter-common intermediate format (QCIF), that is, 176 × 144 pels images, the overall bit rate is further reduced by a factor of 3.2, yielding a 7-kBd system. Alternatively, QCIF representation can be used at 30 frames/s and 22-kBd to achieve fast-motion tracking.

VI. Summary

Our investigations using a 55-kbit/s subband image codec, binary BCH error-correction codecs, and a 16-QAM modem showing that robust moving-image transmissions via microcellular mobile-radio channels are feasible. Monochrome test-image sequences of 320 × 256 pels were coded at a rate of 10 frames/s. Using two different BCH codes in the 16-QAM subchannels results in a moderately complex 22-kBd system, allowing error-free image transmissions at average-channel SNR's in excess of 16–18 dB. Extreme bandwidth efficiency is achieved using QCIF representations at 7-kBd and similar robustness.

References

[1] L. Hanzo, R. Steele, and P. M. Fortune, "A subband coding, BCH coding and 16-QAM system for mobile radio speech communications," *IEEE Trans. Vehicular Tech.*, vol. 39, no. 4, pp. 327–339, Nov. 1990.

[2] L. Hanzo, R. Salami, R. Steele, and P. M. Fortune, "Transmission of digitally encoded speech at 1.2 kBd for PCN," *Proc. Inst. Elec. Eng.*, pt. I, vol. 139, no. 4, pp. 437–447, Aug. 1992.

[3] H. Gharavi, "Subband coding algorithms for video applications: Videophone to HDTV-conferencing, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 2, pp. 174–183, June 1991.

[4] K. H. H. Wong, L. Hanzo, and R. Steele, "Channel coding for satellite mobile channels," *Int. J. Sat. Commun.*, vol. 7, pp. 143–163, 1989.

[5] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1278–1288, Oct. 1986.

[6] H. Gharavi and A. Tabatabai, "Subband coding of digital images using two-dimensional quadrature mirror filtering," in *Proc. SPIE*, vol. 707, Sept. 1986.

[7] ——, "Subband coding of monochrome and color images," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 207–214, Feb. 1988.

[8] H. Gharavi, "Differential subband coding of video signals," in *Proc. ICASSP'89*, May 1989, pp. 1819–1821.

[9] D. Esteban and C. Galand, "Application of quadrature mirror filters to splitband voice coding schemes," in *Proc. ICASSP*, 1977, pp. 191–195.

[10] J. R. Jain and A. K. Jain, "Displacement measurement and its applications in interframe image coding," *IEEE Trans. Commun.*, vol. COM-29, no. 12, Dec. 1981.

[11] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks, in *Proc. ICASSP'80*, Apr. 1980, pp. 291–294.

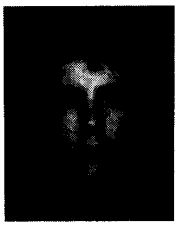
[12] M. Ikura, K. Ohno, Y. Yamao, and F. Adachi, "Field experiments on TDMA mobile radio signal transmissions," in *Proc. of IEEE-VTC'91*, (St. Louis, MO, May 19–22, 1991), pp. 669–674.

[13] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Vehicular Tech.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

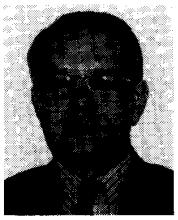
[14] R. Stedman, R. Steele, H. Gharavi, and L. Hanzo, "A 22 kBd mobile video telephone scheme," in *Proc. IEEE VTC'92* (Denver, May 1992), pp. 251–254.

[15] W. Webb, L. Hanzo, and R. Steele, "Bandwidth efficient QAM schemes for rayleigh fading channels," *Proc. Inst. Elec. Eng.*, vol. 138, pt. I, no. 3, pp. 169–175, June 1991.

[16] H. Gharavi, "Efficient multilevel coding scheme for video transmission," *Electron. Lett.*, vol. 28, no. 10, pp. 949–951, May 1992.



Robert Stedman received the B.Eng. degree in electronic engineering from the University of Southampton, Southampton, UK, in 1989. Currently he is working toward the Ph.D. degree in the field of low bit-rate picture coding.



Hamid Gharavi (M'79–SM'90–F'92) received the M.Sc. and Ph.D. degrees from Loughborough University of Technology, Loughborough, UK, in 1975 and 1979, respectively.

Before joining AT&T Bell Laboratories in 1982, he spent periods as a Research Fellow at Loughborough University and as a Lecturer at Auckland University, Auckland, New Zealand. He was an Adjunct Professor at the Polytechnic University (formerly Brooklyn Polytechnic), Brooklyn, New York. In 1984 he was transferred to Bellcore where he is now a Distinguished Member of the Research Staff. He has worked on problems related to digital modulation for satellite communications, bandwidth compression of color-television signals, low bit-rate video coding, high-definition television (HDTV), source coding, and pattern recognition. He holds several patents and has had more than 40 papers published. He was a core member of the CCITT Study Group XV (Specialist Group on Coding for Visual Telephony).

Dr. Gharavi has served as an Associate Editor of the IEEE Transactions on Circuits and Systems. During the period of 1989–1991, he was the Chairman of the CAS VLSI Systems and Applications Technical Committee. He received the 1986 Charles Babbage best paper award of the Institution of Electronics and Radio Engineers and the 1989 Darlington Best Paper Award of the IEEE Circuits and Systems Society.

Lajos Hanzo (SM'91) received the M.Sc. and Ph.D. degrees from the Technical University of Budapest, Budapest, Hungary, in 1976 and 1982, respectively.

From 1976 to 1986, he was with the Telecommunications Research Institute in Budapest, where his main research interests included high-speed data transmission, digital filtering, and modulation. From 1980 to 1981, he was on sabbatical leave working on parallel-DSP modems at the University of Erlangen–Nürnberg, Erlangen, Germany, and between 1981 and 1986 his research endeavors encompassed digital speech and data communications via satellites. In 1986 he joined the academic staff of the Department of Electronics and Computer Science at the University of Southampton where he is teaching one- and two-dimensional signal processing and telecommunications. His recent research interests have included the joint optimization of source (speech, image, etc.) and channel coding as well as modulation schemes for dispersive-fading mobile channels. He has authored and co-authored some 60 papers and contributed to *Mobile Radio Communications*, (London: Pentech Press, 1992). Dr. Hanzo has been awarded various distinctions.

Raymond Steele (SM'80) received the B.Sc. degree in electrical engineering from Durham University, Durham, UK, in 1959, and the Ph.D. and D.Sc. degrees from Loughborough University of Technology, Loughborough, UK, in 1975 and 1983, respectively.

Before receiving the B.Sc. degree, he was an indentured Apprentice Radio Engineer. After research and development posts with E. K. Cole, Cossor Radar and Electronics, and Marconi, he joined the lecturing staff at the Royal Naval College, London. He moved to Loughborough University in 1968 where he lectured and directed a research group in digital encoding of speech and picture signals. During the summers of 1975, 1977, and 1978, he was a consultant to the Acoustics Research Department at Bell Laboratories in the United States, and in 1979 joined the company's Communications Methods Research Department, Crawford Hill Laboratory, Holmdel, NJ. He returned to England in 1983 to become Professor of Communications in the Department of Electronics and Computer Science at the University of Southampton, Southampton, UK, a post he retains. From 1983 to 1986 he was a nonexecutive director of Plessey Research and Technology and from 1983 to 1989 was a consultant to British Telecom Research Laboratories. In 1986 he formed Multiple Access Communications Ltd., a company concerned with digital mobile-radio systems. He is the author of *Delta Modulation Systems* (New York: Halsted, 1975), editor of *Mobile Radio Communications* (London: Pentech Press, 1992), and author of over 150 technical publications. He is the editor of a series of books on digital mobile communications and a senior technical editor of the *IEEE Communications Magazine*.

Dr. Steele and his coauthors were awarded the Marconi Premium in 1979 and 1989, and the *Bell System Technical Journal* Best Mathematics, Communications, Techniques, Computing and Software and Social Sciences Paper in 1981. He has been a conference and session organizer of numerous international conferences and a keynote speaker at many international meetings.