MULTI-MODE WIRELESS VIDEOPHONY - AN OVERVIEW

INVITED PAPER

L. Hanzo and J. Streit

Dept. of Electr. and Comp. Sc., Univ. of Southampton, SO17 1BJ, UK.
Tel: +44 1703 593 125, Fax: +44 1703 594 508
Email: lh@ecs.soton.ac.uk
http://www-mobile.ecs.soton.ac.uk

ABSTRACT

A comparative study of arbitrarily programmable, but fixed-rate videophone codecs using quarter common intermediate format (QCIF) video sequences scanned at 10 frames/s is presented. These codecs were designed to allow direct replacement of mobile radio voice codecs in second generation wireless systems, such as the Pan-European GSM, the American IS-54 and IS-95 as well as the Japanese systems, operating at 13, 8, 9.6 and 6.7 kbps, respectively, although better video quality is maintained over higher-rate, 32kbps cordless systems, such as the Japanese PHS and the European DECT and CT2 systems. From the range of codecs investigated, best overall performance was achieved by our vector-quantised codecs, followed by the discrete cosine transformed and the quad-tree-based schemes, which were characterised by the bitallocation schemes of Table 2. The associated video Peak Signal-to-Noise Ratio (PSNR) was around 30 dB, while the subjective video quality can be assessed under `http://www-mobile.ecs.soton.ac.uk`. A range of multimode wireless transceivers were also proposed, which are characterised by Table 3.

1. MOTIVATION

In recent years the concept of intelligent multi-mode, multimedia transceivers (IMMT) has emerged in the context of wireless systems [1]. Their aim is to provide the best possible compromise amongst a number of contradicting design factors, such as the power consumption of the hand-held portable station (PS), robustness against transmission errors, spectral efficiency, teletraffic capacity, audio/video quality and so forth [1]. Due to lack of space in this treatise we have to limit our discourse to a small subset of the associated wireless video transceiver design issues, referring the reader for a deeper exposure to the literature cited [4]-[10]. A further advantage of the IMMTs of the near future is that due to their flexibility they are likely to be able to reconfigure themselves in various operational modes in order to ensure backwards compatibility with existing, so-called second generation standard wireless systems, such as the Japanese Digital Cellular, the Pan-American IS-54 and IS-95 systems, as well as the Global System of Mobile Communications (GSM) standards. The basic features of these systems are outlined in Table 1.

In order to provide wireless videophony services in the context of these existing so-called second-generation wireless systems, an additional speech channel has to be allocated to videophone users for the transmission of the video information. In this contribution we review some of the associated fixed but programmable-rate video coding aspects. We note, however that a substantial amount of further work has to be carried out in the area of intelligent transducers and interfaces, such as appropriate miniaturized video cameras, low-consumption video displays and in particular in the field of ergonomic hand-held portable multi-media communicator construction and design.

Again, the corresponding speech channel rates of the above existing 2nd generation wireless systems are 6.7, 8, 9.6 and 13 kbps, respectively, as seen in Table 1. For these constant, but rather low video rates there are two possible types of video codecs, which can be realistically invoked, namely fixed-rate proprietary schemes [4]-[7] or the standard H.263 video codec in conjunction with an appropriate rate control scheme, using also adaptive packetisation and a strongly protected transmission packet acknowledgement flag [9]-[10]. Due to lack of space in this treatise we concentrate on the fixed-rate proprietary schemes and refer the interested reader to the literature for a detailed exposure to H.263-based transceivers [9]-[10], which are applicable to any of the above standard systems, but were characterised in GSM-like system environments.

In our former work we designed a range of video codecs [4]-[7] that are capable of operating at a scanning rate of 10 frames/s, while maintaining such low bit rates, provided that low-dynamic head-and-shoulder video-sequences of the 176x144-pixel so-called Quarter Common Intermediate Format (QCIF) or 128x96-pixel Sub-QCIF video resolution are employed. We note, however that for high-dynamic sequences the 32kbps typical speech bitrate of the cordless telephone systems of Table 1, such as the Japanese PHS, the Digital European Cordless Telephone (DECT) or the British CT2 system is more adequate in terms of video quality. Furthermore, the proposed programmable video

---

This treatise is complemented by a demonstration package portraying video sequences at various bit rates, which is downloadable from http://www-mobile.ecs.soton.ac.uk

The financial support of the following organisations is gratefully acknowledged: Motorola ECID, Swindon, UK; European Community, Brussels, Belgium; Engineering and Physical Sciences Research Council, Swindon, UK; Mobile Virtual Centre of Excellence, UK.

	Cellular Systems					Cordless Systems			
System	GSM	DCS-1800	Qualcomm CDMA	IS-54	JDC	CT2	DECT	PHS	PACS
Origin	Europe	Europe	USA	USA	Japan	UK.	Europe	Japan	USA
Voice+FEC Rate(kbps)	22.8	22.8	8/Var.	11.2	13	32	32	32	32
Unprotected Voice Rate(kbps)	13	13	1.2-9.6	7.95	6.7	32	32	32	32
Speech FEC	Conv. (2,1,5)	Conv. (2,1,5)	Conv. Fwd:(2,1,9) Rev: R=1/3	Conv. (2,1,5)	Conv. R=9/17	No	No	CRC	CRC

DCS-1800: GSM-like European system in the 1800 MHz band
IS-95: American CDMA system
PHS: Japanese Personal Handyphone System
JDC: Japanese Digital Cellular
CDMA: Code Division Multiple Access
IS-54: American Digital Advanced Mobile Phone System (DAMPS)
CT2: British Cordless Telephone System
PACS: Personal Access Communications System
GSM: Global System of Mobile Communications
DECT: Digital European Cordless Telephone

Table 1: Speech coding rates of second-generation mobile systems

codecs are capable of multirate operation in the forthcoming 3rd generation Universal Mobile Telecommunications System (UMTS) or the so-called IMT2000 system.

In this short treatise we cannot consider the performance of the proposed video codecs [7, 11] in all of the above 2nd and 3rd generation mobile radio systems, hence we will mainly concentrate on a GSM-like environment. The main goal is to describe the design philosophy of our fixed-rate 8 or 11.36 kbps prototype video codecs and document their performance using two characteristic fixed bit rates, namely 8 and 11.36kbps within the above mentioned typical speech coding rate range, namely around 10kbps. We will show that these 8 and 11.36 kbps codecs exhibit similar video quality, but different error resiliences. Furthermore, we will devote some attention to transmission robustness issues and briefly summarise the features of a range of appropriate transceivers. Speech source coding aspects are beyond the scope of this paper, the reader is referred to for example [2] for the choice of the appropriate speech codecs. Channel coding issues are addressed in Reference [13], while a detailed discussion of reconfigurable modulation is given for example in Chapter 13 of [12].

2. FIXED-RATE VIDEO COMPRESSION

In contrast to existing and forthcoming standard variable-rate video compression schemes, such as the H.261, H.263 and MPEG2, MPEG4 codecs, which rely on bandwidth-efficient but error-sensitive variable-length coding (VLC) techniques combined with a complex self-descriptive bitstream structure, our proposed fixed-rate codecs exhibit a more robust, regular bitstream and a constant bitrate. For the sake of improved robustness it is often advantageous to refrain from using variable-length coding. Hence in this treatise we attempt to offer an overview of a range of constant, but arbitrarily programmable-rate 176×144 pixel head-and-shoulder Quarter Common Intermediate Format (QCIF) video codecs specially designed for videotelephony over existing and future mobile radio speech systems on the basis of a recent research programme [4]-[11].

All candidate codecs share the schematic of Figure 1, although in the Figure Discrete Cosine Transform (DCT) based motion compensated error residual (MCER) coding is shown, as an example. Vector-Quantised (VQ) and Quad-Tree (QT) based MCER compression were the topic of [5] and [6], respectively. We note, however that we employ MC for a given 8×8 pixel block only, if it is deemed sufficiently beneficial in terms of MCER energy reduction. For most of the blocks the MCER entropy reduction does not justify the additional transmission overhead associated with the MV, a principle, which we refer to as cost-gain motivated quantisation [4]. In fact, our results suggested that using the

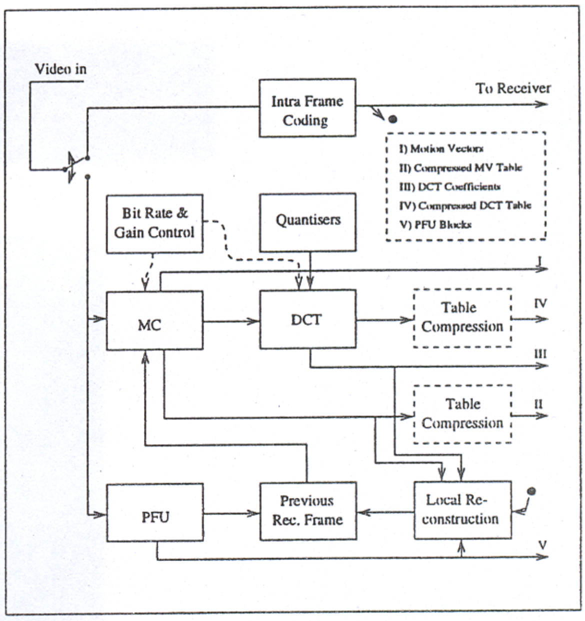

Figure 1: Schematic of the fixed- but programmable-rate DCT codec ©ETT 1997, Streit, Hanzo [7]

most important 30 active MVs out of the 396 blocks MVs was sufficient in order to achieve a near-maximum MCER-reduction. As seen in Table 2, for our 8kbps and 11.36kbps candidate codecs each of the 30 motion-active blocks' MVs were coded using 4 bits, while their position by 9 bits. For the remaining blocks simple frame-differencing was invoked, which does not require MVs. Observe in Table 2 that apart from the above 11.36 kbps DCT-based codec, which we refer to as DCTC2, we also contrived an 8kbps version of this scheme, DCTC1, by Variable-length Coding (VLC) both the motion-activity and the MCER-activity tables to around 350 bits. Lastly, following a similar philosophy, we also designed a VQ-based [5] and a QT-based [6] codec, which are again, characterised in Table 2. The associated transmission issues were also reported in References [4, 5, 6].

The various performance aspects of these codecs were compared in Reference [7], which are summarised in Figures 2, 3 and 4, suggesting the following conclusions:

- The proposed codecs exhibit a similar PSNR performance at 10 kbps to the MPEG-2 codec at 28 kbps. The H-261 codec at 22 kb/s has a PSNR of about 2 dB higher than the 10 kbps fixed-rate codecs.
- The inherent latency of the H-261 codec and that of the 10 kbps codecs is one frame or 100 ms due to MC.

Codec	FAW	PFU	MV Index + MV	DCT Ind. + DCT	VQ Ind. + VQ	QT + PC	Padding	Total
DCTC2	22	22×4	30×9 + 30×4	30× 9 + 30×12	-	-	6	1136
DCTC1	22	22×4	< 350 (VLC)	< 350 (VLC)	-	-	VLC	800
VQC2	22	22×4	38×9 + 38×4	-	31×9 + 31×8	-	5	1136
VQC1	22	22×4	< 350 (VLC)	-	< 350 (VLC)	-	VLC	800
QTC	22	20×4	< 500 (VLC)	-	-	< 565 + 1 or 80	VLC	1136

Table 2: Bit-allocation of fixed- but programmable-rate codecs ©ETT 1997, Streit, Hanzo [7]

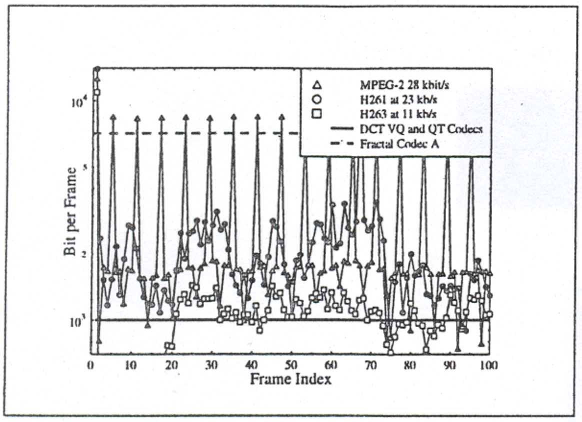

Figure 2: Bit-rate fluctuation versus frame index for the proposed adaptive codecs and the MPEG2, H.261 and H.263 standard codecs ©ETT 1997, Streit, Hanzo [7]

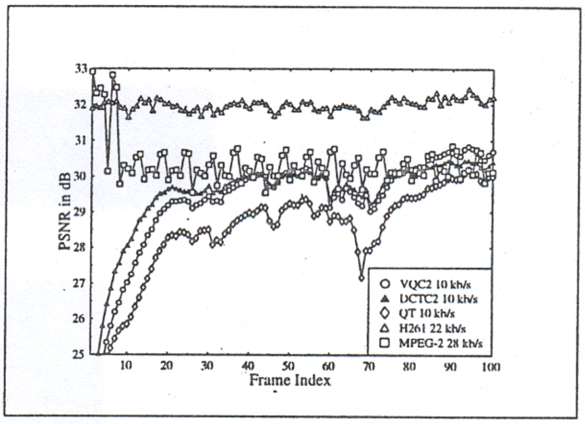

Figure 3: PSNR versus frame index performance of the proposed adaptive codecs and the MPEG2, H.261 and H.263 standard codecs ©ETT 1997, Streit, Hanzo [7]

The latency of the H.263 and MPEG-2 codecs may become significantly higher due to using so-called predicted or P-frames, relying on a number of previously received frames.

- The robustness of the codecs that employ active / passive table compression is quite limited, as is that of the VLC-based standard codecs. The slightly less bandwidth efficient schemes refraining from using table compression exhibit an improved error resilience.
- Overall, the vector quantised codecs VQC1 and VQC2 were shown to constitute the best compromise in terms of video quality, compression ratio and computational complexity, closely followed by the DCTC1 and DCTC2 candidate codecs [7]. The QT codec does not lend itself to robust implementations, since the QT code is rather vulnerable to channel errors, as it is evidenced by Figure 4.

3. TRANSMISSION ASPECTS

On the basis of our preference concerning the VQ codecs, we then contrived a suite of source-sensitivity matched programmable videophone transceivers, which were detailed in Reference [5]. The features of these systems are summarised in Table 3, although here due to lack of space we can only highlight these features with reference to [5].

Specifically, six different systems were designed, which used twin-class error correction coding, protecting the more vulnerable class 1 (C1) video bits more strongly than the less important C2 bits, whenever possible. This is however not always explicit in the Table, as seen for example for System 1, since we exploited that 16-level Quadrature Amplitude Modulation (16QAM) schemes possess two different integrity internal subchannels [12], which - when combined with various BCH codecs [13] - can be invoked for the high-integrity delivery of the more vulnerable video bits. Pilot symbol assisted QAM (PSAQAM) was used [12]. Adaptive QAM schemes, which can reconfigure themselves on a burst-by-burst basis and hence are amenable to IMT implementations were considered for example in Chapter 13 of [12].

All systems have a BCH-coded [13] rate of 20.32 kbps, implying that the lower video coding rates were more strongly BCH-coded, in order to resolve, as to whether it is worthwhile reducing the video codec's rate by variable-length coding, in order to be able to incorporate a stronger BCH code. In some of the scenarios Automatic Repeat Request (ARQ) was invoked, which acted upon either any errors, ie Class One/Class Two video bit errors, or only in case of errors in the more sensitive Class One video bit stream, noting that the packets, which are retransmitted, would occupy additional transmission slots and hence reduce the system's teletraffic capacity. This is also shown in the Table. The corresponding user Baud-rates are 18 or 9KBd, depending on whether 4QAM or 16QAM was invoked by the transceiver. At a multi-user signalling rate of 144KBd, which fits in the 200KHz channel bandwidth, 8 or 16 videophone users could be accommodated without ARQ. When occupying two slots per frame for ARQs, assuming that only one of the users can rely at any moment on re-transmission assistance, the number of users supported by Systems 2-

Feature	System 1	System 2	System 3	System 4	System 5	System 6
Video Codec	VQC1	VQC2	VQC1	VQC2	VQC2	VQC1
Video rate (kbps)	11.36	8	11.36	8	8	11.36
Frame Rate (fr/s)	10	10	10	10	10	10
C1 FEC	BCH(127,71,9)	BCH(127,50,13)	BCH(127,71,9)	BCH(127,50,13)	BCH(127,50,13)	BCH(127,71,9)
C2 FEC	BCH(127,71,9)	BCH(127,92,5)	BCH(127,71,9)	BCH(127,50,13)	BCH(127,50,13)	BCH(127,71,9)
Header FEC	BCH(127,50,13)	BCH(127,50,13)	BCH(127,50,13)	BCH(127,50,13)	BCH(127,50,13)	BCH(127,50,13)
FEC-coded Rate (kbps)	20.32	20.32	20.32	20.32	20.32	20.32
Modem	4/16-PSAQAM	4/16-PSAQAM	4/16-PSAQAM	4/16-PSAQAM	4/16-PSAQAM	4/16-PSAQAM
ARQ	None	Cl. One	Cl. One & Two	Cl. One & Two	None	Cl. One
User Signal. Rate (kBd)	18 or 9	9	18 or 9	18 or 9	18 or 9	9
System Signal. Rate (kBd)	144	144	144	144	144	144
System Bandwidth (kHz)	200	200	200	200	200	200
No. of Users	8-16	(16-2)=14	6-14	6-14	8-16	(16-2)=14
Eff. User Bandwidth (kHz)	25 or 12.5	14.3	33.3 or 14.3	33.3 or 14.3	33.3 or 14.3	14.3
Min. AWGN SNR (dB) 4/16QAM	5/11	11	4.5/10.5	6/11	8/12	12
Min. Rayleigh SNR (dB) 4/16QAM	10/22	15	9/18	9/17	13/19	17

Table 3: Summary of GSM-like video system features ©IEEE 1997, Streit, Hanzo [5]

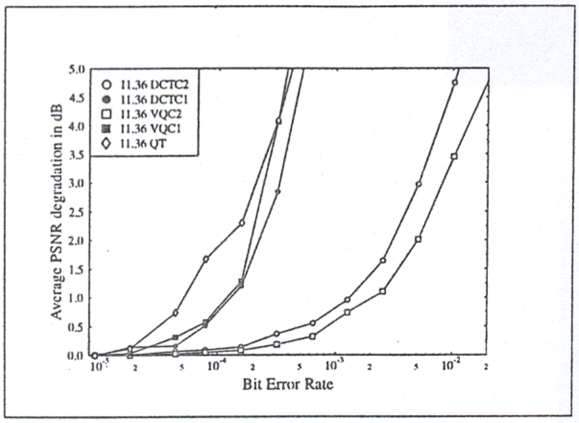

Figure 4: PSNR degradation versus BER for the proposed fixed-rate video codecs ©ETT 1997, Streit, Hanzo [7]

6 is hence reduced by two. The effective user bandwidth was computed by dividing the 200KHz bandwidth by the number of users supported. The minimum required channel SNR values over both Additive White Gaussian Noise (AWGN) and Rayleigh channels was recorded in the Table for second-order switch-diversity reception using the appropriate PSNR versus channel SNR curves of Reference [5], along with the maximum number of videophone users supported within the 200KHz GSM-bandwidth. A range of interesting conclusions can be drawn from Table 3, which were set out more explicitly in Reference [5].

4. SUMMARY AND CONCLUSION

Our discussions centered around the choice of fixed-, but programmable-rate video codecs for constant-rate wireless videophony. The VQ-codecs were found slightly superior to the other fixed-rate candidate codecs and hence the associated system design trade-offs and transmission aspects were summarised in Table 3 in the context of the favoured 8 and 11.36 kbps, 10 frames/s QCIF VQ-codecs. The associated head-and-shoulders videophone quality of various codecs can be studied under `http://www-mobile.ecs.soton.ac.uk` using an MPEG player. The interested reader is referred to [4]-[7] for further details on the various proprietary candidate codecs and for a discussion on the associated transmission aspects. A range of H.263-based videophone schemes were contrived in [9]-[10].

5. REFERENCES

[1] L. Hanzo: Bandwidth-efficient Wireless Multimedia Communications, Proc. of the IEEE, July 1998, Vol. 86, No. 7, pp 1342-1382

[2] *R.A. Salami, L. Hanzo, et. al*: Speech Coding, Chapter 3, pp 186-346, in R. Steele (Ed), *Mobile Radio Communications.* Pentech, 1992.

[3] L. Hanzo, F.C.A. Brooks, J.P. Woodard: Modern Voice Communications: Principles and applications for fixed and wireless channels, IEEE Press, in preparation [1]

[4] J. Streit, L. Hanzo: Vector-quantised low-rate cordless videophone systems, IEEE Tr. on Vehicular Technology, Vol. 42, No.2, May 1997, pp 340-357

[5] L. Hanzo, J. Streit: Adaptive low-rate wireless videophone systems, IEEE Tr. on Circuits and Systems for Video Technology, Aug. 1995, Vol. 5, No. 4, 305-319, ISSN 1051-8215

[6] J. Streit, L. Hanzo: Quadtree-based reconfigurable cordless videophone systems, IEEE Tr. on Circuits and Systems for Video Technology, Apr. 1996, Vol. 6, No. 2, pp 225-237

[7] J. Streit, L. Hanzo: Comparative study of programmable-rate videophone codecs for existing and future multimode wireless systems, European Transcations on Telecomms., Vol.8, Issue No 6, 1997, pp 551-572

[8] L. Hanzo, J. Streit: Multi-mode Wireless Videophony, Invited Paper, Journal of the Institute of Image Electronics Engineers of Japan (JIIEEJ), Japan, 1997, Vol. 26, No.6, pp 648-658

[9] 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

[10] P. Cherriman, T. Keller, L. Hanzo: Orthogonal Frequency Division Multiplex transmission of H.263 encoded video over highly frequency-selective wireless networks `http://www-mobile.ecs.soton.ac.uk/peter/robust-h263/robust.html`, to appear in IEEE Tr. on CSVT, 1999

[11] L. Hanzo, P. Cherriman, J. Streit: Modern Video Communications: Principles and applications for fixed and wireless channels, IEEE Press, in preparation [2]

[12] 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

[13] K.H.H. Wong, L. Hanzo: Channel Coding, pp 347-489, Chapter 4 in R. Steele (Ed.) Mobile Radio Communications,IEEE Press-Pentech Press, London, 1992 ISBN 07 273-1406-8

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

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