A $22KBd$ Mobile Video Telephone Scheme

R.Stedman (1), R. Steele (1), H. Gharavi (2), L. Hanzo (1)

(1) Dept. of Electr., Univ. of Southampton, UK. S09 5NH
(2) Bellcore, Red Bank, New Jersey, USA.

Abstract

The complexity, image quality, bandwidth efficiency and robustness issues of a mobile video telephone scheme for personal communications networks (PCN) are addressed. Our motion-compensated $55Kbit/s$ subband codec (SBC) using seven non-uniformly spaced active bands with bandspecific scanning and runlength coding achieves image peak signal to noise ratios (SNR) of around $38dB$ associated with good communications quality for monochrome common intermediate format (CIF) images sampled at 10 frames per second. The motion dependent bitrate fluctuations are smoothed out by buffering with adaptive quantiser control feedback. The video bits are then sorted into two sensitivity classes and error protected by a twin-class binary Bose-Chaudhuri-Hocquenghem (BCH) scheme. The separately BCH-encoded more significant bits (MSB) and less significant bits (LSB) are transmitted via separate subchannels of the 16-level quadrature amplitude modulator (16-QAM), having different integrity. The overall signalling rate becomes $22KBd$, sufficiently low for typical microcells to fulfil the narrowband channel condition. Clearly, no equaliser has to be used, which considerably reduces the system's complexity, yet unimpaired image quality is achieved for channel SNRs in excess of $20dB$.

1 Introduction

The image and speech codecs of the emerging personal communications network (PCN) of the near future must guarantee good speech and image quality at high spectral efficiency, high robustness against channel errors and low complexity, which requires the joint optimisation of the source and channel codecs as well as modems [1], [2], [3]. High spectral efficiency is important not only to save precious bandwidth, but also to reduce signalling rate to render the channel non-dispersive, thereby removing the need for equalisation and hence reducing complexity and power consumption. In this spirit, our contribution is devoted to the design aspects of the attractive combination of a $55Kbit/s$ SBC image codec amalgamated with a matched pair of BCH error correction codecs and with a 16-QAM modem to yield a robust $22KBd$ mobile video system, yielding image peak SNRs around $38dB$ for channel SNRs in excess of $20dB$.

2 Subband Image Codec

Outline: The general outline of our motion-compensated SBC codec is seen in Figure 1 [3], [4]. The locally reconstructed previous frame is subtracted from the present source frame to be encoded to generate the prediction error frame along with an array of motion vectors, which have to be encoded and multiplexed for transmission in the block MV. Simultaneously, the error frame is split into seven non-uniform subbands $B0 \ldots B6$ by the two-dimensional (2D) quadrature mirror filter(QMF) [6]. Then the subjectively more important low-frequency subbands are allocated more bits than their less important high-frequency counterparts. The 2D subbands are then scanned in the 'SCAN' block to a one-dimensional (1D) stream for encoding. After quantisation (Q) of the 1D-scanned signal run-length coding (RLC) with variable length codewords (VLCW) follows, where the subband scanning algorithms are adjusted to suit the individual subband features to minimise the bitrate by generating long runs. Unfortunately, in contrast to coding and transmission convenience the time-variant motion activity and the bandwidth-efficient VLCWs result in a variable bitrate, which can be smoothed out using buffering with adaptive feedback control. When the 'Buffer fullness', defined as the proportion of the 'Buffer' already filled with bits, is approaching overflow, the quantiser feedback modifies the quantiser parameters to generate less bits in subsequent coding steps to achieve the targeted 'buffer fullness' and constant bitrate.

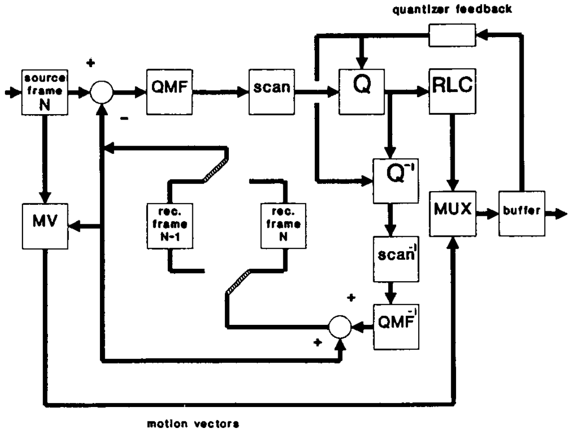

Figure 1: Image subband codec schematic

The SBC-coded motion-compensated prediction error is multiplexed with the motion vectors and conveyed to the forward error correction (FEC) channel codec and the 16-QAM modem. The quantised subband signals are locally decoded, scanned back into

0-7803-0673-2/92 $3.00 1992 IEEE

2D and the full-band motion error-frame is reconstructed in the QMF^{-1} block. This locally reconstructed error frame is added to the previous reconstructed image frame to generate the locally reconstructed cuurent frame to be utilised in the next motion-compensation step.

Interframe Prediction: The motion compensation algorithm used has to represent a good compromise in terms of computational complexity, encoded bitrate and low interframe prediction error. Reducing the block size utilised in our quest for minimum residual energy increases the transmission overhead for motion vectors, while extending the search area increases complexity but reduces the error energy and allows faster motion tracking. A good compromise is represented by a block size of $16 \cdot 16$ pels and a search area of $(\pm 7, \pm 7)$ pels. The search complexity of the optimum full search can be mitigated by Jain's near-optimum 'star-structured' algorithm [7], yielding only 10% higher error energy at the fifth complexity.

Band-splitting: Two-dimensional band-splitting is carried out using a 1D QMF [6], [8] along the rows and then along the columns, resulting in a basic 4-band decomposition. To achieve best compression ratio at highest image quality only the low horizontal and vertical spatial bands are further split to improve their resolution, yielding seven subbands. When the band has a low horizontal content, it must be scanned horizontally (H), while low vertical frequency bands must be scanned vertically (V) to achieve best compression after RLC. Should both the H and V contents be of high or low frequency, zig-zag (Z) scanning is used.

Quantisation: Close scrutiny of the subband probability density functions (PDF) reveals that although they are highly peaked around zero, this peak is indeed due to detrimental low-level camera noise. Clearly, the optimum Max-Lloyd quantiser therefore is unsuited for their quantisation. Our proposed quantiser characteristic [4] seen in Figure 2 hence exhibits a dead-zone (d) and a saturation threshold (t) to remove the camera-noise and clip low-probability extreme PDF peaks. The input values in the active range are then uniformly quantised to L levels. The total number of levels is hence $(L + 3)$.

The quantiser parameters $[d, t, L]_i$, $i = 0 \ldots 6$ have to be matched to the typical subband PDFs, but bearing in mind that the subsequent VLCW-based RLC produces a variable bitrate, output buffering with adaptive dead-zone control is deployed to smooth out bitrate fluctuations.

Run Length Coding: In terms of RLC we refer to zero-valued samples falling in the quantiser's dead-zone domain as 'black', while to non-zero quantised samples in the active quantiser domain as 'white'. Both the run colours and the run lengths are coded in one variable length code word. If a white run is coded then the actual nonzero quantized levels are transmitted as band-dependent fixed length binary codes immediately following the white run code. The runlength codewords are based on the so-called B1 codes [4], [3]. The run length decoding algorithm reads black and white run lengths from the input bit stream and the run lengths are used to reconstruct a scan line of the coded picture. After each line is coded, a specific end of line (EOL) code is inserted into the 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, allows the decoder to correct a single bit error per scanline by tentative bit-flipping and redecoding.

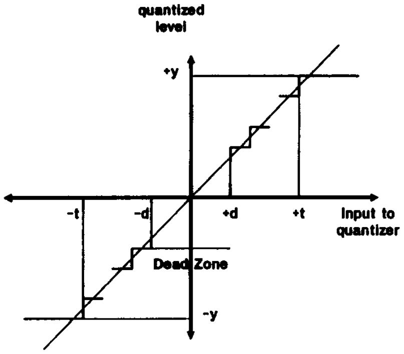

Figure 2: Quantiser characteristic

Adaptive Bitrate Control: To achieve a constant SBC bitrate, it is implementationally convenient to encode each subband into a constant number of bits and monitor their output rate on a line-by-line basis. The instantaneous bitrate of the specific line being encoded is compared against its targeted rate and its dead-zone d_i is appropriately increased, decreased or maintained. After coding each line of length N pels, the coder compares the actual number of coded bits produced against the 'ideal' number predicted by a linear model to arrive at a total of B_{Total} bits. Then a value for the excess number of coded bits B_{Excess} is computed and if B_{Excess} is positive then too many bits are being produced by the coder and hence the dead zone should be increased to lower the output bitrate for the next line, while d is decreased, if B_{Excess} is negative. The saturation clipping thresholds $t_i, i = 0 \ldots 6$ are kept constant for each subband to moderate codec complexity. The choice of $t_i, i = 0 \ldots 6$ is based on the 1% quantiles of the subband PDFs to minimise the effects of quantiser saturation.

i	t	d	L
0	6.0	4.4	1
1	5.0	4.7	1
2	4.8	3.3	1
3	4.8	2.1	1
4	4.8	2.0	1
5	4.8	1.6	1
6	5.8	2.6	1

Table 1: Quantizer initial values for each band

Clearly, the subband SNRs and hence the overall image SNR, as well as the average total bitrate can be conveniently controlled by the parameters $[d, t, L]_i$. After evaluating both the subband SNRs and the number of coded bits vs. $[d, t, L]_i$ for all subbands, finally we settled on the initial choice summarised in Table 1, whereby t_i and L_i were fixed to minimise complexity.

Performance: An average peak video SNR of about 38 dB is maintained by the SBC for a variety of image sequences, such as 'Miss America', yielding pleasant subjective communications quality at a bitrate of $55 Kbps$ or 0.067 bit/pel. Occassionally, 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 quantiser.

3 Modulation and Channel Coding

As expected, the high-compression SBC codec is fairly vulnerable to channel errors due to the interframe predictions and VLCW, resulting in catastrophic error propagation and annoying image artifacts. Therefore a robust, spectrally efficient modem would be required, but robust modems, such as Gaussian Minimum Shift Keying (GMSK) are not sufficiently bandwidth efficient for video transmissions to fulfil the narrowband channel condition. This then necessitates a channel equaliser to combat channel dispersion, which unfortunately contradicts the low equipment complexity and low power consumption requirement.

The key to this problem is the deployment of highly bandwidth efficient 16-QAM modulation, rendering the signalling rate sufficiently low for channel dispersions to become negligible without complex channel equalisers. Unfortunately, 16-QAM requires higher channel SNR, but this is ensured in pedestrian microcells of the emerging third generation personal communication networks (PCN), where typically SNR $\geq$ 25 dB, and the channel coherence bandwidth is $\geq$ 100 KHz.

In our earlier work [1] we have shown that the first two bits of a four bit QAM symbol have dramatically different BERs, from that of the second pair of bits, which we refer to as class 1 (C1), and class 2 (C2), subchannels respectively. We evaluated the C1 and C2 BER for a narrowband Rayleigh-fading channel using a propagation frequency of $1.8 GHz$, pedestrian speed of of $4 mph$, and a signalling rate of $22 KBd$. 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. This prompted us to use 'fade-tracking' automatic gain control (AGC) [2] and second order diversity in line with other prototype systems [9]. The AGC requires 10% transmission overhead and both arrangements slightly increase the system's complexity at dramatic BER improvement. However, the consistent C1 and C2 integrity discrepancy persists, as demonstrated by Figure 4 evaluated for the above conditions. This property will be advantageously exploited in our embedded error correction scheme to protect more important image bits more strongly.

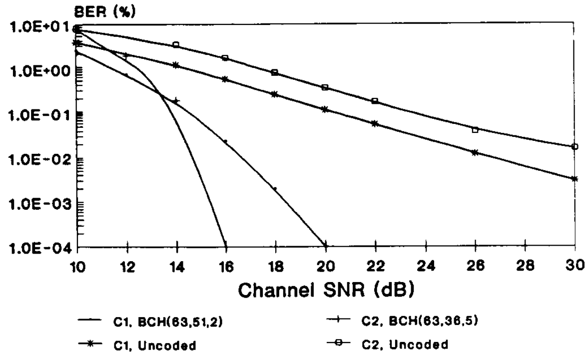

Figure 3: 16-QAM BER vs channel SNR with and without BCH coding using AGC and second-order diversity

Due to our choice of using 16-QAM the selection of robust error protection is quite crucial. High constraint-length convolutional codes are powerful, when combined with soft-decisions, but this becomes prohibitively complex for a 16-level modem. In contrast, low-complexity binary BCH block codecs are powerful with hard-decisions, as long as they are combined with interleavers of sufficient memory to break up long error bursts and render the channel near memoryless [5].

4 Amalgamated Video Schemes

In contriving attractive video schemes our basic philosophy is rather than trying to equalize the differing BERs 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 shown in Figure 5 into two substreams having different integrity requirements. While a full-scale bit-sensitivity analysis is feasible for example for speech source codecs [1], [2], where the frame length usually does not exceed 200 bits, in our 5481 bits 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 subchannel than the less sensitive high-frequency bands. Therefore we arrange for the more significant bits (MSBs) to travel via the better FEC-protected subchannel, while for the least significant bits (LSBs) to be transmitted through the worse subchannel. The BER integrity of the subchannels is controlled by the careful selection of BCH codes used.

The blocks BCHE and INT are optional for the C1 sub-channel, since its BER might be sufficiently low for the transmission of the SBC LSBs constituted by the high-frequency bands, if AGC and second order diversity are used, as evidenced by Figure 4. The higher C2 BER always has to be lowered by some grade of error protection. The FEC coded and interleaved C1 and C2 image streams are assembled in the block ASM QAM for transmission and mapped onto the QAM subchannels. The QAM signal is passed through the non-dispersive microcellular Rayleigh-Fading channel and in the receiver diversity reception ensues, which necessitates two separate fade tracking AGCs. 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 post processing algorithm, whenever the frame is deemed to be corrupted.

For the target channel SNR of about 20 dB the unprotected image LSBs transmitted via the better C1 QAM subchannel using 'fade-tracking' AGC and second order diversity have a BER of around 0.1%. These sparse image LSB errors are non-catastrophic thanks to corrections by the help of the redundancy introduced with the EOL codes, and become rarely subjectively annoying. For the transmission of the sensitive image MSBs we create a thoroughly FEC-protected subchannel by deploying the powerful BCH(63,36,5) code in the worse C2 QAM subchannel. This code encodes 36 image bits to a 63-bit frame and can correct up to 5 errors, which combined with the AGC and second-order diversity effectively removes all residual channel errors in the C2 subchannel at the target SNR of 20 dB.

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

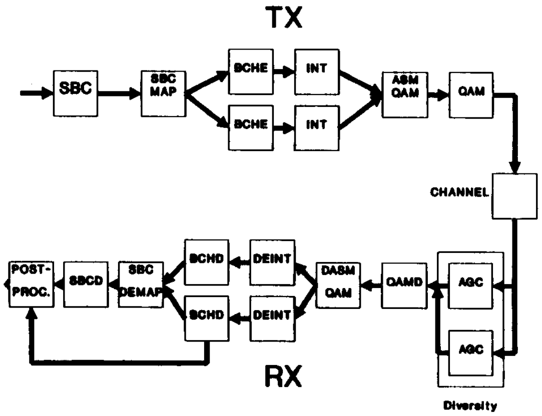

Figure 4: SBC/BCH/QAM video system schematic

bits BCH frame are redundant parity bits. Therefore the image bits have to be subdivided into MSBs and LSBs, in such a way that the number of bits after BCH encoding travelling 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 SNRs in excess of 20 dBs via Rayleigh-fading channels, but for lower values the unprotected image LSBs in the C1 subchannel have a detrimental subjective effect, although the protected C2 bits are error free down to channel SNRs of 16 – 18 *dB*, as seen in Figure 4.

This discovery prompted us to utilise some grade of error correction also in the C1 subchannel, where using the BCH(63,51,2) code removes most of the channel errors from the group of image LSBs as well, although it remains poorer than the BCH(63,36,5)-coded C2 subchannel, as demonstrated by the corresponding BER curves of Figure 4. 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 MSBs to be BCH(63,36,5) coded and 3213 image LSBs to be BCH(63,51,2) coded. The number of FEC-coded bits in both subchannels becomes 3969, yielding a total of 7938 FEC-protected image bits. Therefore the overall bitrate is 79.380 Kbit/s or approximately 20 KBaud. After adding the 10% channel sounding overhead, the signalling rate becomes 22 KBaud, a value, readily accommodated within the coherence bandwidth of microcellular PCN systems.

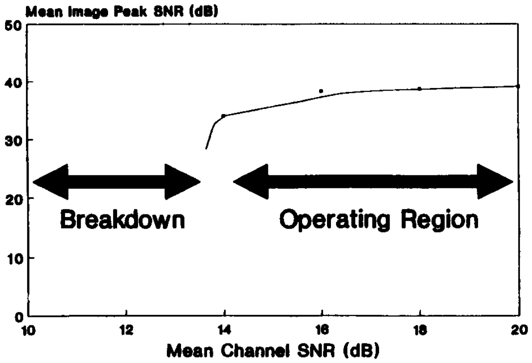

Figure 5: Overall SBC/BCH/QAM video SNR performance vs channel SNR via Rayleigh channel

The overall SBC/BCH/QAM system performance is shown in Figure 6 giving practically error-free performance down to an average channel SNR value of 16 – 18 *dB*, where communications rapidly breaks down with the image SNR becoming an unreliable quality measure. This underlines the need for an efficient and rapid handover algorithm that can be conveniently controlled by the BCH(63,36,5) codec's reliable error detection capability. When system complexity is at premium, slightly less robust performance is achieved at a lower bandwidth requirement using no FEC in the C1 QAM subchannel. When channel capacity, i.e. bitrate is at premium, using the so-called quarter common intermediate format (QCIF), i.e. $176 \cdot 144$ pels images, the overall bitrate is further reduced by a factor of 3.2, yielding a 7 *KBd* system. Alternatively, QCIF representation can be used at 30 *frames/sec* and 22 *KBd* to achieve fast motion tracking.

5 Summary

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

References

[1] L. Hanzo, R. Steele, P.M. Fortune: A Subband Coding, BCH Coding and 16-QAM System for Mobile Radio Speech Communications, IEEE Tr. on VT, Vol. 39, No4, Nov. 1990, pp 327-339

[2] L. Hanzo, R. Salami, R. Steele, P.M. Fortune: Transmission of Digitally Encoded Speech at 1.2 KBd for PCN, accepted by IEE Proc. Part-I, 1991

[3] R. Stedman, H. Gharavi, L. Hanzo, R. Steele: Transmission of Subband Coded Images via Rayleigh-fading Channels, submitted to IEEE Video Technology, 1992

[4] H. Gharavi: Subband Coding of Video Signals, Chapter in 'Subband Image Coding', Ed. J.W. Woods, Kluwer Academic Publishers, 1990

[5] K.H.H. Wong, L. Hanzo, R. Steele:Channel Coding for Satellite Mobile Channels, Int. Journal of Satellite Communications, Vol 7, 1989, pp 143-163

[6] D. Esteban, C. Galand: Application of Quadrature Mirror Filters to Splitband Voice Coding Schemes, Proc. of ICASSP 1977, pp191-195

[7] J.R. Jain, A.K. Jain: Displacement Measurement and its Applications in Interframe Image Coding, IEEE Transactions on Comms., Vol Com-29, No 12, Dec 1981 pp

[8] J.D. Johnston: A filter family designed for use in quadrature mirror filter banks, Proc. ICASSP'80, April 1980, pp 291-294

[9] M. Ikura, K. Ohno, Y. Yamao, F. Adachi: Field Experiments on TDMA Mobile Radio Signal Transmissions, Proc. of IEEE-VTC'91, St. Louis, USA, 19-22 May, 1991, pp 669-674