MULTI-MODE JOINT-DETECTION CDMA/H.263 BASED VIDEO TELEPHONY

P. Cherriman, E.L. Kuan, L. Hanzo

Dept. of Electronics and Computer Science, Univ. of Southampton, SO17 1BJ, UK.
Tel: +44-1703-593125, Fax: +44-1703-594508
Email: pjc@ecs.soton.ac.uk, lh@ecs.soton.ac.uk
http://www-mobile.ecs.soton.ac.uk

ABSTRACT

A multi-mode joint-detection CDMA-based video transceiver is proposed for wireless video telephony, which substantially outperforms the matched-filtering based bench-marker video transceiver. For channel SNRs in excess of about 5 dB near-unimpaired video quality is maintained by the proposed scheme.

1. VIDEO TRANSCEIVER

In this study we transmitted 176x144 pixel Quarter Common Intermediate Format (QCIF) and 128x96 pixel Sub-QCIF (SQCIF) video sequences at 10 frames/s using a reconfigurable Time Division Multiple Access / Code Division Multiple Access (TDMA/CDMA) transceiver, which can be configured as a 1, 2 or 4 bit/symbol scheme shown in Figure 1. The H.263 video codec [1] exhibits an impressive compression ratio, although this is achieved at the cost of a high vulnerability to transmission errors, since a run-length coded stream is rendered undecodable by a single bit error. In order to mitigate this problem, when the channel codec protecting the video stream is overwhelmed by the transmission errors, we refrain from decoding the corrupted video packet in order to prevent error propagation through the reconstructed video frame buffer [4]. We found that it was more beneficial in video quality terms, if these corrupted video packets were dropped and the reconstructed frame buffer was not updated, until the next video packet replenishing the specific video frame area was received. The associated video performance degradation was found perceptually unobjectionable for packet dropping- or transmission frame error rates (FER) below about 5%. These packet dropping events were signalled to the remote decoder by superimposing a strongly protected one-bit packet acknowledgement flag on the reverse-direction packet, as outlined in [4]. Bose-Chaudhuri-Hocquenghem (BCH) [5] and turbo error correction codes [6] were used and again, the CDMA transceiver was capable of transmitting 1, 2 and 4 bits per symbol, where each symbol was spread using a low spreading factor (SF) of 16, as seen in Table 1. The associated parameters will be addressed in more depth during our further discourse. Employing a low spreading factor of 16 allowed us to improve the system's multi-user performance with the aid of joint-detection techniques [7]. We note furthermore that the implementation of the joint detection receivers is independent of the number of bits per symbol associated with the modulation mode used, since the receiver simply inverts the associated system matrix and invokes a decision concerning the received symbol, irrespective of how many bits per symbol were used. **Therefore, joint detection receivers are amenable to amalgamation with the above 1, 2 and 4 bit/symbol modem, since they do not have to be reconfigured each time the modulation mode is switched.**

In this performance study we used the Pan-European FRAMES proposal [2] as the basis for our CDMA system. The associated transmission frame structure is shown in Figure 2, while a range of generic system parameters are summarised in Table 1. In our performance studies we used the COST207 [3] seven-path bad urban (BU) channel model, whose impulse response is portrayed in Figure 3.

Our initial experiments compared the performance of a whitening matched filter (WMF) for single user detection and the Minimum mean square error block decision feedback equalizer (MMSE-BDFE) for joint multi-user detection. These simulations were performed using 4-level Quadrature Amplitude Modulation (4QAM), invoking both binary BCH [5] and turbo coded [6] video packets. The associated bitrates are summarised in Table 2. The transmission bitrate of the 4QAM modem mode was 29.5Kbps, which was reduced due to the approximately half-rate BCH or turbo coding, plus the associated video packet acknowledgement feedback flag error control [1] and video packetisation overhead to produce effective video bitrates of 13.7Kbps and 11.1Kbps, respectively. A more detailed discussion on the video packet acknowledgement feedback error control and video packetisation overhead will be provided in Section 2 with reference to the convolutionally coded multi-mode investigations.

Figure 4 portrays the bit error ratio (BER) performance of the BCH coded video transceiver using both matched filtering and joint detection for 2–8 users. The bit error ratio is shown to increase, as the number of users increases, even

---

The financial support of the Mobile VCE, UK and that of the European Commission is gratefully acknowledged.
ACTS'99, June 1999, Sorrento, Italy
Key words: H.263, reconfigurable modulation, wireless video, video telephony, CDMA

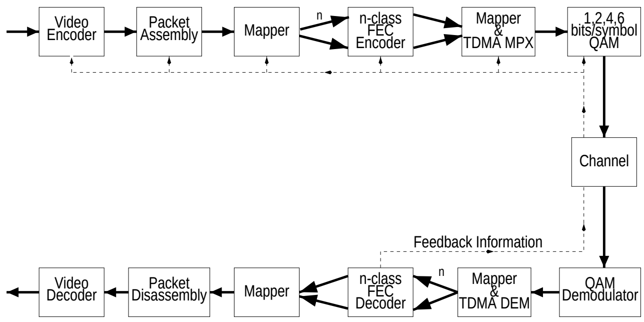

Figure 1: Reconfigurable transceiver schematic

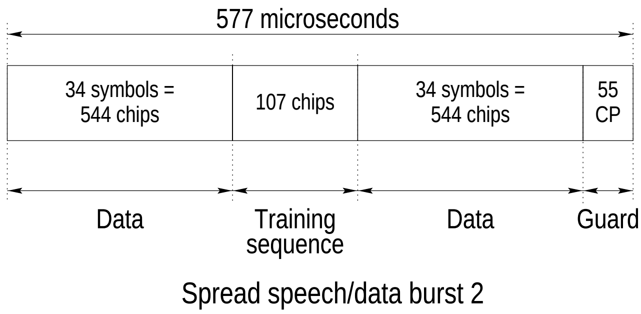

Figure 2: Transmission burst structure of the FMA1 spread speech/data mode 2 of the FRAMES proposal[2]

Parameter	
Multiple access	TDMA/CDMA
Channel type	C0ST 207 Bad Urban
Number of paths in channel	7
Normalised Doppler frequency	3.7×10^{-5} Hz
CDMA spreading factor	16
Spreading sequence	Random
Frame duration	4.615 ms
Burst duration	577 μs
Joint detection CDMA receiver	Whitening matched filter (WMF) or Minimum mean square error block decision feedback equalizer (MMSE-BDFE)
No. of Slots/Frame	8
TDMA frame length	4.615ms
TDMA slot length	577μs
TDMA slots/Video packet	3
Chip Periods/TDMA slot	1250
Data Symbols/TDMA slot	68
User Data Symbol Rate (kBd)	14.7
System Data Symbol Rate (kBd)	117.9

Table 1: Generic system parameters using the Frames spread speech/data mode 2 proposal [2]

Features	BCH coding	Turbo coding
Modulation	4QAM	
Transmission bitrate (kbit/s)	29.5	
Video-rate (kbit/s)	13.7	11.1
Video framerate (Hz)	10	

Table 2: FEC-protected and unprotected BCH and Turbo coded bitrates for the 4QAM transceiver mode

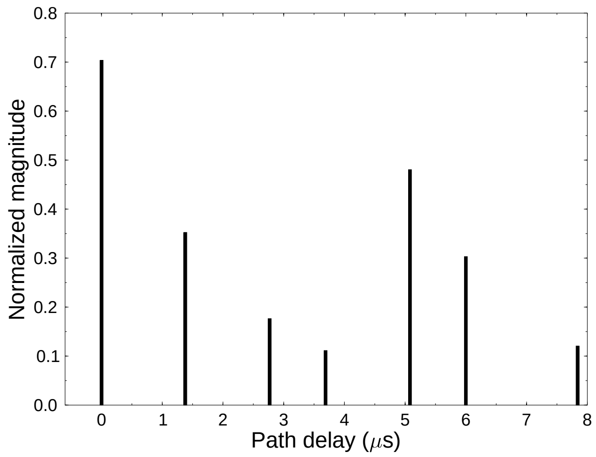

Figure 3: Normalized channel impulse response for the COST 207[3] seven-path Bad Urban channel.

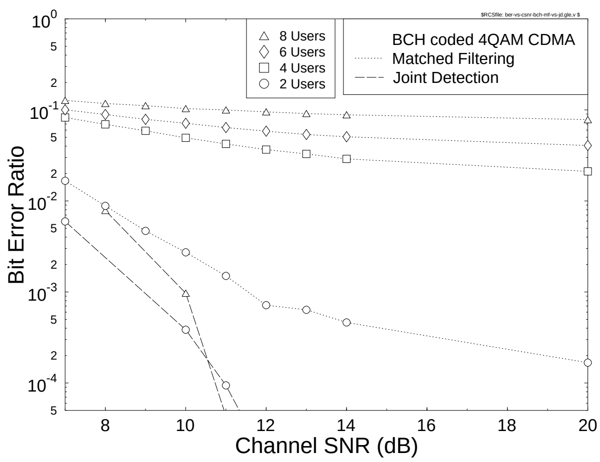

Figure 4: BER versus channel SNR 4QAM performance using BCH coded 13.7Kbps video, comparing the performance of matched filtering and joint detection for 2–8 users.

upon employing the MMSE-BDFE multi-user detector (MUD). However, while the matched filtering receiver exhibits an unacceptably high BER for supporting perceptually unimpaired video communications, the MUD exhibits a far superior BER performance.

When the BCH codec was replaced by the turbo-codec, the bit error ratio performance of both matched filtering and the MUD receiver improved, as shown in Figure 5. However, as expected, matched filtering was still outperformed by the joint detection scheme for the same number of users. Furthermore, the matched filtering performance degraded rapidly for more than two users.

Figure 6 shows the video packet loss ratio (PLR) for the turbo coded video stream using matched filtering and joint detection for 2–8 users. The figure clearly shows that the matched filter was only capable of meeting the target packet loss ratio of 5% for upto four users, when the channel SNR was in excess of 11dB. However, the joint detection algorithm guaranteed the required video packet loss ratio performance for 2–8 users in the entire range of channel SNRs shown. Furthermore, the 2-user matched-filtered PLR performance was close to the 8-user MUD PLR.

2. MULTI-MODE VIDEO SYSTEM PERFORMANCE

Having shown that joint detection can substantially improve our system's performance, we investigated the performance of a multi-mode convolutionally coded video system employing joint detection, while supporting two users. The associated convolutional codec parameters are summarised in Table 3.

Below we now detail the video packetisation method employed. The reader is reminded that the number of symbols per TDMA frame was 68 according to Table 1. In the 4QAM mode this would give 136 bits per TDMA frame. However, if we transmitted one video packet per TDMA frame, then the packetisation overhead would absorb a large percentage of the available bitrate. Hence we assembled larger video packets, thereby reducing the packetisation overhead and arranged for transmitting the contents of a video packet over three consecutive TDMA frames, as indicated in Table 1.

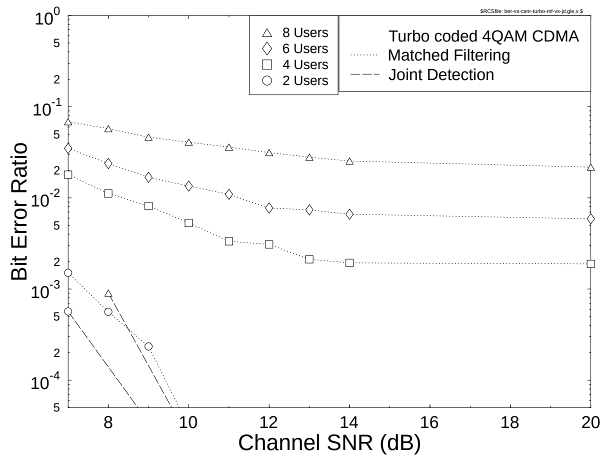

Figure 5: BER versus channel SNR 4QAM performance using turbo-coded 11.1Kbps video, comparing the performance of matched filtering and joint detection for 2–8 users.

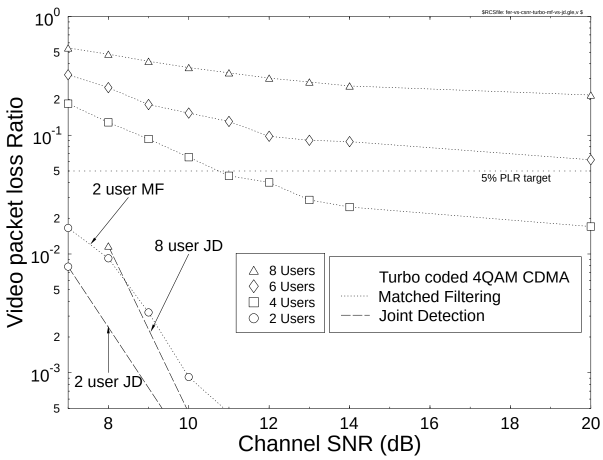

Figure 6: Video packet loss ratio versus channel SNR for the turbo-coded 11.1 Kbps video stream, comparing the performance of matched filtering and joint detection for 2–8 users.

Therefore each protected video packet consists of $68 \times 3 = 204$ modulation symbols, yielding a transmission bitrate of between 14.7 and 38.9 Kbps for BPSK and 16QAM, respectively. However, in order to protect the video data we employed half-rate convolution coding, using actally represented generator polynomials of 5 and 7. The useful video bitrate was further reduced due to the 16 bit Cyclic Redundancy Checking (CRC) used for error detections and the nine-bit repetition-coded feedback error flag for the reverse link. This leaves video packet sizes of 77, 179 and 383 bits for each of the three modulation modes. The useful video capacity was finally further reduced by the video packet header of between 8 and 10 bits, resulting in useful or effective video bitrates ranging from 5 to 26.9 Kbps in the BPSK and 16QAM modes, respectively.

The proposed multi-mode system can switch amongst the 1, 2 and 4 bit/symbol modulation schemes under network control, based upon the prevailing channel conditions. As seen in Table 3, when the channel is benign, the unprotected video bitrate will be approximately 26.9Kbps in the 16QAM mode. However, as the channel quality degrades, the modem will switch to the BPSK mode of operation, where the video bitrate drops to 5Kbps, and for maintaining a reasonable video quality, the video resolution has to be reduced to SQCIF (128x96 pels).

Figure 7 portrays the packet loss ratio for the multi-mode system, in each of its modulation modes for a range of channel SNRs. It can be seen that above 14dB the 16QAM mode offers an acceptable packet loss ratio of less than 5%, while providing an unprotected video rate of about 26.9Kbps. If the channel SNR drops below 14dB, the multi-mode system is switched to 4QAM and eventually to BPSK, when the channel SNR is below 9dB, in order to maintain the required quality of service, which is dictated by the packet loss ratio. The figure also shows the acknowledgement feedback error ratio (FBER) for a range of channel SNRs, which has to be substantially lower, than the video PLR itself. This requirement is satisfied in the figure, since the feedback errors only occur at extremely low channel SNRs, where the packet loss ratio is approximately 50%, and it is therefore assumed that the multi-mode system would have switched to a more robust modulation mode, before the feedback acknowledgement flag can become corrupted.

Features	Multi-rate System		
Mode	BPSK	4QAM	16QAM
Bits/Symbol	1	2	4
FEC	Convolutional Coding		
Transmitted bits/packet	204	408	816
Total bitrate (kbit/s)	14.7	29.5	58.9
FEC-coded bits/packet	102	204	408
Assigned to FEC-coding (kbit/s)	7.4	14.7	29.5
Error detection per packet	16 bit CRC		
Feedback bits / packet	9		
Video packet size	77	179	383
Packet header bits	8	9	10
Video bits/packet	69	170	373
Unprotected video-rate (kbit/s)	5.0	12.3	26.9
Video framerate (Hz)	10		

Table 3: Operational-mode specific transceiver parameters for the proposed multi-mode system

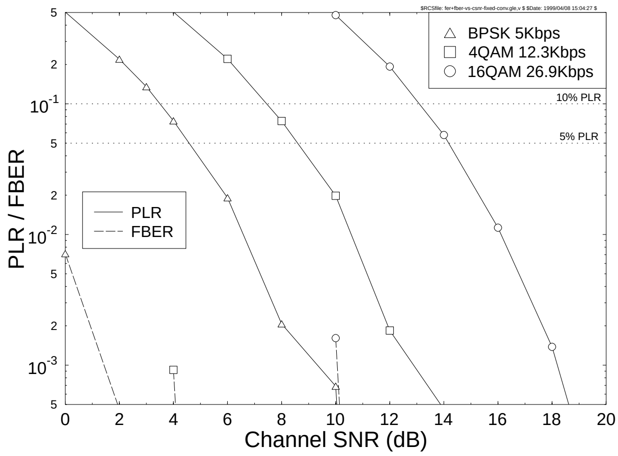

Figure 7: Video packet loss ratio (PLR) and feedback error ratio (FBER) versus channel SNR for the three modulation schemes of the 2-user multi-mode system using joint detection.

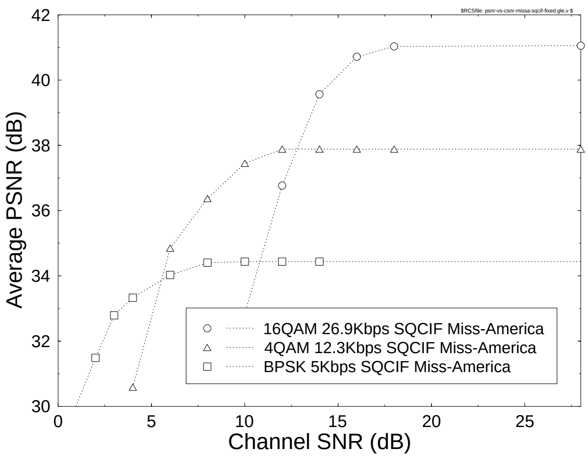

Figure 8: Decoded video quality (PSNR) versus channel SNR for the modulation modes of BPSK, 4QAM and 16QAM supporting 2-users with the aid of joint detection. These results were recorded for the Miss-America video sequence at SQCIF resolution (128x96 pels).

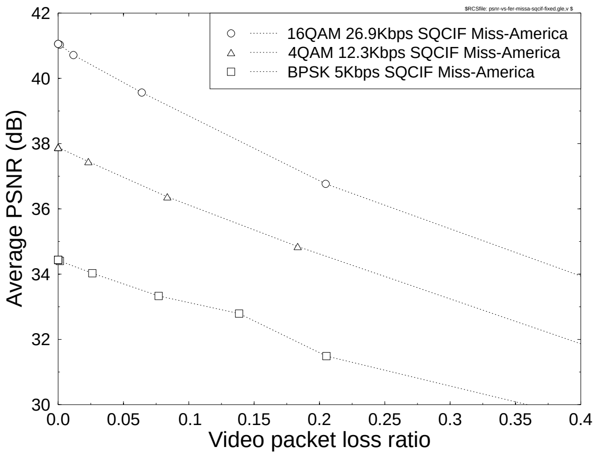

Figure 9: Decoded video quality (PSNR) versus video packet loss ratio for the modulation modes of BPSK, 4QAM and 16QAM, supporting 2-users with the aid of joint detection. The results were recorded for the Miss-America video sequence at SQCIF resolution (128x96 pels).

The video quality is commonly measured in terms of the peak-signal-to-noise-ratio (PSNR). Figure 8 shows the video quality in terms of the PSNR versus the channel SNRs for each of the modulation modes. As expected, the higher throughput bitrate of the 16QAM mode provides a better video quality. However, as the channel quality degrades, the video quality of the 16QAM mode is reduced and hence it becomes beneficial to switch from the 16QAM mode to 4QAM at an SNR of about 14dB, as it was suggested by the packet loss ratio performance of Figure 7. Although the video quality expressed in terms of PSNR is superior for the 16QAM mode in comparison to the 4QAM mode at channel SNRs in excess of 12dB, however, due to the excessive PLR the perceived video quality appears inferior in comparison to that of the 4QAM mode, even though the 16QAM PSNR is higher for channel SNRs in the range of 12–14dB. More specifically, we found that it was beneficial to switch to a more robust modulation scheme, when the PSNR was reduced by about 1dB with respect to its unimpaired PSNR value. This ensured that the packet losses did not become obvious, resulting in a higher perceived video quality, and smoother degradation as the channel quality deteriorated.

The effect of packet losses on the video quality quantified in terms of PSNR is portrayed in Figure 9. The figure shows, how the video quality degrades, as the PLR increases. It has been found that in order to ensure a seamless degradation of video quality as the channel SNR reduced, it was the best policy to switch to a more robust modulation scheme, when the PLR exceeded 5%. The figure clearly shows that a 5% packet loss ratio results in a loss of PSNR, when switching to a more robust modulation scheme. However, if the system did not switch until the PSNR of the more robust modulation mode was similar, the perceived video quality associated with the originally higher rate, but channel-impaired stream became inferior.

3. CONCLUSIONS

In conclusion, the proposed joint-detection assisted multimode CDMA-based video transceiver substantially outperformed the matched-filtering based transceiver. The transceiver guaranteed a near-unimpaired video quality for channel SNRs in excess of about 5 dB over the COST207 dispersive Rayleigh-faded channel. The benefits of the multimode video transceiver clearly manifest themselves in terms of supporting un-impaired video quality under time-variant channel conditions, where a single-mode transceiver's quality would become severely degraded by channel effects.

4. REFERENCES

[1] L. Hanzo, P. Cherriman, J. Streit: Video Compression and Communications over Wireless Channels: From Second to Third Generation Systems, WLANs and Beyond, John Wiley and IEEE Press, in preparation [1]

[2] A. Klein, R. Pirhonen, J. Skoeld, and R. Suoranta, "FRAMES multiple access mode 1 - wideband TDMA with and without spreading," in *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC'97*, vol. 1, (Marina Congress Centre, Helsinki, Finland), pp. 37–41, IEEE, 1–4 Sept 1997.

[3] "COST 207: Digital land mobile radio communications, final report." Office for Official Publications of the European Communities, 1989. Luxembourg.

[4] P. Cherriman, L. Hanzo: Programable H.263-based wireless video transceivers for interference-limited environments, IEEE Tr. on CSVT, June 1997, Vol. 8, No.3, pp 275-286

[5] R. Steele, L. Hanzo (Ed.), *Mobile Radio Communications*, 2nd edition, IEEE Press-John Wiley and Sons Ltd., 1999

[6] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding : Turbo Codes," in *Proceedings, IEEE International Confrence on Communications*, pp. 1064–1070, 1993.

[7] EE L. Kuan, L. Hanzo: Joint Detection CDMA Techniques for Third-generation Transceivers, *Proceedings of ACTS'98*, Rhodos, Greece, June 1998, pp 727-732

[1] For detailed contents please refer to http://www-mobile.ecs.soton.ac.uk