

DIGITAL VIDEO IN A FADING INTERFERENCE WIRELESS ENVIRONMENT

Louis C. Yun and David G. Messerschmitt
 Department of Electrical Engineering and Computer Sciences
 University of California at Berkeley, CA 94720

ABSTRACT

Much research and development related to digital video has targeted storage, wired backbone, and broadcast wireless transport environments. Here we address video transport in a heterogeneous network, especially with interference- and fading-dominated wireless access, such as a cellular local-area or wide-area network. It is expected that multimedia (including video) will be an important application for desktop computers, and this suggests offering similar wireless capabilities. We note a number of deficiencies in existing video compression standards such as MPEG for this environment, and identify a number of challenging new research issues.

1 INTRODUCTION

Experience has shown that users welcome tetherless (no wires), nomadic (freedom in the point of network access), and mobile (access while moving) access to networking computing and communications services. Multimedia, and video in particular, should be no exception. Thus, our working hypothesis is that digital video via broadband backbone networks with wireless access deserves serious attention.

A transport network is rarely exclusively wireless, but usually incorporates wireless access to a broadband backbone network. Wireless access links present a severe bottleneck in provisioning video and multimedia services, and will only become more so as backbone networks continue to increase in speed. Thus, wireless access will become the primary bottleneck, and the entire system, including the backbone network and the video coding should be designed for maximum efficiency on wireless access links, where they exist. We can identify a number of interrelated objectives:

- high subjective quality and low perceptual delay (the latter primarily for the subset of interactive applications);
- high traffic capacity;
- multicast and unicast topology;
- mobility, including transparent inter-cell handoff;
- privacy by end-to-end encryption;
- scalability to heterogeneous terminals (from desktop to hand-held) and transport links (primarily broadband backbone and wireless access).

2 SYSTEM DESIGN ISSUES

Most existing digital video research and standardization focuses on bit rate minimization. A number of other important objectives for the wireless access in a heterogeneous net-

work point to the need for a complete rethinking of digital video coding for this environment.

2.1 Traffic capacity

If we admit a digital video stream to a wireless interference environment, there are three primary parameters of the stream that predict the traffic capacity it consumes, which we now cover in turn.

Bit rate. The distinguishing characteristics of the interference wireless environment are fading (changes in the propagation path as the terminal moves) and external interference from other terminals in the same cell (intra-cell interference) and/or other nearby cells (inter-cell interference). The shared interference precludes unilaterally increasing transmit power to compensate for other impairments, as in a broadcast environment. The primary scarce resource in an interference environment is thus radiated power (which represents interference to other users). The average radiated power is equal to the product of the average bit rate and the energy per bit. Thus, bit rate is one parameter that affects traffic capacity.

Reliability. For digital video in interactive applications with tight delay objectives (more on this later), we cannot afford reliable delivery of the video packets (as in data services), but rather must display corrupted data. The reliability as measured by error rate is an important parameter. The impact of fading and interference is a deteriorating channel reliability (as measured by error rate) as the aggregate traffic increases. Thus, the total traffic capacity is strongly dependent on the reliability objective. For example, the cost in capacity of an information stream in a wireless CDMA environment is the *product* of its reliability requirement (as specified by a signal-to-noise ratio) and its average bit rate [2]. As one benchmark -- in the worst case Rayleigh fading channel using sophisticated concatenated channel coding schemes -- the bandwidth penalty increases in proportion to twice the number of orders of magnitude decrease in error rate [3]. (This is in stark contrast to storage, wired, and non-interference dominated wireless broadcast environments, where reliability is normally not an issue or even a controllable parameter.)

Delay. Digital video is only one service out of many, and the wireless medium is usually only a portion of the total transport environment. The data portion of this traffic, including graphics, is typically very bursty in its rate characteristics. In such a heterogeneous environment, it is efficient to use packet transport with statistical multiplexing. The delay objectives of a given stream then become important, because re-

laxed delay gives more flexibility to choose *when* to transmit each packet. Packet scheduling can be traded for traffic smoothing and greater statistical multiplexing advantage on wireless access links [9].

2.2 Trade-off in bit rate and reliability

Minimizing bit rate in digital video coding is often at the expense of a greatly increased reliability requirement. This is appropriate for a wired and storage environment (where bandwidth and capacity can be expected to increase geometrically with time) but results in serious inefficiency on wireless access links (which will increasingly be the bottleneck). For example, with a compression ratio upwards of 100:1, MPEG-2 may at first glance appear ideal for bandlimited wireless links. However, MPEG bit error rate (BER) requirements are in the range of 10^{-9} to 10^{-12} [6]. Tetherless and mobile users may be more tolerant of impairments, and hence higher BERs (perhaps 10^{-7}) may be acceptable. In contrast, typical BERs on mobile wireless channels are orders of magnitude worse, 10^{-2} to 10^{-3} . Thus, forward error correction (FEC) and ARQ retransmission protocols are required to bridge the reliability gap, substantially lowering the effective compression ratio of MPEG on wireless links (by an order of magnitude or more on the worst-case Rayleigh channel), and also necessitating a more power-consuming receiver in the portable unit (to perform both the complex decompression and the FEC decoding).

2.3 Delay

In global networks, the round-trip propagation delay can be hundreds of milliseconds, which is in itself very significant for interactive applications (such as telephony and video conferencing). For these delay-sensitive applications there is little headroom for additional queuing or signal processing delay. This is doubly problematic for the traditional approach of minimizing bit rate. First, this increases the delay (since compression incorporates time averaging). Second, it increases reliability requirements, and techniques for achieving that reliability (such as ARQ retransmission protocols, interleaving, and forward error correction) inject additional delay.

In an unreliable transmission environment, it is advantageous to recognize the relaxed reliability (but tight delay) requirements of regions with rapid temporal changes, and the relaxed delay (but stringent reliability) requirements of stationary or close to stationary regions. The tight reliability objectives of the latter can be realized by a combination of FEC and ARQ techniques at the expense of delay. More generally, there are many opportunities to exploit the delay dimension in interactive applications to improve subjective quality and traffic capacity, but only if the video compression standards are more delay-cognizant [8][9].

2.4 Joint source/channel coding

For many years, the elegant source/channel separation theorem due to Shannon [11] has been taken too seriously,

because it does not apply to many practical situations [12] nor account for delay. In particular, for multiple access channels and fading channels with side information, the deliberate coordination of the source and transport can lead to substantial increases in traffic capacity. For example, if we design video compression algorithms that are robust in high error-rate environments, even at the expense of bit rate, we may achieve a better trade-off of capacity and delay.

To achieve the highest traffic throughput given subjective quality constraints, there must be a coordination of the source and transport in all the component quality-of-service (QOS) parameters: rate, reliability (corruption and loss), and delay. That is, as the video coder considers trade-offs among these parameters, while maintaining fixed subjective quality, it should consider the relative costs of provisioning each QOS component in the transport. For example, the objective of minimizing bit rate must be balanced against the penalties incurred in increased reliability requirements and delay, and this analysis must be specialized to each transport link. This is further complicated by the concatenation of two or more transport links in heterogeneous networks.

Even greater transport capacity can be achieved by *segmenting* the information in the video stream according to rate, reliability and delay needs, and making this source segmentation visible to the transport. In any video representation, not all data is equally susceptible to errors (or possibly delay as well). The wireless access can take advantage of this segmentation by performing a fine-grained resource allocation, making certain that no data receives a higher QOS (reliability or delay) than necessitated by subjective quality objectives [2]. This careful resource optimization is made practical by the relatively restricted bandwidth (and hence processing requirements) typical of wireless access.

2.5 Modularity

Joint source/channel coding is quite important, but should not be achieved at the expense of good system modularity. Such modularity is an important principle in complexity management and maintaining a flexible and extensible network design. Unfortunately, much past research assumed *tightly coupled* joint source-channel coding; examples include optimally tuning the bit allocation between specific source and channel coders to trade off rate against reliability (but without regard for delay) [4] and, in the case of vector quantization, directly incorporating the channel statistics in the codebook design [5]. This tight coupling is undesirable for two primary reasons:

- A joint source/channel design would have to be done for every pairing of source (e.g. consumer HDTV, interactive video conferencing) and transport (e.g. ATM, digital cellular). This becomes an N^2 solution and suffers from poor scalability.
- Where heterogeneous transport links are concatenated, tight coupling of the video coding and transport then neces-

sitates *transcoding* (the conversion of one coding standard to another) whenever information crosses from one link to the next; for example, at a wireless base station. Such transcoding is problematic in many ways. Here we only mention that it will introduce substantial delay (a critical resource in global networks) in signal processing and by forcing worst-case transport delay accumulation across concatenated links, and it also precludes privacy by end-to-end encryption.

2.6 The substream transport abstraction

In light of modularity considerations, transcoders should be avoided, video coding should be done only once at the edges of the network, and the transport should be transparent. But then the question becomes how to achieve joint source/channel coding. Fortunately, there is an architecture that combines joint source/channel coding and good modularity based on a *substream* abstraction of the transport (Fig. 1) [7].

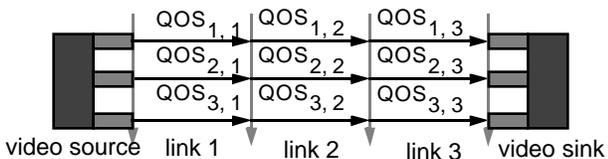


Figure 1. Joint source-channel coding across concatenated links via the substream architecture.

A substream is a flow of information characterized by its own rate, reliability and delay specifications. The transport then provides a different QOS to each substream. For concatenated transport entities, there is the freedom to disaggregate the end-to-end QOS of each substream into compliant QOS parameters for each link of that substream. Thus, each link of each substream potentially has a different set of QOS parameters, all of which may be negotiated at establishment.

To carry out joint source-channel coding, we segment the video source into multiple substreams. As a naive example, inter-frame coded video may be segmented into two substreams, one containing the motion vectors, the other the block differences. There will be a considerable difference in the reliability of these two segments. By employing substreams, the suboptimality inherent in practical source coders becomes an opportunity: the residual redundancy at the encoder output manifests itself as a range in the importance of the data it emits. The transport can appropriate more resources - for example, lower latency and stronger error protection - to the parts of the user information that are more important. Simultaneously, if the video coder knows the level of service that the transport network can provide down to the substream granularity, it may optimally reallocate information bits during source coding to match the transport capabilities. Hence, the substream abstraction supports both mechanisms of joint source-channel coding for increasing system capacity: it permits the matching of the QOS components of the source with the resource allocation trade-offs in the transport, and it

makes the fine grain segmentation of source information according to QOS needs visible to the transport.

The substream abstraction maintains good modularity because the coder flags its variable QOS requirements to the transport in a generic way, through a substream identifier on each packet. The network is not aware (in stark contrast to the transcoder approach) of the underlying semantics (video vs. audio) or syntax (particular compression standard) of the substreams it carries. Another manifestation of good modularity is that end-to-end encryption becomes feasible, as long as each substream is independently encrypted (no joint state). The correspondence between transport end-to-end QOS in the plaintext and encrypted signals is retained, and joint source/channel coding is not obstructed.

2.7 Multicast and scalable video

Wireless terminals, due to limitations in battery life, form factor, etc., will often have lower resolution capabilities than desktop terminals. With multiple video sinks tuned to a single source, the *simulcast* of a distinct video stream for each sink is not scalable, since the source bandwidth and processing power is directly related to the number of sinks. *Multicast*, in which the source outputs a single coded video stream and this stream is replicated within the network as necessary, solves this problem. However, it raises a different issue: how does a single source stream support a collection of sinks with possibly different transport access bandwidths and resolution and processing capabilities?

The multi-resolutions need to be embedded in the same stream, so that a multicast bridge can extract any of the representations. Again, for good modularity the bridge should not be cognizant of the semantic or syntactical content of the stream. The substream architecture elegantly achieves this objective. In addition to using substreams for segmenting information according to QOS, substreams also form a natural vehicle for layered, or multiresolution video representations. For example, the coarse and fine versions of a multiresolution video signal can be split into different substreams. A portable sink with limited display resolution can decode the coarse substream, whereas the high resolution terminal can decode all substreams. Alternatively, we can utilize substreams for simulcasting different rate versions of the video signal, should a gross disparity in the bandwidth provided to the receivers render layered coding techniques impractical. This support of multicasting and terminal heterogeneity is accomplished without compromising privacy by end-to-end encryption, provided the substreams are independently encrypted.

A key challenge is maintaining scalability in the process of QOS negotiation between a single source and multiple sinks. In a unicast environment, both source and sink may negotiate with the transport to obtain the desired trade-off between cost and performance. A direct extension of this approach to multicast -- that is, to have the source and all sinks jointly negotiate the best cost/service trade-off with the

transport each time a sink enters or leaves the session -- is complex and unscalable. The mobility of some sinks magnifies this problem: it is compelling that the source not be required to negotiate with both a changing set of sinks as well as a changing transport connection to some of these sinks. Maintaining scalability for multicast and mobile communications motivates sink-driven QOS negotiation strategies; that is, the source makes available a set of substreams embedding representations tailored to different transport bandwidths, reliability characteristics, and sink terminal capabilities, and leaves each individual sink to negotiate QOS with its dedicated portion of the transport and subscribe to the substreams of its choosing [7]. Solutions based on a combination of simulcast (two versions for backbone and wireless sinks) and multicast (supporting all terminals in each class) are feasible also.

3 VIDEO CODING FOR VARIABLE QOS

Wireless access suggests that simply minimizing bit rate may no longer be desirable, if the goal is to achieve the best trade-off between subjective quality and transport cost. The cost of storage and backbone bandwidth is decreasing rapidly, while reliability and delay requirements are at least as important as rate on wireless access links. Fig. 2 illustrates a generic video coder for future heterogeneous transport networks: it segments information into substreams by rate, reliability, and delay, embeds multi-resolution representations identified by substreams, and integrates rate monitor feedback to achieve rate targets. For an example, see the InfoPad [1] asynchronous video (ASV) [8][9]

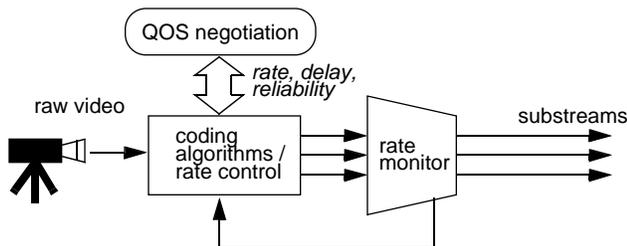


Figure 2. Video coding for variable QOS.

4 CONCLUSION

Assuming wireless access to multimedia networks is an important objective for the future, we have noted that two objectives are paramount: achieving low delay, and maximizing the traffic capacity of wireless access links. Other performance and cost parameters in the overall system will be relaxed over time by advances in electronics and photonics technologies.

The objective of minimizing delay, as well as enabling privacy by end-to-end encryption, implies the need to avoid transcoders and other converters within the network. This in turn places important requirements on the video coding at the source: it must be able to simultaneously scale to the capabilities of the multicast terminals (note the plural!) and their dif-

ferent access bandwidths and transport reliabilities.

High traffic capacity on the wireless link requires a coordination of video coding at the source with the concatenated transport links (again note the plural!) through joint source/channel coding in all dimensions: rate, reliability, and delay.

We have neglected other interesting issues, such as some aspects of mobility. How robust is the coding standard is to the disconnect-reconnect associated with handoff from one cell to another? A poorly designed coder and decoder pair will result in a high latency handoff or major hit in the video presentation.

Overall, there are many interesting research opportunities raised by new and sometimes conflicting requirements in heterogeneous networks. The field of video compression and related channel coding techniques will doubtless be revitalized by this heterogeneous network environment, with its myriad interrelated and challenging requirements.

The authors gratefully acknowledge Richard Han, Johnathan Reason, Junjing Yan and Yuan-Chi Chang for their collaboration on many of the ideas throughout this paper.

REFERENCES

- [1] S. Sheng *et al.*, "A portable multimedia terminal," *IEEE Comm. Mag.*, Vol. 30, No. 12, pp. 64-75, Dec. 1992.
- [2] L.C. Yun and D.G. Messerschmitt, "Variable quality of service in CDMA systems by statistical power control," *Proc. IEEE ICC*, Seattle, WA, June 18-22, 1995, Vol. 2, pp. 713-719.
- [3] J. Proakis, *Digital Communications*, McGraw Hill, 1989.
- [4] J.W. Modestino *et al.*, "Combined source channel coding of images using the block cosine transform," *IEEE Trans. Comm.*, Vol. 29, pp. 1262-1274, Sept. 1981.
- [5] K.A. Zeger and A. Gersho, "Vector quantizer design for memoryless noisy channels," *Proc. IEEE ICC'88*, Philadelphia, PA, June 12-15, Vol. 3, pp. 1593-1597.
- [6] S.-M. Lei, "Forward error-correction codes for MPEG2 over ATM," *IEEE Tran. Cir. Sys. for Video Tech.*, (4):2, April 1994.
- [7] P. Haskell and D.G. Messerschmitt, "In favor of an enhanced network interface for multimedia services", *IEEE Multimedia*, to appear.
- [8] A. Lao, J. Reason, D.G. Messerschmitt, "Layered asynchronous video for wireless services", *IEEE Workshop on Mobile Computing Systems and Applic.*, Santa Cruz, CA., Dec. 1994.
- [9] J.M. Reason, L.C. Yun, A.Y. Lao, D.G. Messerschmitt, "Asynchronous video: coordinated video coding and transport for heterogeneous networks with wireless access," *Mobile Wireless Information Systems*, Kluwer Academic Press, to appear.
- [10] K.S. Gilhousen *et al.*, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Tech.*, (40):2, pp. 303-312, May 1991.
- [11] C.E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, Vol. 27, pp. 379-423, 1948.
- [12] M. Khansari and M. Vetterli, "Transmission of sources with fidelity criteria over time-varying channels with side information," submitted to *IEEE Trans. Communications*, 1995.
- [13] R. Han and D.G. Messerschmitt, "Asymptotically Reliable Transport of Text/Graphics Over Wireless Channels", *Multimedia Computing and Networking*, San Jose, Jan. 29-31, 1996.