# Orthogonal Frequency Division Multiplex transmission of H.263 encoded video over wireless ATM networks

P. Cherriman, T. Keller, and L. Hanzo
Dept. of Electronics and Computer Science,
University of Southampton, SO17 1BJ, UK.
Tel: +44-703-593 125, Fax: +44-703-594 508
Email:lh@ecs.soton.ac.uk
http://www-mobile.ecs.soton.ac.uk

**Abstract: The video performance of the Median wireless asynchronous transfer mode (WATM) system is evaluated for a range of application scenarios using the H.263 video codec and a novel packetisation and acknowledgement scheme. The video resolutions and system parameters used are summarised in Tables 1 and 2. The required channel signal-to-noise ratio for near-unimpaired video quality is about 16dB over the dispersive worst-case channel used.**

## 1 System Overview

The proposed wireless LAN system's schematic is shown in Figure 1, which supports videophone calls. The video signal is compressed using the H.263 video compression standard [1]. The H.263 standard achieves very high compression ratio, however the resulting bitstream is extremely sensitive to channel errors. This sensitivity to channel errors is not a serious problem over benign wireline-based Gaussian channels, but it is an impediment, when used over wireless networks. There have been several solutions suggested in the literature for overcoming this using Automatic Repeat Request (ARQ) [2], dual-level coding [3] and the use of a feedback channel [4]. A range of further robust video schemes were proposed in [5, 6, 7].

As seen in Figure 1, our system uses a feedback channel to inform the encoder of the loss of previous packets. However, we do not retransmit the corrupted packets, since this would reduce the system's teletraffic capacity by occupying additional transmission slots, while increasing the video delay. We found that simply dropping the corrupted packets at both the local and remote decoder results in an extremely high error resilience, in particular in high frame-rate systems, where 30 frames/s high-rate transmissions are facilitated. This allows the local and remote decoder to stay in sychronisation, which is essential

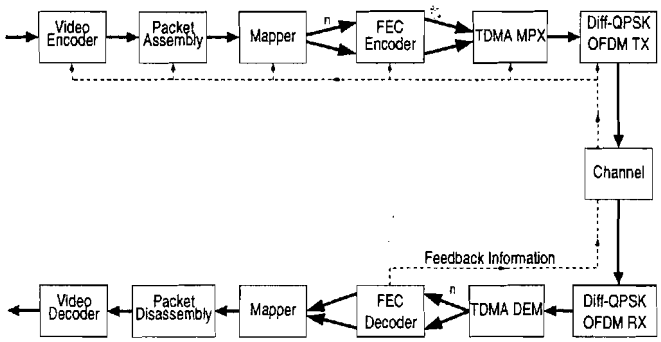


Figure 1: Transceiver schematic

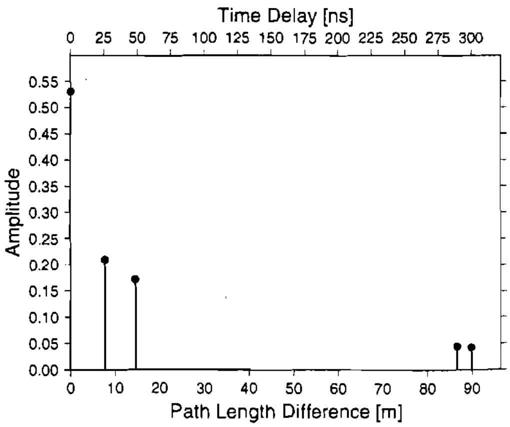


Figure 2: Indoors five-path impulse response

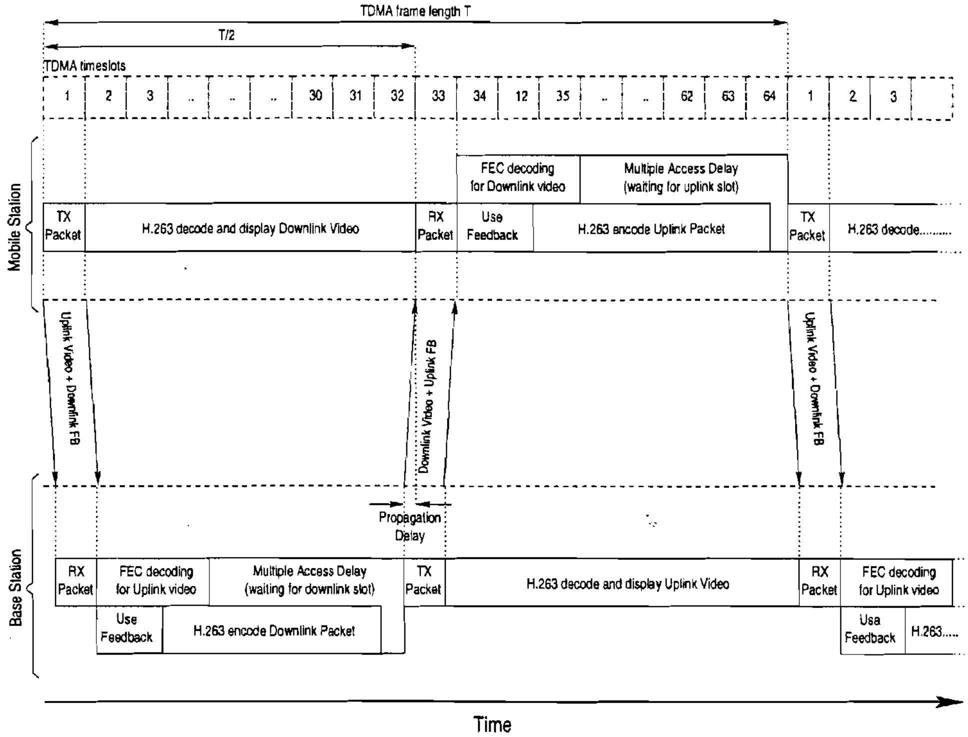


Figure 3: Transmission feedback timing diagram showing the feedback signalling superimposed on the reverse channel video datastream. The tasks that need to be performed in each time interval are shown for both the mobile station and the basestation.

for preventing error propagation through the reconstructed frame buffer. Then, when the instantaneous channel quality improves, the corrupted picture segments of the reconstructed frame buffers are replenished with more uptodate video information. The feedback channel is implemented by superimposing the packet retransmission request on the reverse link, as shown in Figure 3. This Figure shows how the feedback acknowledgement is implemented in the context of a Time Division Multiple Access (TDMA), Time Division Duplex (TDD) system using 32 timeslots, where one video packet was transmitted in each TDD frame. A number of further interesting details become explicit by studying the Figure.

As demonstrated by Figure 1, the H.263 encoded bitstream in passed to the packet assembly block. The packetiser's function is to assemble the video packets for transmission, taking into account the packet acknowledgement feedback information. The packet disassembly block of the Figure ensures that an error-free H.263 bitstream is output to the video decoder, discarding any erronously received packet and using only error-free packets to update the reconstructed frame buffer. Since the transmission packets contain typically fractions of video macroblocks at the beginning and end of the packets, a corrupted packet implies that the previously received partial macroblocks have to be discarded. The loss of the packet is then signalled via the feedback channel to the video encoder and packet assembly blocks. The lost macroblocks are not re-transmitted, but signalling flags are inserted into the video bitstream to signify the macroblocks that have not been updated. This requires one bit per lost macroblock in the next reverse-direction packet. The decoded videostream is error-free, although certain parts of some video frames may be frozen for a few frames due to lost packets. These areas will be usually be updated in the next video frame, and the effect of the lost packet will be no longer visible. An example of a typical scenrio is shown in Figure 4, which is fairly self-explanatory. The packetised video stream is then Forward Error Correction (FEC) coded, mapped to the allocated TDMA timeslot and transmitted using Differential Quarternary Phase Shift Keying (D-QPSK) between adjacent sub-carriers of the Orthogonal Frequency Division Multiplex (OFDM) scheme employed [8]. In the next Section let us now briefly consider the specific system parameters proposed by the Pan-European Wireless Asynchronous Transfer Mode (WATM) consortium Median.

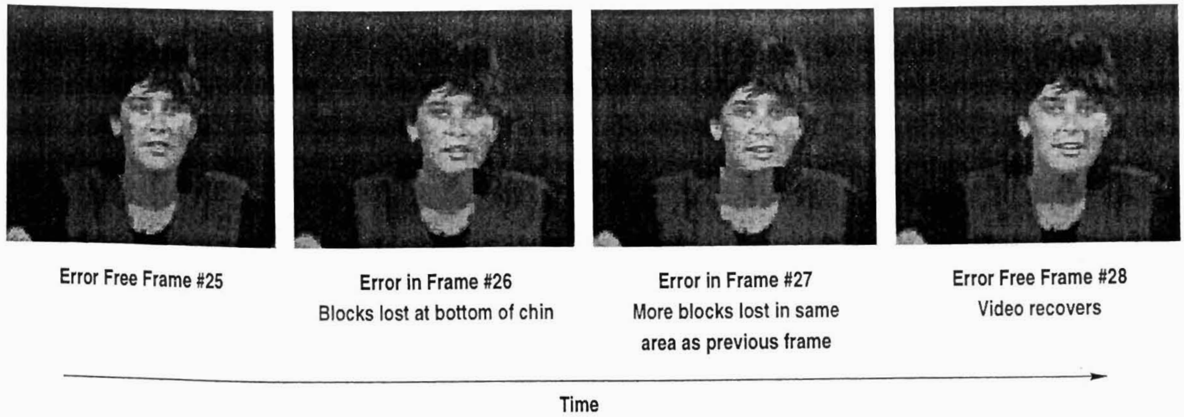


Figure 4: Typical expected video quality of the proposed video transceiver for 16Kbit/s at 10fps, for a frame error rate (or packet dropping rate) of 5%. This figure also shows the error recovery and concealment used in frames 26 and 27.

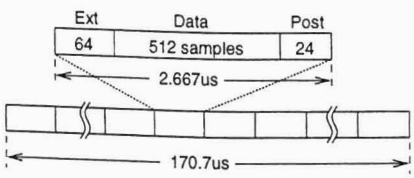


Figure 5: Schematic plot of the ACTS MEDIAN frame structure. A time frame contains 64 timeslots of 2.667$\mu s$ duration. Each timeslot holds the data samples of a 512 point IFFT OFDM symbol, 64 samples of the cyclic extension and a cyclic postamble of 24 samples.

| Feature | Value |
|---|---|
| TDMA frame length | $171\mu s$ |
| Slots/Frame | 64 (61 useable) |
| Slot length | $2.667\mu s$ |
| OFDM carriers | 512 (511 used) |
| Modulation | Differential-QPSK |
| Coded Bits/slot | 1022 bits |
| FEC (1/2 rate) | BCH(255,131,18) |
| Pre-FEC bits/slot | 524 bits |
| System Bandwidth | 225MHz |
| System Symbol rate (symbols/sec) | $186 \times 10^6$ |
| Normalised Doppler Frequency | $1.235 \times 10^{-5}$ |

Table 1: Summary of Median System Parameters

## 2 The Median System

The system employed in our experiments follows closely the ACTS MEDIAN proposal, a system designed as an ATM data rate compatible wireless indoors radio network, operating in the 60GHz band utilising OFDM as modulation technique. The main MEDIAN system parameters are listed in Table 1. Channel access in the MEDIAN system is based on a Time Division Multiple Access / Time Division Duplex (TDMA/TDD) frames having a duration of 170.7$\mu s$. This frame is split into 64 time-slots of 2.667$\mu s$ duration. Two of these time slots are reserved for networking functions, leaving 62 for useful information transfer.

In order to avoid utilising equalisation at the FEC-coded sampling rate of 225Msamples/s, OFDM is employed as baseband modulation technique [8]. Differential Quarternary Phase Shift Keying (D-QPSK) between adjacent sub-carriers is used as frequency–domain modulation scheme. If all 512 sub-carriers are used and only one phase reference sub-carrier is employed, then 1022 bits can be transmitted using a single OFDM symbol. Figure 5 portrays the Median frame- and slot-structure. Each time-slot contains one OFDM symbol, preceeded by a cyclic extension of 64 samples in order to combat interference in wideband channels. A cyclic post-amble is appended to the OFDM data samples in order to simplify symbol timing synchronisation. For our simulations, all 512 sub-carriers per OFDM symbol were used, resulting in the maximum throughput of 1022 bits per OFDM symbol. Variable bit rate users can be accommodated by allocating groups of time-slots per frame. The transmission system can be adapted to different user data packet rates by skipping time frames.

| Feature | Video Resolution | | | |
|---|---|---|---|---|
| | QCIF | CIF | CIF4 | CIF16 |
| Luminance resolution (pixels) | 176x144 | 352x288 | 704x576 | 1408x1152 |
| Crominance Resolution (pixels) | 88x72 | 176x144 | 352x288 | 704x576 |
| Packet separation (in No. of TDMA frames) | 30 | 6 | 1.5 | 1 |
| Packet rate (packets/s) | 195 | 975 | 3900 | 5448 |
| Bits/Timeslot | 1022 | 1022 | 1022 | 1022 |
| Timeslots per active TDMA frame | 5 | 2 | 5 | 7 |
| Bits per active TDMA frame (packet size) | 5110 | 2044 | 5110 | 7154 |
| Channel Bitrate | 1Mbps | 2Mbps | 20Mbps | 41.8Mbps |
| FEC | 20×BCH(255,131,18) | 8×BCH(...) | 20×BCH(...) | 28×BCH(...) |
| Pre-FEC Bits per active TDMA frame | 2620 | 1048 | 2620 | 3668 |
| Pre-FEC Bitrate | 511Kbps | 1Mbps | 10.2Mbps | 21.5Mbps |
| Feedback control bits | 26 | 24 | 26 | 29 |
| H.263 Packetisation header bits | 13 | 12 | 13 | 14 |
| Video bits per active TDMA frame | 2581 | 1012 | 2581 | 3625 |
| Useful Video Bitrate | 503Kbps | 1Mbps | 10Mbps | 21.2Mbps |

Table 2: Summary of Video parameters for Median System

# 3 The Channel Model

The channel model employed for the experiments is a five–path, Rayleigh fading indoors channel. The impulse response as shown in Figure 2 was obtained by ray–tracing for a $100 \times 100m^2$ hall or warehouse environment, and every path in the impulse response was faded independently according to a Rayleigh fading narrow band channel with a normalised Doppler frequency of $f'_d = 1.235 \cdot 10^{-5}$, corresponding to the 60GHz propagation frequency and a worst-case indoor speed of 30mph.

# 4 Video Formats

The Median system is a high bandwidth wireless LAN system, and therefore it constitutes an ideal medium for high resolution video transmission. In order to assess the system's ability to support various application scenarios, we simulated the transmission of four different resolution video material, ranging from 176x144 pixel Quarter Common Intermediate Format (QCIF) video telephony sequences to high definition television (HDTV) sequences, corresponding to an image resolution of 1408x1152 pixels. The video resolution of all the video frame sizes employed are summarised in Table 2. The video packetizer operated most efficiently, when the packet generation rate was neither too high nor too low. If the packet-generation rate is too high, each packet may contain less than a whole macroblock, leading to an increased buffering in the de-packetizer. If the packet generation rate is too low, then each packet contains a high number of macroblocks, and therefore when a packet is lost, a large porportion of the video frame is lost. Therefore the packet generation rate for each video resolution was adjusted experimentally, taking into account that as the video resolution was increased four-fold, corresponding to increasing the video resolution from QCIF to CIF, the number of macroblocks per frame or per time unit increased by the same factor, resulting in a corresponding increase in terms of the packet generation rate.

The packet generation rate for each of the four video resolutions used is shown in Table 2 in terms of both the number of TDMA frames between video packets or, synonymously, the TDMA packet separation, as well as in terms of packets per second. For example, for QCIF resolution video the packet generation rate is 195 per second, which corresponds to one packet every 30 TDMA frames. After setting the packet generation rates for each video mode, the video target bitrates were set to give high quality for the majority of video sequences. As mentioned previously, the OFDM system can transmit 1022 bits in every timeslot. Hence for the QCIF mode, with one timeslot every 30 TDMA frames the channel bitrate would be $1022 \times 195 = 200Kbit/s$. Upon using half-rate channel coding the video bitrate becomes 100Kbit/s. Since we required around 500Kbit/s for high quality QCIF video for a wide range of video sequences, we decided to use 5 timeslots once every 30 TDMA frames. Therefore the channel bitrate became $5 \times 1022 \times 195 = 1Mbit/s$, providing a bitrate of 500Kbit/s for video source coding. The interested reader is referred to `http://www-mobile.ecs.soton.ac.uk` for some examples of coded sequences, which can be viewed using an MPEG player.

For the Median experiments we decided to use embedded binary Bose-Chaudhuri-Hocquenghem

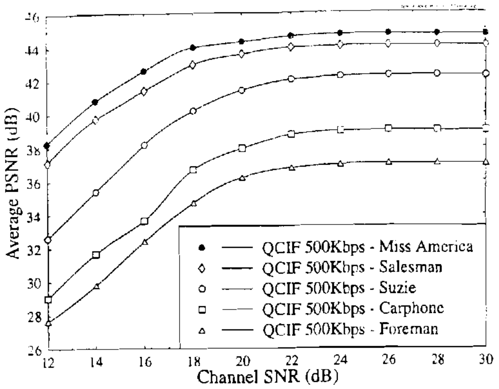


Figure 6: Video Quality in PSNR versus channel SNR for QCIF video using 500Kbit/s at 30fps, using the Median System, for various video sequences.

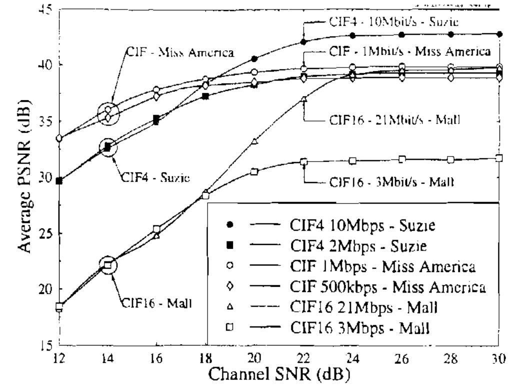


Figure 7: Video Quality in PSNR versus channel SNR for CIF, CIF4, and CIF16 resolution video at 30fps, using the Median System, for various video sequences.

(BCH) block coding, since it is capable of both error correction and error detection. For all the modes we used the near half-rate code of BCH(255,131,18) was employed. The corresponding pre-FEC bitrates for the various modes are shown in Table 2. In conjunction with this channel coding scheme the pre-FEC bitrate for the QCIF mode is 511Kbit/s.

The videophone system requires some additional overhead for its operation, since the feedback information for the reverse link is concatenated with the information packet, requiring between 24-29 bits/packet. In addition, the H.263 packetisation adds a header to each packet, which is dependent on the number of bits in each packet and for our simulation this header was between 12-14 bits per packet. Therefore the number of useful video source bits in each packet used for video transmission was about 40 bits less than the actual number of bits/packet. The corresponding video source bitrate for each of the modes is shown in Table 2, which was 503Kbit/s for QCIF resolution video, when taking into account the above-mentioned transmission overheads. Having highlighted the salient system features, let us now consider the achievable system performance in the next Section.

# 5 System Performance

Figure 6 shows the video quality in terms of the average Peak-Signal-to-Noise-Ratio (PSNR) versus the channel SNR for QCIF resolution video transmitted over the Median wireless LAN. The Figure shows the video quality of a range of video sequences from the highly motion active "Foreman" and "Carphone" sequences to the more compressable "Miss America" sequence. For all the video sequences the PSNR starts to drop, when the channel SNR falls below about 20dB. Due to lack of space the frame error rate (FER) versus channel Signal-to-Noise Ratio (CSNR) performance of the system is not explicitly characterised in this treatise, but our records show that the frame error rate around this CSNR value is about 3%. The corresponding visual quality appears unimpaired and the effects of dropping do not become evident for CSNRs in excess of 16dB. At 16dB the frame error rate is 17%. However, the effects of this packet loss is only becoming 'just noticeable' at a CSNR of 16dB. The effect of the packet loss, is that parts of the picture are frozen, but usually for only one video frame duration of about 30ms at 30frames/s, which is not sufficiently long for these artifacts to become objectionable. However, if the part of the picture that was lost contains a moving object, the effect of the loss of the packet becomes more obvious. Therefore for more motion active sequences, the effect of packet loss is more pronounced.

In order to portray the expected system performance in other application scenarios, where higher video quality is expected, in Figure 7 we portrayed the average PSNR versus channel SNR performance for a range of video resolutions from CIF to 16xCIF HDTV quality. At CIF resolution the 'Miss America' sequence was used encoded at both 500kbps and 1Mbps. For 4CIF resolution the 'Suzie' sequence was encoded at 2 and 10Mbps, while for 16CIF resolution the 'Mall' video clip was transmitted at 3 and 21 Mbps. This Figure shows results for using multiple timeslots per active TDMA frame, as suggested by Table 2, down to just a single timeslot per active TDMA frame. Notice that the high and low bitrate modes for each resolution seem to converge to a similar PSNR, when the channel SNR is low. This is likely to be due to the packetisation algorithm reducing the bitrate, as the packet dropping rate increases, in order to maintain a constant video framerate.

The CIF16 simulations seems to be more vunerable to packet loss, however this is because the packet generation rate is not four times that of the CIF4 simulations, and therfore each CIF16 video packet contains approximately 2.5 times more macroblock per packet than the CIF and CIF4 resolution video packets. Therefore the effect of packet loss is more noticable and this is manifested in the faster reduction of the PSNR, as the channel SNR degrades.

# 6 Conclusions

In this contribution the expected video performance of the Median WATM systems was quantified in a variety of applications scenarios, using a range of video resolutions and bitrates. The high-efficiency H.263 video codec was employed to compress the video signal. The video formats used were summarised in Table 2 along with their associated target bitrate figures. The proposed system ensures extremely robust video communications using the Median WATM systems in a dispersive Rayleigh-fading environment even at the worst-case vehicular speed of 30mph, requiring CSNRs in excess of only about 16dB for near-unimpaired video transmission. The associated networking aspects and the system's user capacity are summarised in the companion paper[9].

# 7 Acknowledgement

The financial support of Motorola ECID, Swindon, UK and that of the European Commission in the framework of the Median project is gratefully acknowledged. The authors are also indebted to the Median and First Consortium partners for fruitful discussions and for their friendship.

# References

[1] ITU-T, *Draft Recommendation H.263: Video coding for low bitrate communication.* Due to be approved 1996.

[2] N. MacDonald, "Transmission of compressed video over radio links," *BT technology Journal*, vol. 11, pp. 182–185, April 1993.

[3] M. Khansari, A. Jalali, E. Dubois, and P. Mermelstein, "Low bit-rate video transmission over fading channels for wireless microcellular systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 1–11, February 1996.

[4] N. Färber, E. Steinbach, and B. Girod, "Robust H.263 video transmission over wireless channels," in *Proc. of International Picture Coding Symposium (PCS)*, (Melborne, Australia), pp. 575–578, March 1996.

[5] P. Cherriman and L. Hanzo, "Programmable H.263-based wireless video transceivers for interference-limited environments," in *Proc. of IEEE ICCS'96 / ISPACS'96*, vol. 3, (Singapore), pp. 1369–1373, 25-29 November 1996.

[6] P. Cherriman and L. Hanzo, "Power-controlled H.263-based wireless videophone performance in interference-limited scenarios," in *Proc. of IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications*, vol. 1, (Taipei, Taiwan), pp. 158–162, 15-18 October 1996.

[7] P. Cherriman and L. Hanzo, "Programable H.263-based wireless video transceivers for interference-limited environments." accepted for publication in IEEE Tr. on CSVT, 1997.

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

[9] J. Brecht and L. Hanzo, "Statistical packet assignment multiple access for wireless asynchronous transfer mode systems." 7-10 October 1997.