A NOVEL CODING SCHEME FOR THE TRANSMISSION OF ON-LINE HANDWRITING

H. Yuen, L. Hanzo

Dept. of Electr. and Comp. Sc., Univ. of Southampton, SO17 1BJ, UK
Tel.: +44 1703 593 125, Fax: +44 1703 593 045
Email:lh@ecs.soton.ac.uk, http://www.ecs.soton.ac.uk

ABSTRACT

The proposed adaptive fixed-length differential chain coding (FL-DCC) scheme can be configured as a lower rate, lower resolution or higher rate, higher resolution source codec. The programable-rate FL-DCC codec is embedded in an adaptive transceiver invoking more bandwidth efficient 16-level Quadrature Amplitude Modulation (16QAM) over benign microcellular channels, while employing more robust 4QAM in more hostile propagation scenarios. The minimum required channel signal to noise ratio (SNR) over AWGN and Rayleigh channels in its 16QAM mode is about 11 and 15 dB, respectively and the low-rate graphical stream can be accommodated during silent speech spurts.

1. INTRODUCTION

This treatise is concerned with a channel-adaptive re-configurable graphical communications scheme, which is embedded in a voice, video multimedia system context. The advantage of the fixed-length differential chain coding (FL-DCC) scheme proposed is that the number of bits used to quantise the differential vectors of the so-called coding ring is dynamically adjusted under system's control, in order to match the prevailing bitrate, graphical quality and/or channel quality constraints. Although during low bitrate operation lossy graphical quantisation is used, for a small coding ring size of $M = 8$ near-unimpaired subjective quality is maintained, which is similar to that of differential chain coding (DCC) [1]-[4].

Section 2 introduces the lossy FL-DCC concept, Section 3 suggests a technique for improving its robustness against channel errors, while the system's architecture is briefly discussed in Section 4, before characterising the system's performance in Section 5.

2. FIXED LENGTH DIFFERENTIAL CHAIN CODING

In chain coding (CC) a square-shaped coding ring is slid along the graphical trace from the current pixel, which is the origin of the legitimate motion vectors, in steps represented by the vectors portrayed in Figure 1. The bold dots in the Figure represent the next legitimate pixels during the graphical trace's evolution. In principle the graphical trace can evolve to any of the surrounding eight pixels and hence a three-bit codeword is required for lossless coding. Differential chain coding [1]-[4] (DCC) exploits that the most likely direction of stylus movement is a straight extension, corresponding to vector 0 and with a gradually reducing probability of sharp turns corresponding vectors having higher indeces. Explicitly, we have found the while vector 0 typically has a probability of around 0.5 for a range of graphical source signals, including English and Chinese handwriting, a Map and a technical Drawing, the relative frequency of vectors ± 1 is around 0.2, while vectors $\pm 2, \pm 3$ have probabilities around 0.05. This suggests that the coding efficiency can be improved using the principle of entropy coding by allocating shorter codewords to more likely transitions and longer ones to less likely transitions.

In this treatise we embarked on exploring the potential of a graphical coding scheme dispensing with variable length coding, which we refer to as fixed length differential chain coding (FL-DCC). FL-DCC was contrived in order to comply with the time-variant resolution- and/or bit rate constraints of intelligent adaptive multimode terminals, which can be re-configured under network control to satisfy the momentarily prevailing tele-traffic, robustness, quality, etc system requirements. In order to maintain lossless graphics quality under lightly loaded traffic conditions, the FL-DCC codec can operate at a rate of $b = 3$ bits/vector, although it has a higher bit rate than DCC. However, motivated by the success of perceptually unimpaired lossy voice and video coding, we embarked on exploring the potential of the reconfigurable FL-DCC codec under $b < 3$ low-rate, lossy conditions.

Based on our findings above mentioned as regards to the relative frequencies of the various differential vectors, we decided to evaluate the performance of the FL-DCC codec using the $b = 1$ and $b = 2$ bit/vector lossy schemes. As demonstrated by Figure 1, in the $b = 2$-bit mode the transitions to pixels -2, -3, +2, +3 are illegitimate, while vectors 0, +1, -1 and +4 are legitimate. In order to minimise the effects of transmission errors the Gray codes seen in Figure 1 were assigned. It will be demonstrated that, due to the low probability of occurance of the illegitimate vectors, the associated subjective coding impairment is minor. Under degrading channel conditions or higher tele-traffic load the FL-DCC coding rate has to be reduced to $b = 1$, in order to be able to invoke a less bandwidth efficient, but more robust modulation scheme or to generate less packets contending for transmission. In this case only vectors +1 and -1 of Figure 1 are legitimate. The subjective effects of

GLOBECOM'95, SINGAPORE

0-7803-2509-5/95 US$4.00 © 1995 IEEE

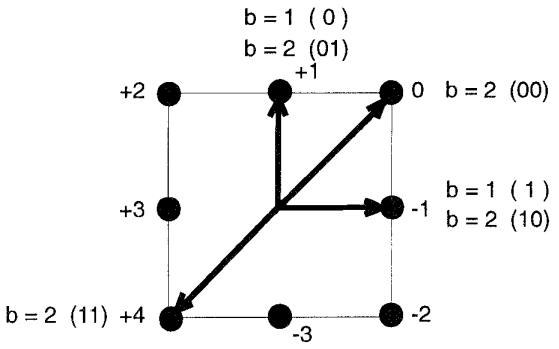

Figure 1: Coding ring

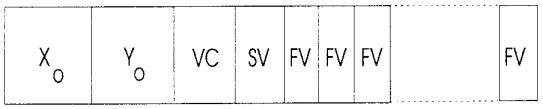

Figure 2: Coding Syntax

the associated zig-zag trace will be removed by the decoder, which can detect these characteristic patterns and replace them by a fitted straight line.

In general terms the size of the coding ring is given by $2n\tau$, where $n = 1, 2, 3 \ldots$ is referred to as the order of the ring and τ is a scaling parameter, characteristic of the pixel separation distance. Hence the ring shown in Figure 1 is a first order one. The number of nodes in the ring is $M = 8n$.

The data syntax of the FL-DCC scheme is displayed in Figure 2. The beginning of a trace can be marked by a typically 8 bit long pen-down (PD) code, while the end of trace by a pen-up (PU) code. In order to ensure that these codes are not emulated by the remaining data, if this would be incurred, bit stuffing must be invoked. We found that in complexity and robustness terms using a 'vector counter' (VC) constituted a more attractive alternative for our system. The starting coordinates X_0, Y_0 of a trace are directly encoded using for example 10 and 9 bits in case of a video grahics array (VGA) resolution of 640 × 480 pixels.

The first vector displacement along the trace is encoded by the best fitting vector defined by the coding ring as the starting vector (SV). The coding ring is then translated along this starting vector to determine the next vector. A differential approach is used for the encoding of all the following vectors along the trace, in that the differences in direction between the present vector and its predecessor are calculated and these vector differences are mapped into a set of 2^b fixed length b-bit codewords, which we refer to as 'fixed vectors' (FV). We will show that the coding rate of the proposed FL-DCC scheme is lower for $b = 2$ and $b = 1$ than that of DCC.

When a curve is encoded by FL-DCC, it is sliced by the coding ring into small segments. Consider a sampled curve segment s. Let v be the vector link produced by the coding ring. The coding rate of a chain code is defined [2],[3] in

	$b = 1$ bit/vector	$b = 2$ bit/vector	DCC bit/vector
English script	0.8535	1.7271	2.0216
Chinese script	0.8532	1.7554	2.0403
Map	0.8536	1.7365	2.0396
Drawing	0.8541	1.7911	1.9437
Theoretical	0.9	1.80	2.03

Table 1: Coding rate comparison

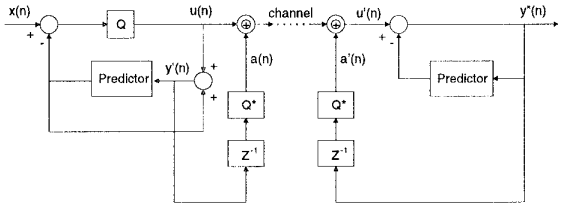

Figure 3: Robust FL-DCC scheme

bits per unit length of the curve segment as

$$r = \frac{E[b(s,v)]}{E[l_n(s)]} \tag{1}$$

where $b(s,v)$ is the number of bits used to encode a vector link v, $l_n(s)$ is the length of the curve segment s, while $E(x)$ represents the expected value of a random variable x. It has been shown [4] that for the set of all curves, the product of a segment length $l_n(s)$ and the probability $p(\alpha)$ that this segment occurs with a direction α must be constant. Thus the expected curve segment length for a ring of order n is given by [1]:

$$E[l_n(s)] = \int_0^{\pi/4} \frac{8 \cdot n \cdot \tau}{\cos\alpha} \cdot p(\alpha) d\alpha = \frac{\pi \cdot n \cdot \tau}{2 \cdot \sqrt{2}}. \tag{2}$$

Therefore, the theoretical coding rate of FL-DCC becomes:

$$r = \frac{E[b(s,v)]}{\pi \cdot n \cdot \tau / (2 \cdot \sqrt{2})}. \tag{3}$$

The theoretical and experimental coding rates of the $b = 1$ and $b = 2$ FL-DCC schemes are shown in Table 1.

3. ROBUST FL-DCC

It is well understood that the objective and subjective effects of transmission errors is different in memoryless coding schemes, such as for example Pulse Code Modulation (PCM), and in predictive schemes, such as Differential PCM (DPCM) [5]. Our proposed FL-DCC scheme is to a certain extent related to DPCM, since it uses the previous vector as a prediction of the current one. This similarity prompted us to invoke a hybrid PCM-DPCM-like technique suggested by Bull [5] in order to improve the robustness of the FL-DCC codec using the schematic of Figure 3.

Bull's hybrid scheme [5] is adopted in our FL-DCC scheme as follows. As it is demonstrated by the Figure, the b-bit FL-DCC codeword $u(n)$ is modified before transmission by modulo-2^b adding the quantity $a(n)$, which is derived from

the locally decoded graphical signal $y'(n)$ by delaying it and taking the b Most Significant Bits (MSBs) of it, which is carried out by the quantiser Q^*. Note that the modulo-2^b addition truncates the result of the operation, which introduces some grade of arithmetic inaccuracy, but ensures that the transmitted codeword is still represented by b bits. Under error-free conditions the encoder and decoder are perfectly 'aligned' with eachother, implying that $a(n) = a'(n)$, $u(n) = u'(n)$ and $y''(n) = y'(n)$, which naturally allows unimpaired graphical signal reconstruction, since the decoder carries out the inverse operations of the encoder.

In case of channel errors, however, the encoder and decoder become misaligned, but the proposed arrangement mitigates the propagation of transmission errors in the feedback loop, as we will highlight in our forthcoming discourse. Let us assume that the decoded graphical signals at the encoder and decoder, namely $y'(n)$ and $y''(n)$, respectively, now differ at the sampling instant n due to transmission errors by the non-zero error quantity

$$e(n) = y'(n) - y''(n). \tag{4}$$

Since the quantities $a(n)$ and $a'(n)$ are derived from $y'(n)$ and $y''(n)$ by coarse quantisation using Q^*, the difference $a(n+1) - a'(n+1)$ at instant $n+1$ consitutes a good measure of the error's influence at instant $n+1$, because the decoded signals are delayed by one sampling interval. Bull [5] showed that the error's effect can be substantially mitigated at sampling instant $n+1$, which will be demonstrated below, for the sake conceptual simplicity assuming infinite precision addition, rather than modulo-2^b truncation. From Equation 4 we have:

$$y''(n) = y'(n) - e(n) \tag{5}$$

and by the help of the decoder's inner predictor loop of Figure 3 we can proceed to sampling instant $n+1$ to arrive at:

$$y''(n+1) = u'(n+1) + y''(n) \tag{6}$$

and substituting $y''(n)$ from Equation 5 into Equation 6 yields:

$$y''(n+1) = u'(n+1) + y'(n) - e(n). \tag{7}$$

We argued above that the difference $a(n+1) - a'(n+1)$ is a good measure of the error's influence at instant $n+1$, hence $u'(n+1)$ in Equation 7 can be written as:

$$u'(n+1) \approx u(n+1) + a(n+1) - a'(n+1), \tag{8}$$

and substituting Equation 8 back in Equation 7 leads to:

$$\begin{aligned} y''(n+1) &= u(n+1) + a(n+1) - a'(n+1) \\ &\quad + y'(n) - e(n) \\ &\approx u(n+1) + y'(n) \\ &= y'(n+1), \end{aligned} \tag{9}$$

which implies that the error effects are approximately cancelled at instant $n+1$. Above we have assumed an infinite precision arithmetic, rather than modulo-2^b arithmetic, but in practice the accuracy of the error cancellation is degraded by the modulo-2^b truncation. Having characterised the proposed FL-DCC codec let us now focus our attention briefly on the associated transmission issues.

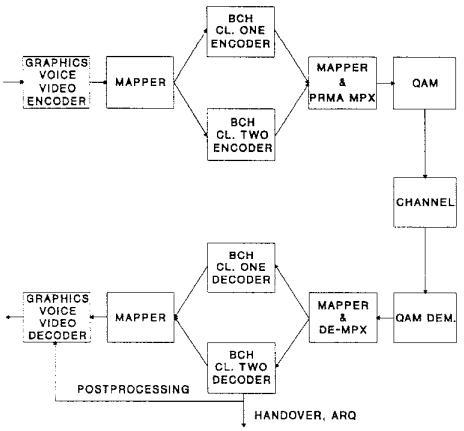

Figure 4: System's Schematic

4. TRANSMISSION SCHEME

The proposed graphical communications arrangement is embedded in a multimedia system, which is depicted in Figure 4. As seen in the Figure, the voice, video and graphics multimedia source streams are protected by a twin-class sensitivity-matched binary Bose-Chaudhuri-Hocquenghem (BCH) forward error correction (FEC) scheme. In our system the 8 kbps Code Excited Linear Predictive (CELP) CCITT speech codec is assumed [9] for the purposes of networking studies, but speech coding aspects are not considered here. On the same note, the 8 kbps videophone scheme of reference [8] is assumed, but the discussion of specific videophony aspects is beyond the scope of this treatise.

Then both the more vulnerable Class One and the more robust Class Two multimedia source bits are fed to the Packet Reservation Multiple Access (PRMA) [7] Multiplexer (MPX) and queued for transmission to the Base Station (BS). A unique feature of our PRMA multiplexer is that it allocates regularly spaced slots on a Time Division Multiple Access (TDMA) basis to interactive videophone users in order to maintain a low delay and lip-synchronisation, while speech and graphical users are served under the control of the Voice Activity Detector (VAD) on a demand basis. The VAD detects inactive speech spurts and multiplexes a number of additional speech users as well as graphical FL-DCC users onto the passive timeslots.

Depending on the channel conditions, Pilot Symbol Assisted (PSA) 4-level or 16-level Quadrature Amplitude Modulation (QAM) [7] was invoked, allowing for the system to increase the number of bits per symbol within the same bandwidth and hence to improve the voice, video or graphics communications quality under benign channel conditions [6]. The programable $b = 1 \ldots 3$-bit FL-DCC codec supports the operation of the adaptive transceiver. In the 16QAM mode the BCH(255,131,18) and BCH(255,91,25) codes, while in the 4AM mode the BCH(255,111,21) code is employed [6]. The receiver carries out the inverse functions and recovers the transmitted information. Should the communications quality unacceptably degrade, then the error detection capability of the stronger BCH code can be used to initiate a handover or Automatic Repeat Request (ARQ). Specific algorithmic issues of the transceiver, including BCH FEC

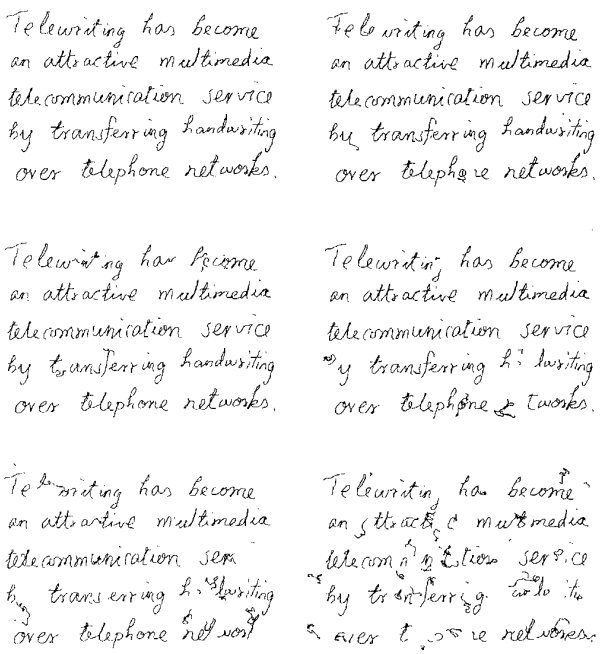

Figure 5: Subjective effects of transmission errors for the $b = 1$ 16QAM, RD, TX3 scheme for PSNR values of (left to right, top to bottom) 49.47, 42.57, 37.42, 32.01, 27.58 and 21.74 dB, respectively.

coding [6], PRMA multiplexing [7], adaptive QAM transmissions [7] etc are beyond the scope of this treatise, for further details the interested reader is referred to the references quoted.

5. RESULTS AND DISCUSSION

The performance of the proposed FL-DCC schemes was evaluated for a range of dynamographical source signals, including an English script, a Chinese script, a drawing and a map using a coding ring of M = 8. Table 1 shows the associated coding rates produced by FL-DCC for $b = 1$ and $b = 2$ as well as by DCC along with the corresponding theoretical coding rates. Both FL-DCC schemes achieve a lower coding rate than DCC. The corresponding subjective quality is portrayed in Figure 5 for $b = 1$ under both error-free and impaired channel conditions, when using the previously mentioned English Script of Table 1. At this stage let us concentrate on the quality of the top left segment only, corresponding to unimpaired channel conditions, where the FEC scheme managed to remove all transmission errors and the corresponding Peak Signal to Noise Ratio (PSNR) became 49.47 dB. The transceiver was configured for this particular set of graphical outputs to employ 16QAM and three re-transmission attempts (TX3) over Rayleigh channels using diversity (RD), in order to improve the system's robustness. The corrupted segments exhibiting reduced PSNR values in the Figure correspond to various channel conditions under the same transceiver configuration. Here we used the PSNR values, rather than the channel SNR values in order to differentiate amongst the different scenarios, since the portrayed subjective quality would be characteristic of a similar PSNR value, but a different channel SNR value in the 4QAM operating mode of the transceiver.

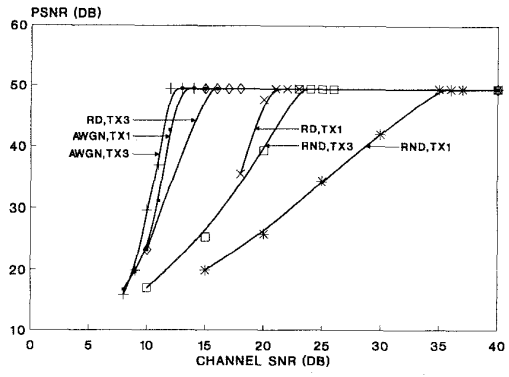

Figure 6: Graphical PSNR versus channel SNR performance of the $b = 1$-bit FL-DCC/16QAM mode of operation over various channels

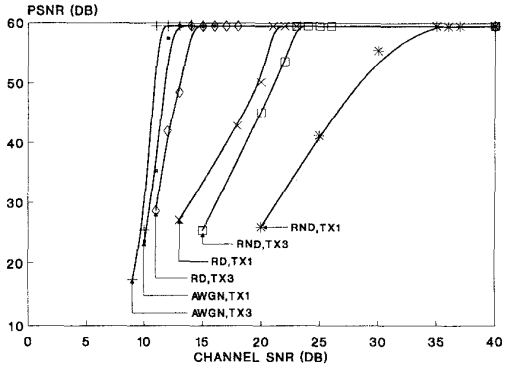

Figure 7: Graphical PSNR versus channel SNR performance of the $b = 2$-bit FL-DCC/16QAM mode of operation over various channels

Let us now focus our attention on the robustness aspects of the proposed graphical transceiver, when used over various wireless channels. Our experimental channel conditions were based on a worst-case Rayleigh-fading scenario using a propagation frequency of 1.9 GHz, signalling rate of 133 kBd and vehicular speed of 30 mph. The graphical representation quality was evaluated in terms of both the mean squared error (mse) and the Peak Signal to Noise Ratio (PSNR). In analogy to the PSNR in image processing, the graphical PSNR was defined as the ratio of the maximum possible spatial deviation 'energy' to the coding error 'energy'. When using a resolution of 640 × 480 pixels, the maximum spatial deviation energy is $640^2 + 480^2 = 800^2$. The reconstruction error energy was measured as the mean squared value of the pixel-to-pixel spatial distance between the original graphical input and the FL-DCC graphical output. For perfect channel conditions the $b = 1$ and $b = 2$ FL-DCC codec had PSNR values of 49.47 and 59.47 dB,

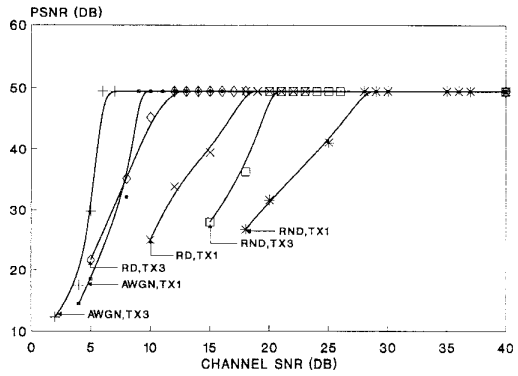

Figure 8: Graphical PSNR versus channel SNR performance of the b = 1-bit FL-DCC/4QAM mode of operation over various channels

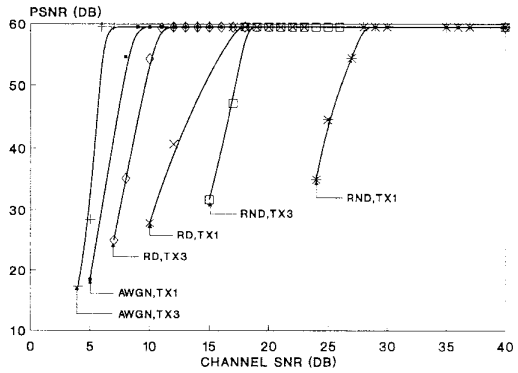

Figure 9: Graphical PSNR versus channel SNR performance of the b = 2-bit FL-DCC/4QAM mode of operation over various channels

respectively.

The system's PSNR versus channel SNR performance, when using the lower-rate b = 1 FL-DCC codec in conjunction with the 16QAM and 4QAM modes of operation is characterised by Figures 6 and 8, while the PSNR performance of the higher-rate b = 2 codec is portrayed in Figures 7 and 9, respectively. Observe that transmissions over Rayleigh channels with diversity (RD) and with no diversity (RND) are portrayed using both one (TX1) and three (TX3) transmission attempts, in order to improve the system's robustness. Similarly, over AWGN channels one or three transmission attempts were invoked. When using the lower-rate, lower-quality b = 1 mode, similar required channel SNR values apply, but the unimpaired PSNR is around 50 dB, ie some 10 dB lower.

6. SUMMARY AND CONCLUSIONS

An adaptive FL-DCC coding scheme was proposed for dynamographical communications, which has a lower coding rate and similar graphical quality to DCC in case of b = 1 and 2. The codec can be adaptively re-configured to operate at b = 1, b = 2 or even at b = 3 in order to comply with the prevailing network loading and/or propagation conditions, as well as graphical resolution requirements. The proposed FL-DCC codec was employed in an intelligent, re-configurable wireless adaptive multimedia communicator, which was able to support a mixture of speech, video and graphical users. In the more bandwidth efficient 16QAM mode of operation the 200 kHz system bandwidth accommodated a signalling rate of 133 kBd and allowed us to support nine multimedia users transmitting speech, video and graphical information. The multimedia user bandwidth became 200/9=22.2 kHz. The minimum required channel SNR in the ARQ-assisted 16QAM mode was 11 dB and 15 dB over AWGN and diversity-aided Rayleigh channels, respectively. The system's robustness was improved and its tele-traffic capacity dropped, when 4QAM was invoked.

7. REFERENCES

[1] K. Liu and R. Prasad, "Performance analysis of differential chain coding, " European Transaction on Telecommunications and Related Technology, vol. 3, pp. 323-330, Jul-Aug. 1992

[2] D.L. Neuhoff and K.G. Castor, "A rate and distortion analysis of chain codes for line drawings, " IEEE Tran. Information Theory, vol. IT-31, pp. 53-68, Jan. 1985.

[3] A. B. Johannessen, R. Prasad, N.B.J. Weyland and J.H. Bons, "Coding efficiency of multiring differential chain coding, " IEE Proceedings, Part I, vol. 139, pp. 224-232, April 1992.

[4] R. Prasad, J.W. Vivien, J.H. Bons, J.C. Arnbak, "Relative vector probabilities in differential chain coded line drawings", Proc. of IEEE Pacific Rim Conference on Comms., Computers and Signal Processing, pp 138-142, Victoria, Canada, June 1989

[5] M.C.W. Van Bull: Hybrid D-PCM, a combination of PCM and DPCM, IEEE Tr. on Comms., March 1978, Vol. Com-26, pp 362-368

[6] R. Steele (Ed.) Mobile Radio Communications, IEEE Press-Pentech Press, London, 1992

[7] W.T. Webb, L. Hanzo: Modern quadrature amplitude modulation: Principles and applications for fixed and wireless channels, IEEE Press-Pentech Press, 1994, ISBN 0-7273-1701-6, p 557

[8] J. Streit, L. Hanzo: Vector-quantised low-rate cordless videophone systems, submitted to IEEE VT, 1995

[9] R.A. Salami, C. Laflamme, J-P Adoul, D. Massaloux: A toll quality 8 Kb/s speech codec for the personal communications system (PCS), IEEE Tr. on Veh. Tech., Vol. 43, No. 3, Aug. 1994, pp 808-816

[10] W. Wong, D. Goodman: A packet reservation access protocol for integrated speech and data transmission, Proc. of the IEE, Part-I, Dec. 1992, Vol 139, No 6, pp 607-613.