WIRELESS MULTI-LEVEL GRAPHICAL CORRESPONDENCE

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

Indexing terms: Graphical communications, Image processing, Chain coding, Multimedia communications, Intelligent multimode terminals

ABSTRACT

An adaptive fixed-length differential chain coding (FL-DCC) scheme is proposed for the transmission of line graphics, which can be re-configured as a lower rate, lower resolution or higher rate, higher resolution source codec. The dual-rate re-configurable FL-DCC codec is embedded in a voice, video multimedia communicator system, which allocates the FL-DCC graphics packets to idle speech packets with the aid of the Packet Reservation Multiple Access (PRMA) multiplexer. In a bandwidth of 200 kHz, which is characteristic of the Pan-European GSM system, nine speech, video and graphics multimedia users can be accommodated using bandwidth efficient 16-level Quadrature Amplitude Modulation (16QAM) over benign microcellular channels. In the diversity-aided and automatic repeat request (ARQ) assisted 16QAM mode of operation the multimedia user bandwidth becomes 22.2 kHz and the minimum required channel signal to noise ratio (SNR) over AWGN and Rayleigh channels is about 11 and 15 dB, respectively.

1. INTRODUCTION

In this contribution an intelligent re-configurable graphical communications scheme is proposed, which is embedded in a voice, video multimedia system context. In the proposed fixed-length differential chain coding (FL-DCC) scheme 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 quantisation is used, we will demonstrate that in case of small coding rings near-unimpaired subjective quality is maintained, which is similar to that of differential chain coding (DCC) [1].

The paper is structured as follows. Section 2 highlights the system's architecture, leading to Section 3, which describes the proposed re-configurable graphical source codec. Transmission issues, including error control, modulation and packetisation aspects are discussed in Section 4. Section 5 is devoted to issues related to multiplexing graphical source signals with voice and video, before reporting on the system's performance in Section 6. Our conclusions are offered in Section 7.

ICUPC'95, 6-10 NOV. 1995, TOKYO, JAPAN

0-7803-2955-4/95/$4.00 ©IEEE

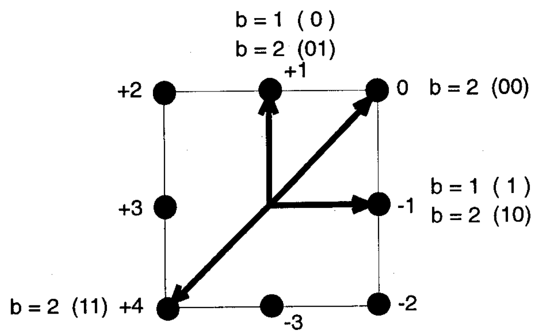

Figure 1: Coding ring

2. SYSTEM ARCHITECTURE

In the proposed transceiver the voice, video and graphics source encoders' bit streams are mapped in two sensitivity classes and sensitivity-matched binary Bose-Chaudhuri-Hocquenghem (BCH) forward error correction (FEC) coded [2]. Then both the more vulnerable Class One and the more robust Class Two source bits are fed to the Packet Reservation Multiple Access (PRMA) [3] Multiplexer (MPX) and queued for transmission to the Base Station (BS). Depending on the channel conditions, Pilot Symbol Assisted (PSA) 4-level or 16-level Quadrature Amplitude Modulation (QAM) [3] is 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 [2]. 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). In our system the 8 kbps Code Excited Linear Predictive (CELP) CCITT speech codec is assumed [5] 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 [4] is assumed, but the discussion of specific videophony aspects is beyond the scope of this treatise.

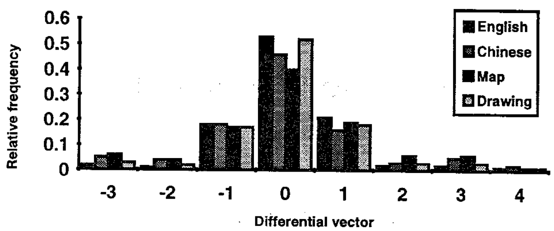

Figure 2: Relative frequency of differential vectors for a range of dynamographical source signals

3. 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] (DCC) exploits that the most likely direction of stylus movement is a straight extension, with a diminishing chance of 180^o turns. 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. This argument is supported by the histogram of the differential vectors of a range of graphical source signals, including English and Chinese handwriting, a Map and a technical Drawing, portrayed in Figure 2, where vectors 0, +1 and -1 are seen to have the highest relative frequency.

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 chan 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, since in voice and video coding typically perceptually unimpaired lossy quantisation is used, we embarked on exploring the potential of the re-configurable FL-DCC codec under $b < 3$ low-rate, lossy conditions, while using a coding ring of $M = 8$ pixels.

Based on our findings in Figure 2 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 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.

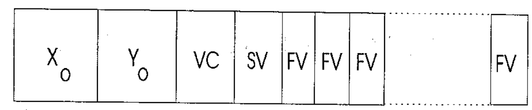

Figure 3: Coding Syntax

	$b = 1$ 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

The data syntax of the FL-DCC scheme is displayed in Figure 3. 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 $\times$ 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). Table 1 demonstrates that the coding rate of the proposed FL-DCC scheme is lower for $b = 2$ and $b = 1$ than that of DCC.

4. TRANSMISSION ISSUES

4.1. Modulation Aspects

As evidenced by the American IS-54, the Japanese Digital Cellular (JDC) and the Handyphone (PHP) systems, spectrally efficient multilevel modulation schemes are becoming popular. Their further advantage is that they conveniently lend themselves to channel-quality or teletraffic-motivated signal constellation re-configuration. In order to complement the proposed multi-rate FL-DCC graphical source codec we designed a re-configurable modem. The most robust but least bandwidth efficient 4-level Quadrature Amplitude Modulation (4QAM) [3] mode can be used in outdoors scenarios in conjunction with the $b = 1$ mode of the FL-DCC codec. The less robust but more bandwidth efficient 16QAM mode may be invoked in friendly indoors cells in order to accommodate the $b = 2$ mode of operation of the FL-DCC codec. When the channel conditions are extremely favourable, the modem can also be configured as a 64QAM scheme, in which case it can deliver $b = 3$ bits per FL-DCC vector, allowing lossless coding.

In our former work we found [3] that the rectangular 16QAM constellation exhibits two independent 2-bit subchannels having different bit error rates (BER), which naturally lend themselves to un-equal protection coded modulation. The BER of the lower quality class 2 (C2) subchannel was found a factor 2-3 times higher than that of the higher integrity class 1 (C1) subchannel. Hence the more vulnerable FL-DCC coded bits will be transmitted via the C1 subchannel, while the more robust bits over the C2 subchannel. While some previously proposed PSAM schemes used either a low-pass interpolation filter or an approximately Gaussian filter, Cavers deployed an optimum Wiener filter to minimise the channel estimation error variance $\sigma^2{}_e(k) = E\{e^2(k)\}$, where $E\{\ \}$ represents the expectation. However, this computationally demanding technique has a similar performance to the conventional sinc-interpolator. In our quest for the best interpolator we found that a simple polynomial interpolator outperformed the above mentioned schemes in terms of both mean squared estimation error as well as BER performance.

4.2. Packetisation Aspects

In order to determine the desirable length of the transmission packets we evaluated the histogram of the FL-DCC trace length, which exhibited a very long low-probability tail. However, most traces generated less than a few hundred bits, even when $b = 2$ was used. In order to be able to use a fixed packet length, while maintaining robustness against channel errors and curtailing transmission error propagation across packets and/or traces, we decided to tailor the number of bits per trace to the packet length of 222 bits. If a longer trace was encountered, it was forcidly truncated to this length and the next packet started with a new 'artificial' pen-down code. If, however, a shorter trace was encountered, a second trace was also fitted in to the current packet and eventually truncated to the required length for transmission. Unfortunately, the additional forcidly included VC code and the X_0, Y_0 coordinates portrayed in Figure 3 increased the number of bits generated but mitigated the error propagation effects. The proportion of the bit rate increase evaluated in terms of % for various packet lenghts in case of the FL-DCC $b = 1$ scheme is portrayed in Figure 4.

The BCH(255,131,18) and BCH(255,91,25) codes were found to ensure the required balance between the more and less robust FL-DCC bits, when employed in the C1 and C2 16QAM subchannels, respectively. The $2\times255=510$ BCH-coded bits constitute 128 4-bit symbols. After adding 14 pilot symbols according to a pilot spacing of $P = 10$ and concatenating 4 ramp symbols for smooth power amplifier ramping in order to minimise the out-of-band emissions the

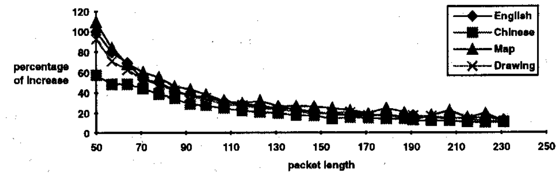

Figure 4: Proportion of bit rate increase due to fixed-length trace termination versus packet length for the FL-DCC $b = 1$ scheme

resulting 146-symbol packets are queued for transmission to the BS. The same packet format can be used for the voice and video users. The corresponding single-user voice and video signalling rate becomes 146 symbols/20ms=7.3 kBd. Graphical users assemble their 146-symbol transmission packets over a longer period and have a significantly lower source rate. The issue of maximum achievable graphical data rate will be addressed in the Results Section.

When using a bandwidth of 200 kHz, as in the Pan-European mobile radio system known as GSM [2] and a modulation access bandwidth of 50 % or a Nyquist roll-off factor of 0.5, a signalling rate of 133 kBd can be accommodated. This allows a conventional Time Division Multiple Access (TDMA) system to support 18 7.3 kBd speech or video users in the 16QAM mode of operation. Since both the speech and video rates are 8 kbps, alternatively 9 voice/video users can be supported. If the channel quality degrades, the more robust 4QAM mode can be invoked under BS control. This, however, halves the number of voice/video users supported. In this situation the FL-DCC graphical source encoder can reduce the number of bits from $b = 2$ to 1, while providing adequate graphics quality. In order to maintain the same 222-bit long framing structure we used two codewords of the shortened BCH(255,111,21) code in conjunction with the 4QAM modem scheme.

5. GRAPHICS, VOICE AND VIDEO MULTIPLEXING

In the proposed multimedia system packet reservation multiple access (PRMA) is invoked in order to multiplex graphics, voice and video packets for transmission to the base station (BS). Apart from carrying out this function the PRMA multiplexer facilitates a more efficient exploitation of the transmission medium by allocating packets on a flexible demand basis to the graphics, voice and video users. Since our video codec [4] delivers a constant-rate 8 kbps stream, it requires regularly-spaced physical packets, if no additional delay and buffering is tolerated. In contrast, human speech has bursty statistics and the proportion of active speech spurts is around 40% , which can be exploited to support a number of additional speech users or to accommodate a mixture of speech and data users [6].

The voice activity detector (VAD) queues speech packets for transmission to the Base Station (BS) and if the Portable Station (PS) gets permission to transmit, it reserves the time-slot for the duration of the current active speech spurt. If a passive speech spurt is detected, the time-slot is surrendered and can be reserved by other speech, data or video users, who are becoming active. If, however, due to packet collisions a speech user cannot get a reservation within the maximum tolerable delay of about 30 ms, the validity of the current speech packet expires, since new speech packets may be awaiting transmission. Hence this initial packet must be dropped but the dropping probability must be kept below $P_{drop} = 1\%$ [6] in order to minimise the speech degradation. Fortunately this initial spurt clipping is hardly perceivable.

In contrast, graphical data packets cannot be dropped, but tolerate longer delays and can be allocated to slots, which are not reserved by speech users in the present frame. In our experiments all but one graphical data users were obeying a negative exponential packet generation model, while one user transmitted graphical traces generated by a writing tablet. A less than unity permission probability either allowed or disabled permission to contend during any particular slot for a reservation within the current PRMA frame. Wong and Goodman noted [6] that it is advantageous to control data packet contentions on the basis of the fullness of the contention buffer. Specifically, contentions are disabled, until a certain number of packets awaits transmission, which reduces the probability of potential packet collisions due to the frequent transmission of short graphical data bursts. While speech communications stability can be defined as a dropping probability $P_{drop} < 1\%$, data transmission stability is specified in terms of maximum graphical data delay or buffer requirement.

6. 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 two of the input signals previously used in Table 1. Observe that for $b = 2$ no subjective degradation can be seen and the degradation associated with $b = 1$ is also fairly low. This is due to the fact that the typical fuzzy granular error patterns inflicted by the $b = 1$ FL-DCC scheme, when a straight line section is approximated by a zig-zag pattern, can be detected and smoothed by the decoder.

As mentioned before, the PRMA channel rate was 133 kBd, and 18 1.11 ms slots can be created in a 20 ms frame, which can accommodate 7.3 kBd source streams constituted by the pair of BCH(255,131,18) and BCH(255,91,25) C1 and C2 packets in the 16QAM mode. We dedicated nine time slots for video users, who had a unity permission probability, since our video codec generated a continuous, constant rate output stream and the delay had to be minimised. However, if less than nine videophone users are present at any moment, their slots can be offered to voice and graphics data users.

The remaining nine slots were offered to mixed speech and graphical data users. We found that in this scenario the optimum speech and graphical data permission probabilities were $P_{sp} = 0.6$ and $P_d = 0.05$, respectively. The significantly lower P_d value ensures that the voice users are not disadvantaged by agressively contending graphical data packets, which can more readily tolerate slight delays than

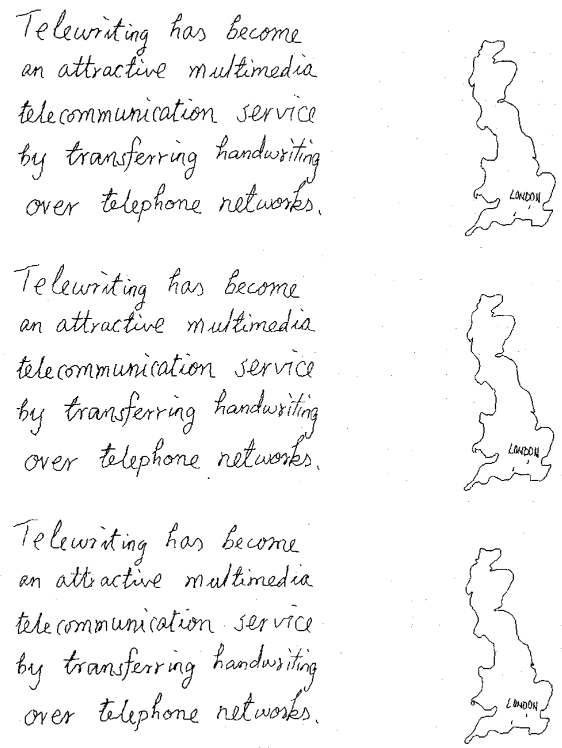

Figure 5: Decoded information for FL-DCC with $b = 1$ (bottom), $b = 2$ (centre) and DCC (top)

speech users. System stability was defined for speech users as $P_{drop} < 1\%$, and for graphical data users as a maximum data buffer length of 100 packets, each delivering a 146 symbol packet. The total maximum storage required was then $100 \times 146 \times 4/8$=7.3 KBytes. In order to limit the variety of traffic loading scenarios, similarly to the number of videophone users, we stipulated the number of graphical data users as nine.

The system's traffic loading was varied by adjusting the data transmission rate between 0.5 and 4 kbps in steps of 0.5 kbps and the number of speech users was set to 9, 12, 13 and 14. The minimum number of graphical data packets in the contention buffer to enable data contention was set to two, which slightly increased the mean data delay, but significantly reduced the chance of packet collision. Due to lack of space the traffic loading performance of the PRMA multiplexer is not studied here, but our results showed that due to speech burstyness and with the advent of PRMA the nine-slot system can support up to 5 additional speech users with a dropping probability around 1% plus nine 1 kbps graphical users. This corresponds to a system capacity improvement of $5 \times 8 + 9 \times 1$=49 kbps, or a relative capacity increase of $49/(9 \times 8) = 68$ %. 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

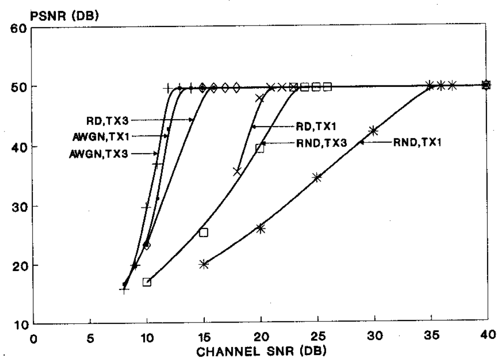

Figure 6: Graphical PSNR versus channel SNR performance of the b = 1-bit FL-DCC/16QAM mode of operation over various channels

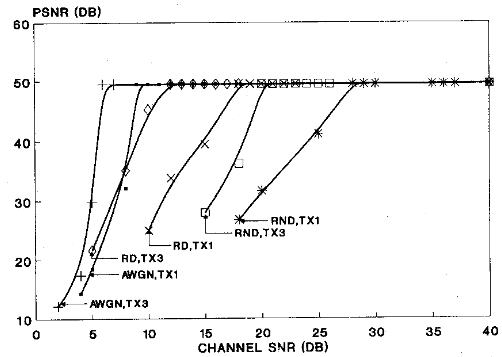

Figure 7: Graphical PSNR versus channel SNR performance of the b = 1-bit FL-DCC/4QAM mode of operation over various channels

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 lossy coding 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, respectively. As we showed in Figure 5, the subjective effects of a 10 dB PSNR degradation are not severe in terms of readability and a few dB further degradation due to channel effects is hardly perceivable.

The system's PSNR versus channel SNR performance, when using the low-rate $b = 1$ FL-DCC codec in conjunction with the 16QAM and 4QAM modes of operation is characterised by Figures 6 and 7. 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 higher-rate, higher-quality $b = 2$ mode, similar required channel SNR values apply, but the unimpaired PSNR is around 60 dB.

7. 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.

8. ACKNOWLEDGEMENT

The financial support of the EPSRC, UK in the framework of the research contract GR/J46845 is gratefully acknowledged.

9. REFERENCES

[1] R. Prasad, P.A.D. Spaargaren and J.H. Bons: "Teletext reception in a mobile channel for a broadcast teleinformation system", IEEE Tr. on VT., Vol. 42, No. 4, Nov. 1993, pp 535-545

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

[3] 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

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

[5] 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

[6] 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.