Performance Evaluation of H.263–Based Video Transmission in an Experimental Ad–Hoc Wireless LAN System*

Matías Freytes

Laboratorio de Comunicaciones Digitales
Universidad Nacional de Córdoba
Argentina
`mfreytes@gtwing.efn.uncor.edu`

Abstract. This paper analyzes different packetization strategies that significantly improve the quality of H.263 coded video transmission in wireless local area networks (WLANs). We show that a considerable improvement can be obtained with the proper combination of error concealment techniques and transmission unit (TU) sizes. Moreover, we present performance evaluation results on critical system parameters for interactive video over Ad–Hoc WLANs, and propose a simple rule to specify TU sizes. We use *Kinesis*, a novel system architecture for packet video, as a software measurement tool to analyze the effects of packetization policies, distance, network offered load, and interference from co–located WLAN devices on overall video quality. *Kinesis* supports IP multicast extensions, overcoming delay issues introduced by the complex retransmission schemes in the IEEE 802.11 MAC sublayer, which are not acceptable for real–time services. It implements real–time transport protocol functions to manage synchronization and QoS, and performs software–only real–time H.263 video encoding.

In this paper we address most common Ad–Hoc WLAN configurations, and present experimental results on Packet Error Rates, Frame Error Rates, frame delays and latency, and Peak Signal–to–Noise Ratio for well–known test video sequences.

1 Introduction

The increasing trend toward networked portable computers, and recent advances in WLAN technology and video compression algorithms, have stimulated the demand for real–time video transmission services over wireless packet switched networks. Real–time interactive video systems over Ad–Hoc WLANs constitute a different and complex scenario that requires further analysis and experimental research. Physical layers supported by the IEEE 802.11 standard for WLANs have limited connection ranges, present higher error rates and have time–varying and

---

* This work has been supported in part by SeCyT, Universidad Nacional de Córdoba, and by a grant from the Agencia Córdoba Ciencia.

E. Gregori et al. (Eds.): NETWORKING 2002, LNCS 2345, pp. 539–550, 2002.

asymmetric propagation properties. These characteristics have direct impact on performance degradation in real–time services: transmission systems in WLANs have to deal not only with network congestion as their wired counterparts, but also with corrupted packets due to higher bit error rates. To address these limitations, a complex Medium Access Control (MAC) protocol is required for adequate transport layer performance. The MAC protocols specified in IEEE 802.11 hide packet losses caused by bit errors by including a retransmission algorithm for corrupted packets which introduces significant delay and packet overhead [1].

Kinesis, a novel system architecture for packet video developed at our laboratory, overcomes this mechanism at the expense of higher frame dropping by forcing multicast addresses even in point–to–point communications. ARQ mechanisms introduce intolerable delays in interactive video systems. Although frame dropping results in video quality degradations, mechanisms are available to compensate for this effect, whereas nothing can be done about the excessive delays introduced by the IEEE 802.11 MAC ARQ. Link layer error checking (checksums) is useful for network services where no errors are tolerated, but it is too strict for applications where some degree of quality degradation is acceptable. Video applications should tolerate frame bit errors and be able to efficiently utilize non–corrupted regions of the bit stream, thus reducing information loss and enhancing overall video quality.

Recent publications investigate the performance of the transfer protocols proposed in this standard as well as voice and video services on packet switched infrastructure networks. In [2], Kamerman and Aben present throughput performance of 802.11 WLANs in relation to overhead from header fields in physical, medium access, and transport protocols. The authors focus on TCP file transfers in infrastructure WLANs, but do not address interference from co–located wireless networks, or real–time protocols and services. Weinmiller et al. ([3]) present comparative analysis on the performance of the access protocols in IEEE 802.11 and ETSI Hiperlan, regarding the impact of hidden stations, number of stations, and packet sizes. Quality of Service capabilities offered by point coordination function (PCF) access in 802.11 infrastructure networks are studied, but no results are presented for time–bounded services, or Ad–Hoc scenarios. Different performance–limitation factors in the two standards are considered in [4]. Two separate simulation scenarios are used, consisting of 10 and 100 stations organized in Ad-Hoc networks. Performance of MAC protocols is compared considering bit error rate (BER) of the fading air medium and the number of stations. However, no real–time experimental results are presented. Experimental throughput measurements comparing frequency hopping (FHSS) and direct sequence spread spectrum (DSSS) physical layers in several 802.11 infrastructure office environments are provided in [5]. Although the authors evaluate the effects of interference from adjacent WLANs, they do not address Ad–Hoc networks or time–bounded services. A high performance TCP protocol for lossy wireless links is presented in [6]. Congestion and random loss are differentiated, and window sizes, as well as TCP timers, are managed according to these two cases. Although this optimized protocol reports higher throughput and lower end–to–end delay, retransmission schemes in TCP are still too complex to be considered in time–bounded situations. Therefore, this is not a suitable solution for real–time interactive video applications. Bahl describes the challenges of supporting digital video in wireless networks in [7]. A custom video coding algorithm, a resource reservation scheme, and a software architecture are presented. The author focuses on the problem of providing high quality video on centralized wireless networks, but does not address Ad-Hoc configurations. Sachs et al. ([8]) propose an interesting hybrid ARQ system for streaming media over WLANs, and better performance results are reported for media servers attached through APs. However, this proposal is not suitable for interactive transmission due to buffering and decoding delays, and Ad-Hoc configurations are not discussed.

In this paper we analyze a packetization scheme that improves video quality at the receiver, and propose a simple rule to specify TU sizes. We also present performance evaluation results on critical system parameters obtained during *Kinesis* video sessions. We show the impact of distance between stations, interference from co–located WLANs, network offered load, and frame sizes on throughput and video quality. We measure Peak Signal–to–Noise Ratio (PSNR) and investigate the influence of INTRA and INTER frame errors by means of objective analysis.

The rest of this paper is organized as follows. In Section 2 we make a brief introduction to *Kinesis* and the general communication framework for networked multimedia. Section 3 discusses performance evaluation results, and in Section 4 we present our concluding remarks.

2 Kinesis

Kinesis is the real–time video transmission system developed at the Digital Communications Research Laboratory (DCRL). It was originally conceived as a software measurement tool to study the behavior of interactive video systems over WLANs. Initial successful results led to the development of a highly modular and extensible architecture, which would support RTP–based networked multimedia applications.

Kinesis is an object–oriented multi–threaded real–time video system made of autonomous and reusable modules. It incorporates our own implementation of the RTP protocols and our software–only H.263 video codec. It is easily extendable to accommodate new media types and offers an abstract interface to network connections. *Kinesis* supports IP multicast extensions, real–time video rendering on X terminals, video recording, and diverse video producers (frame grabbers, USB cameras, video disk files). Object oriented design and software construction are powerful tools to manage the complexity inherent in the development process of multimedia and networking distributed systems. Protocol modules, network connections and media codecs, for instance, are conveniently represented by system classes, providing a set of autonomous and reusable building blocks. Multi–threading led to the use of simplified abstract processors, in the form of active objects ([9] and [10]). It allows low latency on single–processor platforms, and efficient utilization of hardware resources on symmetric multi–processing computers. We have defined three such module–classes in a *Kinesis* session: media producer, media encoder, and media decoder, synchronized by media buffers, being the media encoder one of the most CPU–intensive tasks in a real–time video system.

2.1 Real–Time H.263 Video Encoder

H.263+ is the first international video coding standard specifically designed to work on different network technologies. It is a backward compatible extension of H.263 providing twelve optional modes to improve video quality in error-prone and non guaranteed QoS networks. Wenger et al [11] recommended and theoretically justified a combination of error-resilience optional modes for five scenarios based on wired networks. Although H.263+'s error-resilience modes were designed for wired networks they can be beneficial for wireless networks too, where video quality is severely degraded due to higher packet error rates.

Kinesis incorporates a new software–only H.263 encoder [12], which implements one of the most advanced discrete cosine transforms for fast video coding in interactive systems. Motion estimation is based on the Advanced Center Biased Three Step Search algorithm proposed in [13]. Compression tests on several well known video sequences report outstanding coding times. In Table 1 we present comparative results with the University of British Columbia's (UBC) H.263 Reference codec, version 3.2.

Table 1. Average video coding times (ms)

Frame type	Miss America		Susie		M & D		Foreman	
	DCRL	UBC	DCRL	UBC	DCRL	UBC	DCRL	UBC
INTRA	21.40	52.60	21.55	52.80	21.86	52.73	22.35	52.77
INTER	26.84	104.49	27.41	113.35	26.35	118.20	28.45	126.99

The video encoder incorporated in *Kinesis* performs much faster coding at the expense of a negligible impact on video quality, thus becoming an appropriate tool for interactive video conferencing systems. Although the current version only implements an H.263 video encoder, new media types can be easily added. Its flexible and modular architecture provides an excellent test–bed for new protocol proposals and makes it easily adaptable to new environments. In the next section we analyze performance results obtained during *Kinesis* video sessions.

3 Performance Evaluation

In this section we present experimental results obtained with *Kinesis* during video conferencing sessions over Ad–Hoc WLANs at the engineering building in our campus. We analyze the effects of video error concealment, data transmission unit sizes, offered network load, packet error rates (PER) and frame error rates (FER), INTRA and INTER frame losses, distance between stations, and interference from other WLAN devices in the same area, on the average video quality in wireless systems. We show how these system parameters affect the PSNR (1), as a measure of video quality. Although the average PSNR is an objective measure and may not reflect human perceived quality, it has been widely adopted as a distortion measure. In order to provide "subjective" information, we have also included some representative video frames for the case studies. We also include a detailed description of the equipment used, its set up, and measurement procedures.

In error–free transmissions, the PSNR of the video sequence reproduced at the receiver is given by:

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{N} \sum_{i=1}^{N} D_{sc}(i)} \ , \tag{1}$$

where N is the number of frames in the sequence, and D_{sc} is the video source coding distortion given by:

$$D_{sc} = \frac{1}{n \times m} \sum_{x=1}^{n} \sum_{y=1}^{m} \left| C_i(x,y) - \hat{C}_i(x,y) \right|^2 \ .$$

$C_i(x,y)$ and $\hat{C}_i(x,y)$ are the transmitted and received frames, (x,y) the pixel coordinates in the frame, and $(n \times m)$ the frame size. Transmission errors introduce additional distortion at the receiver, known as channel distortion (D_{ch}). Thus, the overall distortion of the decoded video sequence is given by $D_{sc} + D_{ch}$:

$$PSNR_d = 10 \log_{10} \frac{255^2}{\frac{1}{N} \sum_{i=1}^{N} D_{sc}(i) + D_{ch}(i)} \ , \tag{2}$$

The loss of picture quality, defined as:

$$\begin{aligned} \Delta PSNR &= PSNR_d - PSNR \\ &= 10 \log_{10} \frac{D_{sc}}{D_{sc} + D_{ch}} \ , \end{aligned} \tag{3}$$

is used as a measure of video degradation.

3.1 Experimental Environment

The experimental Ad-Hoc WLAN configured at the engineering building is conformed by Pentium III desktop and laptop PC computers running GNU/Linux. The stations are equipped with DSSS IEEE 802.11b wireless medium interfaces configured on channel 3 (2.422 GHz). Video producers consist on file producers of the well–known video test sequences "Mother&Dougther", "Carphone", "Foreman", and "Deadline". Each one of the results represent approximately

50000 QCIF frames (176 × 144 pixels) at 15 f/s. Measurements were taken on the second floor of the building, which consists primarily of office and laboratory rooms.

In wireless Ad-Hoc scenarios, transmission errors degrade video quality at the receiver. INTRA coded video sequences provide good streaming quality at the expense of inefficient bandwidth usage. There is no error propagation, and encoding algorithms are simpler and faster. Differential (INTER) coding, on the other hand, allows better use of the available bandwidth, but suffers from significant performance degradation. In order to mitigate the effects of transmission errors, a periodic INTRA refresh has been widely used. ITU–T H.263 specifies INTRA Macro Block (MB) coding once every other 132 times the MB is encoded, and several authors propose INTRA frame updates (e.g. [14],[15]), also known as Full Intra Refreshments (FIR).

H.263 provides means to insert synchronization words at the picture level, and, optionally, at the GOB level. The latter allows resynchronization in case of errors and inserts GOB sync headers at the beginning of each MB row. This is exploited by the error concealment technique used in the decoder, which discards corrupted GOBs and replaces the corresponding image content with data from the previously decoded frame. Although this technique has demonstrated very good results for non–moving parts of the sequence, it introduces noticeable distortion in moving image regions.

In Fig. 1 we show a typical ΔPSNR profile obtained during *Kinesis* video sessions. We plot ΔPSNR and packet errors for 1370 "Deadline" frames. As we can see, picture distortion caused by transmission errors (i.e. channel distortion introduced by dropped RTP packets) is usually more noticeable than the distortion produced by the propagation of errors inherent in differential coding techniques. This characteristic has severe consequences in packet switched

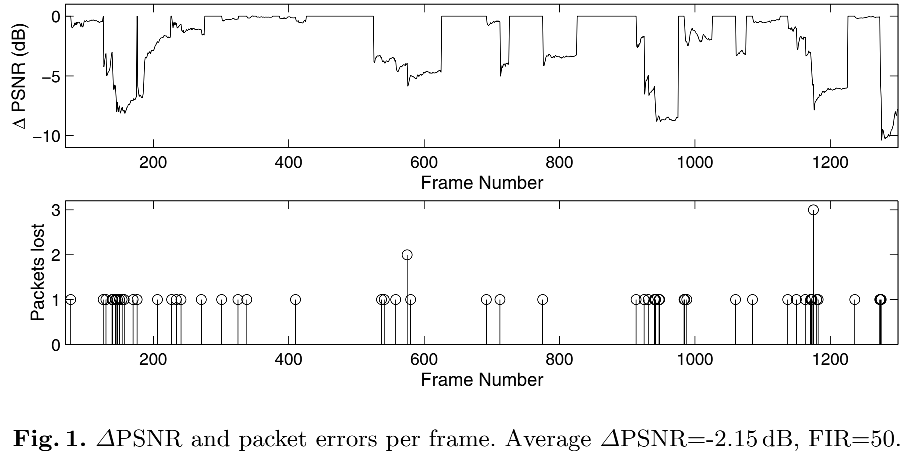


Fig. 1. ΔPSNR and packet errors per frame. Average ΔPSNR=-2.15 dB, FIR=50.

networks, where complete packets are generally lost by congestion on the network elements (e.g. routers), or at the local interfaces. Packet loss rates and their effects on overall video quality become critical in packet erasure channels with high bit error rates like the 802.11 Ad–Hoc WLAN under consideration. The consequences of transmission errors become more serious when corrupted data packets correspond to INTRA coded frames, producing deep PSNR peaks and seriously degrading quality. In Fig. 2 we show immediate error propagation effects after a lost frame compared to an error–free received sequence.


(a) (b) (c)

Fig. 2. Original frame 2(a), source–coding distortion 2(b), and source–coding plus channel distortion 2(c).

3.2 Experimental Results

The first set of experiments analyzes the effects of distance between Ad–Hoc 802.11 stations and interference from co–located WLAN systems on the average video quality for indoor applications. Towards this end we performed several measurements placing portable stations at 22 m, 44 m, 66 m, 88 m, and 110 m from each other. In all cases we verified average PSNR degradation less than 0.2 dB. The FER measured in INTRA coded sequences is an order of magnitude larger than the values obtained for INTER sequences due to larger frame sizes and longer bursts. However, less degradation is perceived because there is no error propagation. These results show that the effects of distance between stations in indoor configurations with line of sight is negligible.

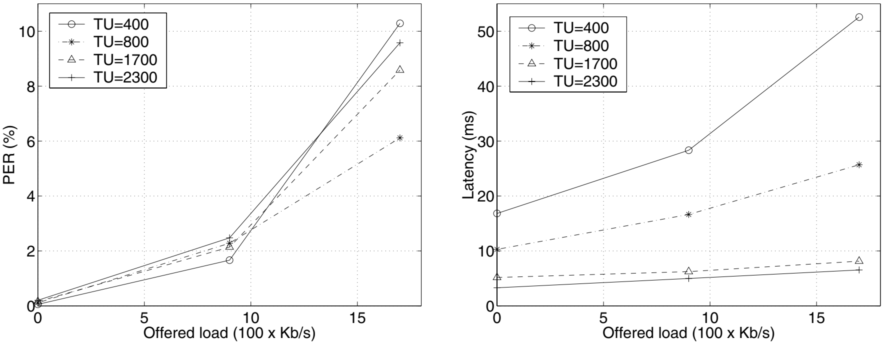
Then, we analyzed the influence of video transmission systems running over co–located 802.11 WLANs on FER and overall video quality. The interfering system was configured on channels 1, 5, 6, and 7, and the video session stations on channel 3 (2.422 GHz). In Table 2 we present average ΔPSNR and FER results for the tested channels. It should be noted that video quality is severely degraded when adjacent 802.11 channels are being used. These results demonstrate that a detailed planning scheme of channel reuse policies must be considered in order to provide acceptable digital video quality on Ad-Hoc 802.11 WLANs.

Table 2. Interference from adjacent WLANs

Channel	ΔPSNR (dB)	FER (%)
1 (2.412 GHz)	-12.84	53.35
5 (2.432 GHz)	-13.47	56.00
6 (2.437 GHz)	-0.41	0.6
7 (2.442 GHz)	-0.40	0.5

The second set of experiments analyzes the overall video quality at the receiver for different network loads TU sizes. We use INTRA and INTER coded sequences with FIRs at 44 and 132 frames. In both cases, we use a GOB replacement concealment mechanism to mitigate the effects of transmission errors. Our conclusions propose a simple rule to set the TU size for INTRA and INTER–coded real–time video in 802.11 WLAN systems.

INTRA–coded sequences. In Fig. 3(a) we show the effects of network load on the PER for different TU sizes. Our results show how small–TU error rates rapidly rise, even above medium and large TU values, as network load increases. At high offered loads, medium size TUs outperform small and large TUs. This is justified by the fact that INTRA–coded sequences, with large average GOB sizes (i.e. 300 Bytes) demand many small TUs to transport one video frame. Arrival time for all of the packets conforming one image is time bounded to guarantee acceptable decoding and rendering times at the receiver, thus forcing packet bursts at the transmitter (see Fig. 3(b)). Large TUs, on the other hand, require fewer packets to convey a full video frame. However, we also show that large TUs are more vulnerable, under low network traffic, than small ones due to wireless channel impairments.

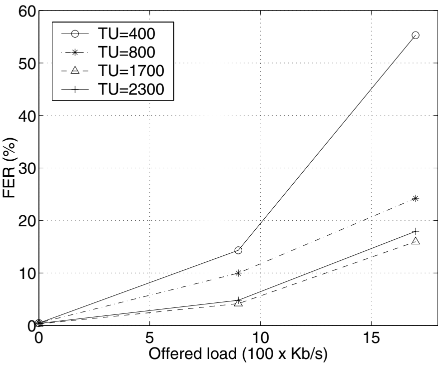


(a) PER vs. offered load.

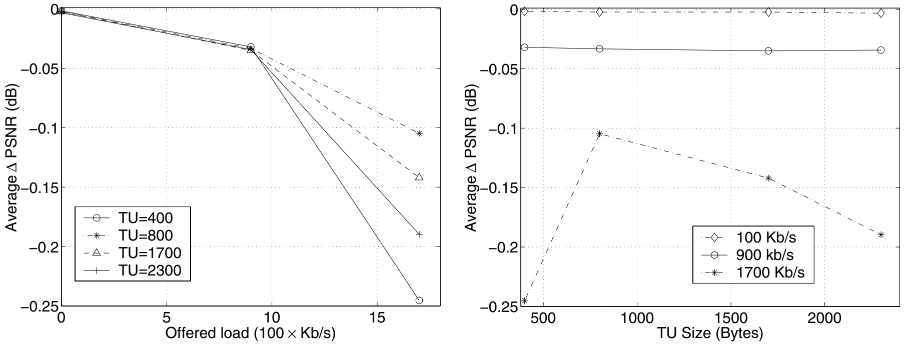
(b) Frame latency vs. offered load.

Fig. 3. PER and latency vs. offered load for INTRA–coded sequences.

Fig. 4 analyzes the overall ΔPSNR for the previous case studies. We verified that medium and large TUs outperform small ones by an almost negligible gain as offered load increases. Although high FERs have been reported under heavy network load conditions for small TU sizes (Fig. 4(a)), the ΔPSNR is still acceptable. INTRA–coded sequences guarantee no error propagation, and the error concealment technique previously described reduces the impact of lost packets, improving overall video quality. If we use frame level synchronization with a follow–on encapsulating packetization policy, a single bit error in the transmitted data will damage the whole frame since there are no resynchronization points. No error concealment methods are usually used, and the FER gives a complete idea of the received video quality. However, when using GOB–



(a) FER vs. offered load.



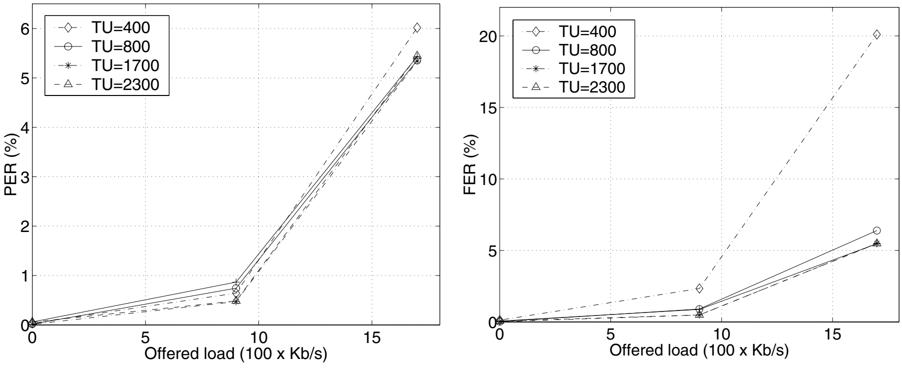
(b) ΔPSNR vs. offered load.

(c) ΔPSNR vs. TU size.

Fig. 4. ΔPSNR for INTRA–coded sequences.

level synchronization, although the FER could become larger, measured PSNR could indicate better performance results. Unlike the previous case, a bit error in the transmitted frame just damages the GOB containing the corrupted bit, and the receiver can resync at the next GOB. This way, only damaged GOBs have to be replaced with the corresponding GOBs from previous frames, keeping non–damaged GOBs in the current frame.

INTER–coded sequences. The following results correspond to INTER–coded sequences with FIRs at 44 and 132 frames, and sync at the GOB level. Although we allow large TUs, the whole unit would only be used in FIRs. INTER–coded frames are significantly smaller (average GOB sizes from 40 to 50 Bytes) and the TU adapts itself to these smaller sizes. For this reason, we introduce a smaller TU (100 Bytes) and the results are shown in Fig. 5. It should be noted that all of the TUs present similar performance results, except for TU=100, where packet transmission bursts can not be avoided.

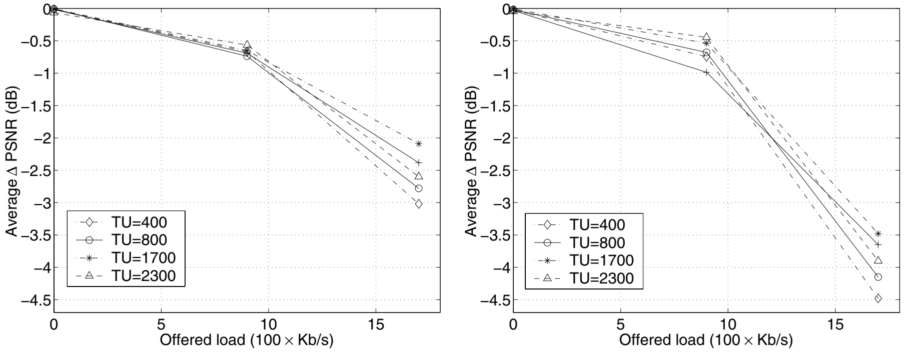


(a) PER vs. offered load. (b) FER vs. offered load.

Fig. 5. PER and FER vs. offered load for INTER–coded sequences (FIR=132).

In Fig. 6 we present the average ΔPSNR for INTER–coded video sequences. According to the results presented in Fig. 5(b), we verified that the overall video quality for medium and large TUs outperforms quality obtained with small TUs. Note that this difference becomes smaller as the offered load of the system decreases, and access to the transmission medium becomes more reliable.

When the system is in an unloaded condition, the channel distortion (D_{ch}) is negligible for all TU sizes since the FER is very small. However, ΔPSNR values for both, FIR=44 and FIR=132, report considerable degradation in overall video quality introduced by channel distortion as network offered load rises.



(a) FIR=44. (b) FIR=132.

Fig. 6. ΔPSNR vs. offered load for INTER–coded sequences.

At high load levels, an average advantage of 1 dB can be obtained for most common test sequences when selecting correct TU sizes. This video quality improvement becomes larger for complex sequences such as "Carphone", where we measured a 2 dB gain. TUs should be large enough to avoid packet bursts, and sufficiently small to overcome wireless channel impairments. The TU should be set to a multiple of the sync block size (i.e. GOB), avoiding frame encapsulation, although this incurs in longer frame latencies at the receiver.

4 Summary

In this paper we presented experimental results for real–time video transmission in Ad–Hoc WLAN systems. We have considered packetization schemes for several network load conditions and video coding policies, impact of distance between wireless stations, and interference from adjacent WLAN systems.

We have analyzed the effects of these system parameters on overall video quality and presented objective results based on ΔPSNR, FER, and frame latency. We have shown that a significant improvement can be achieved with proper error concealment techniques and packetization strategies, and we proposed a simple rule to set TU sizes. We have also discussed the effects of INTRA and INTER frame errors by means of objective and subjective analysis.

Frame error rates and their effects on overall video quality become critical in packet erasure channels with high bit error rates like the IEEE 802.11 WLANs. Current 802.11 MAC sublayer protocols implement an error checking mechanism (FCS) which completely discards whole frames when bit errors are detected. Interactive video applications over WLANs can take advantage of error correction and error resilience techniques like RESCU [16] to protect INTRA–coded frames or to alleviate error propagation. Networked video applications can partially compensate for higher bit error rates in wireless environments by capturing corrupted payload data and either correcting bit errors by means of a FEC scheme, or identifying non damaged regions. This implies an integral solution affecting link, transport and application layer protocols, together with new video coding techniques. We are currently working on 802.11 MAC error–checking policies and on lightweight real–time transport protocols to provide the necessary support for interactive video over WLAN networks.

References

1. IEEE Study Group 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. (1997) Draft Standard P802.11D5.3.
2. Kamerman, A., Aben, G.: Net throughput with ieee 802.11 wireless lans. In: IEEE Wireless Communications and Networking Conference (WCNC). (2000) 747–751
3. Weinmiller, J., Schlager, M., Festag, A., Wolisz, A.: Performance study of access control in wireless lans – ieee 802.11 dfwmac and etsi res 10 hiperlan. Mobile Networks and Applications **2** (1997) 55–67
4. Hadzi-Velkov, Z., Spasenovski, B.: Performance comparison of ieee 802.11 and etsi hiperlan type 1 under influence of burst noise channel. In: IEEE Wireless Communications and Networking Conference (WCNC). (2000) 1415–1420
5. Henty, B.E., Siew, J., Rappaport, T.S.: Frequency hop spread sperctrum and direct sequence spread spectrum study in ieee 802.11 and 802.11b wireless lan systems. Mobile and Portable Research Group – Virginia Tech (2001)
6. Parsa, C., Garcia-Luna-Aceves, J.J.: Differentiating congestion vs. random loss: A method for improving tcp performance over wireless links. In: IEEE Wireless Communications and Networking Conference (WCNC). (2000) 90–93
7. Bahl, P.: Supporting digital video in a managed wireless network. IEEE Communications Mag. (1998) 94–102
8. Sachs, D.G., Kozintsev, I., Yeung, M., Jones, D.L.: Hybrid arq for robust video streaming over wireless lans. In: Information Technology: Coding and Computing (ITCC). (2001) 317–321
9. Meyer, B.: Object-oriented Software Construction. Second edn. Prentice Hall International (1997)
10. Gibbs, S.: Composite multimedia and active objects. In: ACM SIGPLAN Workshop on Object-Based Concurrent Programming (OOPSLA). (1991) 97–112
11. Wenger, S., Knorr, G., Ott, J., Kossentini, F.: Error resilience support in h.263+. IEEE Transactions on Circuits and Systems for Video Technology **8** (1998) 867–877
12. Eschoyez, M., Freytes, M., Rodríguez, C.: Hardware–independent h.263 video encoder for interactive video conferencing systems. In: Multimedia Technology and Applications Conference (MTAC). (2001) 89–94
13. Nisar, H., Choi, T.S.: An advanced center biased three step search algorithm for motion estimation. In: IEEE International Conference on Multimedia and Expo. Volume 1. (2000) 95–98
14. Turletti, T., Huitema, C.: Videoconferencing on the internet. IEEE Trans. on Networking **4** (1996) 340–350
15. Chan, N., Mathiopoulos, P.: Efficient video transmission over correlated nakagami fading channels for is–95 cdma systems. IEEE JSAC **18** (2000) 996–1010
16. Rhee, I., Joshi, S.R.: Error recovery for interactive video transmission over the internet. IEEE JSAC **18** (2000) 1033–1049