H261 AND H263-BASED PROGRAMABLE VIDEO TRANSCEIVERS

P. Cherriman, L. Hanzo

Dept. of Electr. and Comp. Sc., Univ. of Southampton, SO17 1BJ, UK.
Tel: +44-703-593 125, Fax: +44-703-594 508
Email: pjc94r@ecs.soton.ac.uk, lh@ecs.soton.ac.uk
http://www-mobile.ecs.soton.ac.uk

ABSTRACT

The video quality versus bitrate performance of the H.261 and H.263 codecs is characterised and a bitrate control algorithm is proposed in order to maintain a selectable near-constant video bitrate. The source codecs are operated at four different bitrates, which result in a constant symbol rate after error correction coding. This system constellation allowed us to maintain a constant video user bandwidth requirement, but benefit from improving channel conditions in terms of better video quality, when invoking more bandwidth efficient modem schemes ranging from 1 to 6 bits per symbol.

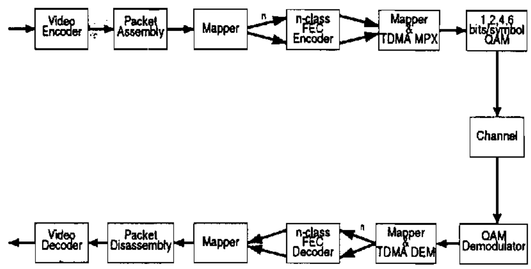

Figure 1: System's Schematic

1. MOTIVATION, SCOPE AND BACKGROUND

Motivated by the rapid emergence of wireless visual communications as a major research area [1] -[8], in this treatise we set out to contrive and investigate an intelligent, adaptively re-configurable wireless video transceiver. The proposed wireless video communicator can adapt to a range of service requirements, video quality, bitrate and robustness constraints by invoking the best possible system configuration under time-variant optimisation criteria.

Although the main target of the recently formed MPEG4 working group is to standardise a very low-rate videophone codec suitable also for mobile applications, currently no wireless video standard exists. In our former work we contrived a range of proprietary video codecs and embedded them in appropriate wireless transceivers [4]-[8]. In this treatise we embarked on investigating the feasibility of and documenting the performance of two wireless candidate systems based on the standard H261 and H263 ITU codecs.

The challenge of these endeavours was to overcome the deficiencies of these codecs due the fact that they were designed for benign Additive White Gaussian Noise (AWGN) channels, while maintaining compatibility with the Recommendations. Hence both the H261 and the H263 video codecs were subjected to rigorous bit sensitivity analysis and their robustness was improved using sensitivity-matched embedded binary Bose-Chaudhuri-Hocquenghem (BCH) coding combined with a re-configurable Pilot Symbol Assisted (PSA) and Automatic Repeat Request (ARQ) aided Quadrature Amplitude Modulation (QAM) modem [9].

The proposed system shown in Figure 1 can re-configure itself under network control in a variety of modes of operation, which span a wide range in terms of video quality, bit rate, robustness against channel errors and implementational complexity. Specifically, the transceiver operates using highly bandwidth efficient 64-level Pilot Symbol Assisted Quadrature Amplitude Modulation (64-PSAQAM) in a benign indoors cordless environment, where high signal-to-noise ratio (SNR) and signal-to-interference ratio (SIR) prevail. The number of modulation levels is dropped from 64 to 16, when the portable station (PS) is handed over to an outdoors street microcell, and can be further reduced to 4 in less friendly propagation scenarios.

THIS TREATISE IS COMPLEMENTED BY A DEMONSTRATION PACKAGE PORTRAYING VIDEO SEQUENCES AT VARIOUS BIT RATES, WHICH IS DOWN-LOADABLE FROM HTTP://WWW-MOBILE.ECS.SOTON.AC.UK

ICCS'96 NOV. 1996, SINGAPORE

2. PERFORMANCE OF THE H.261 AND H.263 CODECS

For reasons of space economy here we refrain from detailing the H.261 and H.263 coding algorithms and rather concentrate on the associated transceiver aspects. The H.263 standard supports five different video resolutions. In addition to the 288×352 pixel Common Intermediate Format (CIF) the so-called Quarter CIF (QCIF) Sub-QCIF (SQCIF), 4×CIF and 16×CIF formats can be used. Table 1 summarises the legitimate video formats and the associated source rates supported by the H.261 and H.263 standards.

In our comparative studies the H.261 simulations were carried out with and without motion compensation. When motion vectors were used, the optional so-called loop filter-

Picture Format	Luminance pixels	Luminance lines	H.261 support	H.263 support	Uncompressed bitrate (Mbit/s)			
					10 frame/s		30 frame/s	
					Grey	Colour	Grey	Colour
SQCIF	128	96		Yes	1.0	1.5	3.0	4.4
QCIF	176	144	Yes	Yes	2.0	3.0	6.1	9.1
CIF	352	288	Optional	Optional	8.1	12.2	24.3	36.5
4CIF	704	576		Optional	32.4	48.7	97.3	146.0
16CIF	1408	1152		Optional	129.8	194.6	389.3	583.9

Table 1: Picture Format supported

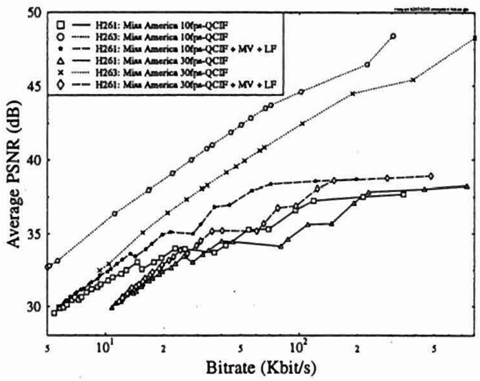

Figure 2: Image quality (PSNR) versus coded bitrate, for H.261 and H.263 simulations using grey scale QCIF "Miss America" video sequences at 10 and 30 frames/s

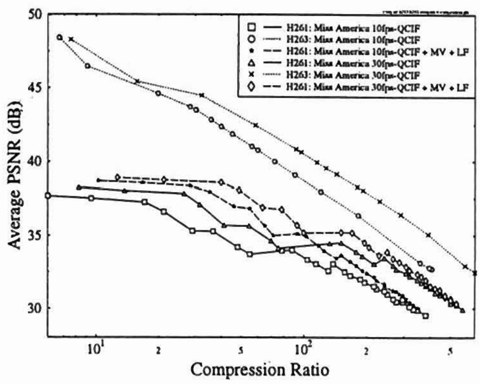

Figure 3: Image quality (PSNR) versus compression ratio, for H.261 and H.263 simulations using grey scale QCIF "Miss America" video sequences at 10 and 30 frames/s

ing was turned on in order to reduce the effects of block edge artifacts. In our H.263 simulations we used none of the so-called advanced negotiable options, but the codec employed motion vectors and inter-frame prediction for motion compensation. Our simulations were carried out at 10 and 30 frames/s, using different resolution versions of the same video sequence, resulting in a wide range of bitrates, in order to quantify the useful operating bitrate range of the codecs.

Figure 2 shows a comparison of the H.261 and H.263 video codecs in terms of image quality expressed in Peak Signal-to-Noise Ratio (PSNR) versus bitrate (Kbit/s). As the pair-wise comparisons reveal, irrespective of the bitrate, the performance of the H.263 codec is significantly higher than that of H.261 scheme in all scenarios.

Table 2 shows the bitrate reduction achieved for the same PSNR, when using the H.263 scheme instead of the H.261 arrangement. The reductions are relative to the H.261 simulations at 10 frames/s without motion compensation. The Table includes a column for the best H.261 performance, which uses motion vectors (MV) and loop filtering. Our next results summarised in Table 3 show the increase in image quality (PSNR) achieved at the same bitrate. The gains are relative to the H.261 simulations at 10 frames/s without motion compensation.

Figure 3 provides a comparison of the H.261 and H.263 video codecs from a different perspective, as image quality (PSNR) versus compression ratio. As can be seen from the Figure, the H.263 codec is significantly more efficient than the H.261 scheme in all investigated scenarios. Let us now highlight the concept of controlling the bitrate of the source codec in order to maintain a near-constant channel signalling rate.

3. BITRATE CONTROL

As with most video codecs using variable-length coding techniques, the bitrate of the H.263 codec is inherently time variant. However, most existing mobile radio systems transmit at a fixed bitrate. Our proposed multi-mode system maintains a constant signalling rate or symbol rate, leading to a different constant bitrate for each of the modulation schemes invoked.

The PSNR versus bitrate performance of the video codecs was quantified in Figure 2 and the portrayed relationship plays an important role in the codecs' bitrate control. The control algorithm modifies the quantiser in the video codec in order to maintain the target framerate and bitrate. In these investigations a fixed framerate of 10 frames/s was employed. The target bitrate was set according to the the past history of the modulation schemes used and to the packet dropping frequency experienced. The target bitrate R was updated after every dropped packet or successful transmission according to Equation 1:

$$R = \frac{S_4 B_4 + S_{16} B_{16} + S_{64} B_{64}}{S_4 + S_{16} + S_{64} + D_4 + D_{16} + D_{64}} \quad (1)$$

where S_n is the number of successful packet transmissions,

Fixed PSNR(dB)	Percentage Bitrate Reduction (%)				
	10fps+MV+LF	H263-10fps	30fps+NoMC	30fps+MV+LF	H.263-30fps
33	19.53	62.28	-61.08	-41.31	26.90
35	60.57	82.99	-85.16	33.34	70.30
37	59.46	87.86	-72.33	10.19	77.78

Table 2: Percentage reduction in bitrate required to achieve the same PSNR for the H.261 and H.263 schemes, compared with the H.261 performance at 10 frames/s using no motion compensation

Fixed Bitrate(Kbit/s)	PSNR increase (dB)				
	10fps+MV+LF	H263-10fps	30fps+NoMC	30fps+MV+LF	H.263-30fps
20	1.54	5.20	-0.98	-0.51	2.76
30	1.51	6.43	-0.57	0.39	3.98
50	2.55	7.50	-0.44	0.35	4.91
100	1.73	7.81	-1.56	0.32	5.59
190	1.22	8.49	-0.41	1.18	7.04

Table 3: PSNR(dB) improvement achieved at given bitrates by the H.261 and H.263 schemes, compared with the H.261 PSNR at 10 frames/s without motion compensation

when using the n-QAM modulation scheme, D_n is the number of dropped packets in the n-QAM mode of operation, while B_n represents the various fixed bitrates for error-free transmission, again, using the n-QAM modulation scheme. If, for example, the videophone system is restricted to operate in the 4QAM mode, the target bitrate is simplified as shown in Equation 2:

$$R = \frac{S_4 B_4}{S_4 + D_4}. \quad (2)$$

When the channel is near error-free, then the number of dropped packets D_4 will be virtually zero and the target bitrate will tend to the fixed bitrate B_4 stipulated for the 4QAM transmission mode. If however the channel was causing 50% frame errors then the number of dropped transmissions D_4 would be similar to the number of successful transmissions, namely to S_4. This would cause the traget bitrate to reduce to approx. 50% of the fixed bitrate for the 4QAM scheme. Therefore the target bitrate reduces in proportion to the channel frame error rate. If the video codec can meet the target bitrate requirements, the frame rate of the transmitted video becomes also nearly constant, although under severe channel conditions the system may have to reduce the video frame rate, in order to arrive at a constant bitrate. The operation of this bitrate control algorithm is characterised by the normalised bitrate histograms of Figures 4 in the transceiver's different modes of operation using 64QAM, 16QAM, 4QAM and Binary Phase Shift Keying (BPSK), respectively. Observe in the Figure that in the lowest capacity BPSK mode due to the stringent bitrate constraint we opted for transmitting in the SQCIF image format. The corresponding frame-rate versus time behaviour of the system is characterised by Figure 5. Observe that after a slight initial delay, which is essentially due to the intra-frame coded bitrate surge of the first frame, the frame rate reaches the target of 10 frames/s, typically within less than 1s. Let us now focus our attention on some of the error control issues.

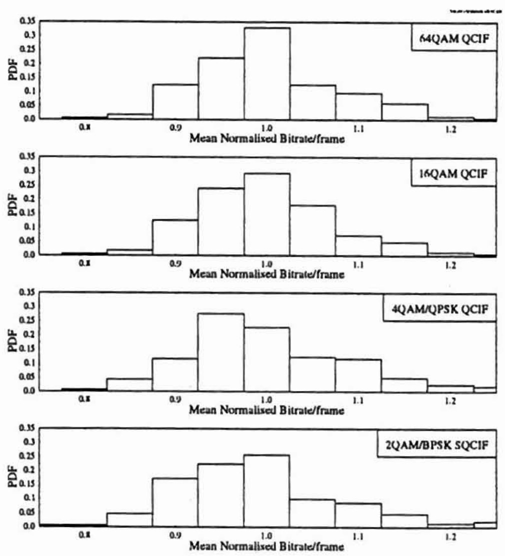

Figure 4: Bitrate control histogram for transmissions at 4,16, and 64QAM using the QCIF resolution Miss America Sequence, and 2QAM using the SQCIF resolution Miss America Sequence, both in error-free conditions

4. FORWARD ERROR CORRECTION CODING

Again, let us consider transmitting QCIF images, where the video codecs were programmed to generate 3560, 2352 and 1176 bits per frame, which at a scanning rate of 10 frames/s resulted in bitrates of 35.6 kbps, 23.52 kbps and 11.76 kbps, respectively, when using 61QAM, 16QAM and 4QAM.

In our earlier work we have shown that in QAM schemes the bits can be assigned to a number of different integrity classes [9]. The number of integrity classes depends on the number of modulation-levels, and in 4-QAM there is only one integrity class, in 16-QAM there are 2, while in 64-QAM

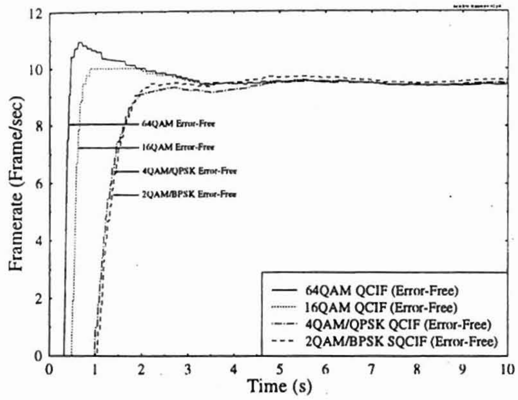

Figure 5: Framerate versus time behaviour of the frame-rate control algorithm, for transmissions at 4,16, and 64QAM using QCIF resolution Miss America Sequence, and 2QAM using SQCIF resolution Miss America Sequence, both in error-free conditions

Modulation scheme	FEC codes used
4 QAM	BCH(255,147,14)
16 QAM	Class 1: BCH(255,179,10) Class 2: BCH(255,115,21)
64 QAM	Class 1: BCH(255,199,7) Class 2: BCH(255,155,13) Class 3: BCH(255,91,25)

Table 4: FEC codes used for 4, 16 and 64 QAM

there are 3 classes, often also referred to as sub-channels. By using different strength FEC codes on each QAM sub-channel it is possible to equalise the probability of errors on the sub-channels. This means that all FEC-protected sub-channels will exhibit similar BERs. This is desirable, if all bits to be transmitted are equally important. Since our datastreams are variable length coded, one error can cause a loss of synchronisation. Therefore in this case most bits are equally important, and so equalisation of the QAM sub-channels' BER is desirable. The FEC codes used in our system are summarised in Table 4, while the corresponding transmission packet sizes are highlighted in Table 5 along with the initial quantiser identifiers, which predetermine the achieveable bitrate. After including a control header, pilot and ramp symbols [9], the FEC-coded single-user signalling rates became 11.84 kBd in all three operating modes. In case of a Nyquist excess bandwidth of 50% this implies a single-user bandwidth requirement of 17.74 kHz. For example in the 200 kHz bandwidth of the Pan-European GSM system 8 voice-only users or supported, which corresponds to a 25 kHz user bandwidth. Consequently, our videophone stream can replace a speech stream, making wireless videophony realistic, if the cost of an additional timeslot is acceptable to the users.

Modulation	TX. packet size	Initial Quantiser
4 QAM	147	9
16 QAM	294	6
64 QAM	445	4

Table 5: H.263 packet size in terms of bits and initial quantiser used for 4, 16, and 64 QAM

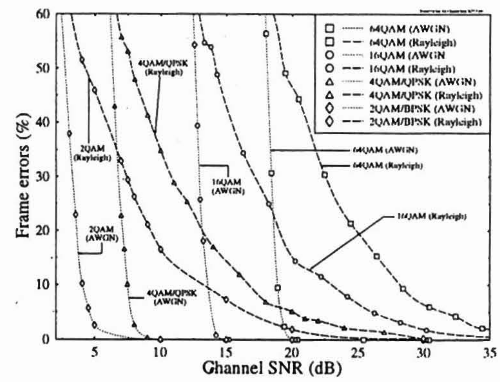

Figure 6: Frame Error Rate (FER) versus Channel SNR (dB) for 2, 4, 16, and 64QAM over both Gaussian and Rayleigh fading channels

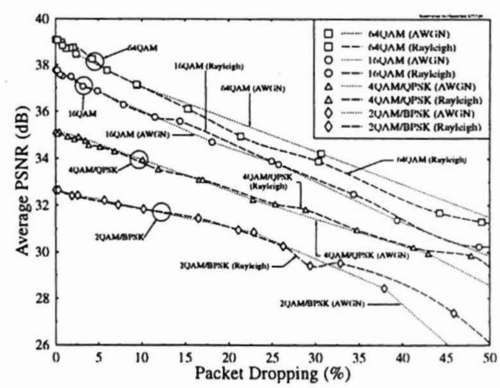

Figure 7: Average PSNR of decoded video versus Frame Error Rate (FER), for 2, 4, 16, and 64QAM overover both gaussian and rayleigh fading channels. QCIF resolution Miss America video sequence used for all transmission modes.

4

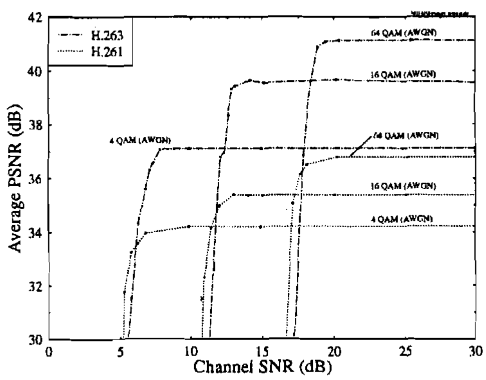

Figure 8: Performance comparison of the proposed adaptive H261 and H263 transceivers over AWGN channels

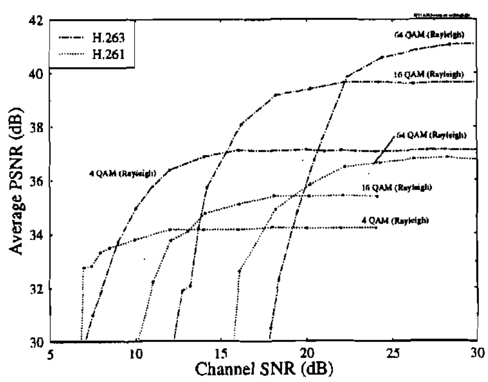

Figure 9: Performance comparison of the proposed adaptive H261 and H263 transceivers over Rayleigh channels

5. SYSTEM PERFORMANCE

The video system performance was evaluated under the propagation conditions of a vehicular speed of 30 mph, signalling rate of 11.84 kBd and propagation frequency of 1.9 GHz. Figure 6 displays the Frame Error Rate (FER) versus channel SNR performance of the system for 2, 4, 16, and 64QAM over both Gaussian and Rayleigh fading channels. The graphs clearly indicate that extremely robust system performance can be maintained upon invoking the more robust modes of operation, when the channel conditions degrade. The performance curves can be related to the video quality exressed in terms of PSNR using Figure 7.

Finally, in the various operating modes investigated the PSNR versus channel SNR curves of Figures 8 and 9 were obtained for Additive White Gaussian Noise (AWGN) and Rayleigh channels, respectively. Since both the H.261 and H.263 source codecs have had similar robustness against channel errors, and their transceivers were identical, the associated 'corner SNR' values, where unimpaired communications broke down were virtually identical for both systems over both AWGN and Rayleigh channels. However, as expected, the H263 codec again exhibited always higher video quality at the same bitrate or system bandwidth.

Our current endeavours are focussed on exploring the quality versus bitrate performance of both systems for various image resolutions, in order to be able to provide the required video quality, bit rate, frame rate, image size and resolution on a demand basis in intelligent adaptive multimode transceivers.

6. ACKNOWLEDGEMENT

The financial support of the EPSRC, UK in the framework of the contract GR/K74043 is gratefully acknowledged.

7. REFERENCES

[1] *IEEE Transactions on Circuits and Systems for Video Technology, Special issue on Very Low Bit Rate Video Coding*, vol. 4, pp. 213–357, June 1994.

[2] M. Khansari, A. Jalali, E. Dubois, and P. Mermelstein, "Robust low bit-rate video transmission over wireless access systems," in *Proc. of International Comms. Conf. (ICC)*, pp. 571–575, 1994.

[3] R. M. Pelz, "An un-equal error protected px8 kbit/s video transmission for dect," in *Proceedings of IEEE VTC '94* [10], pp. 1020–1024.

[4] R. Stedman, H. Gharavi, L. Hanzo, and R. Steele, "Transmission of subband-coded images via mobile channels," *IEEE Tr. on Circuits and Systems for Video Technology*, vol. 3, pp. 15–27, Feb 1993.

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

[6] L. Hanzo and J. Streit, "Adaptive low-rate wireless videophone systems," *IEEE Tr. on Video Technology*, Aug 1995.

[7] R. Stedman, H. Gharavi, L. Hanzo, and R. Steele, "A 22 kbd microcellular video telephone scheme," in *Proceedings of IEEE VTC '92*, (Denver, U.S.A.), pp. 251–254, 1992.

[8] J. Streit and L. Hanzo, "A fractal video communicator," in *Proceedings of IEEE VTC '94* [10], pp. 1030–1034.

[9] W. Webb and L. Hanzo, *Modern Quadrature Amplitude Modulation: Principles and Applications for Wireless Communications*. IEEE Press-Pentech Press, 1994.

[10] *Proceedings of IEEE VTC '94*, (Stockholm, Sweden), June 8-10 1994.