Multimedia Wireless Networking

Rajeev Jain, Abeer Alwan, Mario Gerla, Len Kleinrock, John Villasenor,
Ben Belzer, Walter Boring, Steve Molloy, Sean Nazareth, Marcio Siqueira, Joel Short, Jack Tsai

School of Engineering and Applied Sciences
University of California, Los Angeles*

ABSTRACT

Current wireless network systems (e.g. metropolitan cellular) are constrained by fixed bandwidth allocations and support only a narrow range of services (voice and low bit-rate data). To overcome these constraints and advance the state of the art in wireless multimedia communications, we are developing variable-rate video and speech compression algorithms, and wireless node architectures that will enable peer-to-peer multimedia networking even with very low bandwidth.

To support this objective, each wireless node must support new applications (for multimedia), advances in networking and source coding to support multimedia under limited bandwidth conditions (wireless), advances in physical layer design to support robust, low power, high packet throughput links, low power DSP for multimedia compression, and an architectural strategy to integrate these components into an efficient node. The algorithms and architectures to support this functionality are presented here, together with some preliminary results on network performance.

Keywords: multimedia networking, wireless systems

1. INTRODUCTION

Current wireless systems (e.g. metropolitan cellular) are constrained by fixed bandwidth allocations, fixed network configurations, and reliance on a tethered infrastructure. In addition, existing systems support only a narrow range of services (voice and low bit-rate data). To overcome these constraints and advance the state of the art in wireless multimedia communications, we are developing new multimedia applications, video/speech compression and network control algorithms, and wireless node architectures that will enable peer-to-peer multimedia networking amongst an arbitrary grouping of mobile nodes. To support this objective, each wireless node must support the functionality shown in Fig. 1, i.e. new applications (for multimedia), advances in networking to support mobility and multimedia under limited bandwidth conditions (wireless), advances in physical layer design to support robust, low power, high packet throughput links, low power DSP for multimedia compression, and an architectural strategy to integrate these components into an efficient node.The algorithms and architectures to support this functionality are presented below, together with some preliminary results on network performance.

2. ALGORITHMS

2.1. Video compression

The goals of an adaptive wireless video compression algorithm include 1) low hardware complexity, 2) rate-adaptivity, and 3) robustness against high bit error rates (BERs). In order to reduce hardware complexity and avoid inter-frame error propagation, we implement video compression without motion compensation. Our coding scheme consists of a 6 level discrete wavelet transform (DWT) followed by scalar quantization, run-length coding (RLC), and Huffman coding, with error protection provided by rate-compatible punctured codes (RCPC). Each of these steps has been modified as described below to achieve the desired functionality in the wireless network.

Since the DWT represents the most computationally intensive step in a subband coder, the length and nature (floating point versus integer) of the filter coefficients have a significant effect on the overall power requirements. Using new filter evaluation methods, we have identified several filters that are short (a total of 8 low pass and high pass taps) and have integer coefficients,[1] resulting in a reduction of chip area and power consumption by a factor of three (important for hand-held wireless nodes) compared to the

***** This work is being supported by ARPA/CSTO under contract #J-FBI-93-112

more commonly used floating-point filters.[2] The compression performance of these filters is within 0.5 dB peak SNR of that of the conventional 16-tap filter bank.

Bandwidth adaptation, i.e. coding rate adaptivity, is achieved by varying the quantizer Q-factor, which normalizes all subband quantizer step sizes. We have chosen scalar quantization, performed adaptively on a subband-specific basis based on a “Q”-factor similar to that used in the JPEG standard. The quantizer is simplified for low-power implementation by choosing the quantization step sizes Q_i for each subband as a function of the subband average absolute values A_i rather than the (theoretically optimal) subband variances. After normalization by the user “Q-factor”, the Q_i are rounded to a power of two so that a shift rather than a multiply performs the actual quantization. The resulting quantizer implementation is multiplierless, and requires only 1 add and 7 shifts per pixel. The average reduction in image quality due to the simplifying approximations in the quantizer is about 0.3 dB PSNR. We use a table of 32 Q-factors capable of achieving bit-rates of between 0.1 and 0.8 bpp over a wide range of 8 bit greyscale images; rate control to within 0.02 bpp is achieved by feedback from previously transmitted frames (Figure 2). While bandwidth adaptation may be counterproductive in fast-varying channels because the resulting rapid changes in video quality might be annoying to the user, such adaptation is certainly of value in slow-varying channels because it enables the video quality to slowly improve as available channel bandwidth increases. In addition, bandwidth adaptation is necessary even for constant bandwidth channels in order to produce a constant output rate while the statistics of the input video are changing.

To achieve the high compression ratios required for 60 kbps video, we use RLC followed by Huffman coding. The two variable length codes together give lossless compression of about 6:1. To protect against decoder sync errors when the variable-length codewords are corrupted by channel noise, we use a non-variable length coded End-of-Frame marker to ensure frame sync, and a unique End-of-Subband symbol, EOS. The EOS symbol prevents sync errors from propagating across subbands. To ensure a low bit error rate in the received bitstream, we use rate-compatible punctured convolutional (RCPC) codes that are punctured in accordance with the importance of the subbands. The compression system, including the shifted integer filters, the simplified quantizer, entropy coding and hard-decision RCPC codes gives 31.4 dB PSNR on the 512x512 greyscale “lena” image at 0.201 bits per pixel, and gives intelligible video at channel BERs as high as 0.02. An example of coded video from the 256x256 greyscale “football” sequence is shown in Figure 3.

2.2. Voice compression

Speech codec design is typically driven by bandwidth efficiency considerations; CELP-based coders, for example, are popular because of their low bit rates. The performance of these coders, however, deteriorates significantly in the presence of background noise. While CELP-based coders may be adequate for telephonic applications, future applications such as wireless multimedia personal communication systems will demand high quality speech under varying channel conditions. As a result, new standards for personal communication services are likely to use high quality, medium bit-rate speech codecs such as the one we are developing.

Our coder[4] is an embedded, perceptually based, subband coder, and has the following architecture: a low-delay IIR analysis/synthesis filterbank, a perceptual metric, an adaptive bit-allocation scheme, and an embedded quantization scheme. The coder is fully scalable --- increasing the bit rates improves the quality of encoded speech. In the current implementation, the coder processes input frames of 160 samples (20 ms at 8 kHz). The analysis/synthesis filterbank is an 8-channel IIR QMF bank, which is designed using the audibility of phase distortion as a cost function The perceptual metric ensures that encoding is optimized to the human listener, and is based on calculating the signal-to-noise ratio in short-time frames of the input signal. The bit-allocation scheme translates the signal-to-quantization noise ratio prescribed by the metric into a bit assignment. Finally, the subband samples are quantized with an embedded scalar quantizer.

Subjective listening tests, using quiet and noisy input signals, indicate that the proposed coder produces high-quality speech when operating at 12 kbps or higher. In error-free conditions, our coder has comparable performance to that of QCELP or GSM coders. For speech in background noise, however, our coder outperforms QCELP significantly at 12 kbps, and for music, it outperforms both QCELP and GSM.

2.3. Network protocols

Given an untethered mobile wireless environment and multimedia application requirements, the network protocol architecture must satisfy the following requirements:

1. data/voice/video traffic support
2. mobile, multihop networking
3. ability to handle rapidly changing radio channel conditions
4. protection from network congestion
5. graceful performance degradation as load increases, radio channel deteriorates and topology changes

Traditional radio networks can satisfy only a subset of these requirements. For example, the ARPA Packet Radio network could handle a rapidly deployable, multihop infrastructure, with reliable datagram transmissions; but, it could not support guaranteed bandwidth, multimedia traffic. In contrast, present cellular systems are based on a fixed infrastructure and provide circuit switched type service (mainly voice), without efficient datagram support. The challenge in our project is to combine instant reconfiguration, datagram support of packet radio, and guaranteed bandwidth, connection oriented service of cellular systems.

To meet this challenge, several alternative network approaches have been developed and evaluated. In this paper, we will limit our discussion to the TDMA Cluster approach,[4] which was selected for early implementation for its efficiency and simplicity. In the full paper, we will present an overview and comparison of all the alternatives.[5] The key features of the TDMA Cluster scheme are: (1) partitioning of the network into clusters, with code separation among clusters to reduce interference (see Fig 4); (2) time-division access in each cluster; namely, a slotted TDMA frame supports slot allocation (and thus guaranteed bandwidth) for real time voice/video connections, and random access for datagrams; (3) dynamic routing with shortest path computation and bandwidth/channel quality reporting to the source; this enables call admission control at the source, adaptive rate control and adaptive source coding of voice/video sources; (4) dynamic reconfiguration of clusters and routes to adjust to node movements; (5) support of Virtual Circuit connections for voice/video with guaranteed bandwidth and with dynamic rerouting (using Fast Reservations) in case of path disruption due to mobility; (6) dynamic rate reduction of voice/video streams at intermediate nodes via selective discarding of packets (based on priority defined by the hierarchical encoding scheme) in order to adjust to increased loads during path rerouting.

3. Wireless Multimedia Node Software/Hardware Architecture

3.1. Network operating system: WAMISNOS

WAMISNOS is a Network Operating System (NOS) that implements the network control algorithms such as clustering (described above). It is a multi-tasking operating system which runs on a PC-based computer running DOS. The WAMIS Network Operating System looks like an application in DOS. All of the kernel functions, applications, and networking protocols are compiled together into this application. Any protocol that is written for WAMISNOS is actually part of WAMISNOS. WAMISNOS is a command line based operating system, similar to DOS, but has multi-tasking capabilities and supports common network functionality. With WAMISNOS, the user is able to start an application from the command line prompt, as well as initiate or terminate any background process or network protocol. Each network protocol written for WAMISNOS can be its own process. However, a protocol is able to span multiple processes, or a process can contain multiple protocols. All of the scheduling and task switching functions are controlled by the WAMISNOS kernel. More details on the available library functions can be found in the appendix of.Ref. [7]. The WAMIS Network Operating System brings together the various components of the wireless multimedia node so that the applications, custom as well as standard networking algorithms and protocols, and communications hardware can communicate efficiently. The WAMISNOS kernel provides the underlying operating system functionality and the interfaces which allow the system to be flexible for many different configurations and environments. In Ref. [7] we find the steps necessary for a programmer to add a networking protocol into WAMISNOS.

3.2. Node architecture

The first generation wireless node architecture has been implemented on a PC notebook platform (Fig. 5), and consists of a spread spectrum radio,[8] a video compression/decompression system implemented with FPGAs,[9] and a commercially available video digitizer, each connected to the notebook PC through the ISA bus. The CPU executes the network operating system described above. This approach allowed for the partitioning of the node into a set of functional blocks which are optimized for functionality in isolation of other blocks. While a shared-bus node architecture provides modularity and compatibility with existing peripherals, it does not allow system-level optimization of power dissipation or throughput. Power dissipation is increased due to constant communication over the shared-bus, redundant bus interfaces required on every card on the bus, and the lack of hardware sharing between functional blocks. Throughput is also significantly degraded by forcing all data communication over a single shared resource, increasing bus contention. This becomes a major bottleneck as the available bus bandwidth quickly saturates when real-time multimedia data is exchanged between functional blocks. This node has been used to implement video/data/speech peer-to-peer networking, as described below.

4. APPLICATIONS

4.3. Video/data networking

Video Talk (VTalk) is a multimedia application implemented on WAMISNOS, which can set up a video and data transfer link between a pair of nodes in the wireless network. The video is sent using the User Datagram Protocol (UDP) which is a connection-less non-guaranteed delivery (but error free) transport protocol. Specification on the rate and size of the video image are accepted and statistics about the link and video rate are shown in the upper right hand corner of the display (see Fig. 6). Below the video display, there are two boxes in which a “chat” session can be conducted that uses TCP for transfer of the characters typed in at the keyboard. TCP packages a set of characters together into a packet and sends it on the link. TCP is a connection-oriented guaranteed error-free transport protocol. If any of the characters are lost in a packet, the characters are retransmitted in subsequent packets. No characters are displayed until they have been verified to be in order and error free. This protocol uses a built in flow control sliding window protocol which helps to adapt for congestion in the network during such peak traffic conditions as file transfers.

The current version of VTalk handles full-duplex video and a full duplex “chat” (data) session. With the Proxim frequency-hopped spread spectrum radio, the video throughput varies from .8 frames/sec to 3 frames/sec, depending on the frame size. With the clustering algorithm running, the frame rate drops significantly because of the fixed TDMA structure built into the algorithm. The VTalk application is then only allowed to send 1 segment of a video frame every TDMA frame time, which is 1200 milliseconds. It usually takes 3 segments to make up an average size video frame, therefore vtalk would get about 1 video frame every 3600 milliseconds. Methods to improve the performance while supporting the clustering and TDMA scheme are under investigation.

4.4. Speech networking

The speech coder is currently being implemented using a TMS320C50 Digital Signal Processor that is interfaced with a PC by a dual-port memory which is used for transferring data and status flags. Protocol routines have been developed on both the DSP and the PC for both speech coding and decoding. The first version of the variable-rate coder achieves near-toll quality at 16 Kbps and higher, and currently uses packet sizes of 256 bytes. With the proposed node hardware/software architecture, it is possible to guarantee constant throughput with a constant delay in the order of 1.5s. This delay is higher than 1.2s, the least-noticeable delay value for two-way conversations. The current implementation is therefore useful for half-duplex transmission. Methods to reduce the delay for full-duplex communications are being investigated.

5. ACKNOWLEDGEMENT

The authors gratefully acknowledge the technical contributions of Greg Pottie, Ben Tang, Albert Shen, and Brian Schoner

7. REFERENCES

[1] J. Villasenor, B. Belzer, and J. Liao, “Wavelet filter evaluation for efficient image compression”, *IEEE Trans. Image. Proc.*, vol. 4, pp. 1053-1060, August 1995.

[2] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transform”, *IEEE Trans. on Image Proc.*, vol. 1, pp. 205-220, 1992.

[3] Lelewer, D.A. and Hirshberg, D.S., “Data Compression”, *ACM Computing Surveys* **19**, pp. 261-296, 1987.

[4] A. Shen, B. Tang, A. Alwan, and G. Pottie, “A Robust and Variable-Rate Speech Coder”, *Proc. Int. Conf. Acoustic Speech SIgnal Processing (ICASSP) 1995*, Vol. 1, pp. 249-252.

[5] M. Gerla and J. Tsai “Multicluster, mobile, multimedia radio network”, to appear in *ACM-Baltzer Journal of Wireless Network*, 1995.

[6] N.D. Bambos, S.C. Chen, and G.J. Pottie “Radio link admission algorithms for wireless networks with power control and active link quality protection”, in IEEE *INFOCOM*, pp.97-104, 1995.

[7] C.R. Lin and M. Gerla “Multimedia transport in multihop dynamic packet radio networks”, to appear in *Proceedings of 1995 International Conference on Network Protocols*.

[8] W. Boring, J. Short, "UCLA WAMISNOS Network Protocol Programmers Guide", UCLA Computer Science Dept.

Technical Report 950010, April 1995.
[9] C. Chien, R. Jain, E.G. Cohen, and H. Samueli, “A Single-Chip 12.7 Mchips/s Digital IF BPSK Direct Sequence Spread-Spectrum Transceiver in 1.2 µm CMOS”, *JSSC*, Vol. 29, no. 12, December 1994.
[10] B. Schoner, J. Villasenor, S. Molloy and R. Jain, “Techniques for FPGA Implementation of Video Compression Systems,” ACM/SIGDA International Symposium on Field-Programmable Gate Arrays, Monterey, CA, 1995.

8. FIGURES

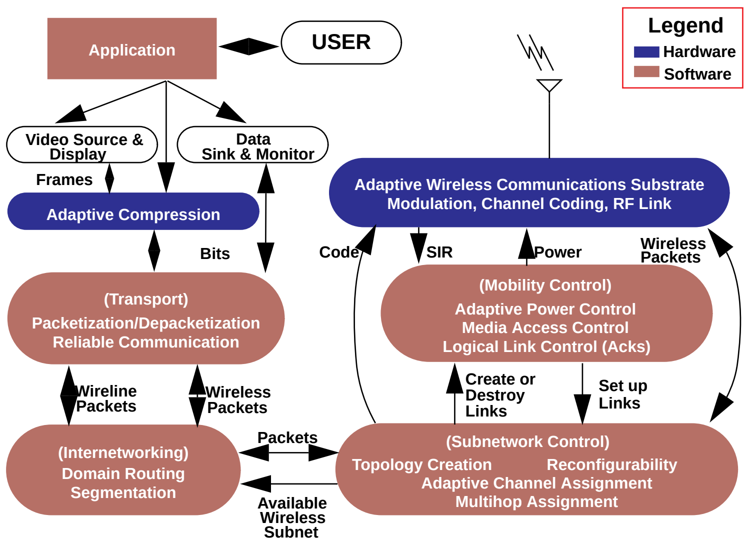

Figure 1: Wireless multimedia node functions

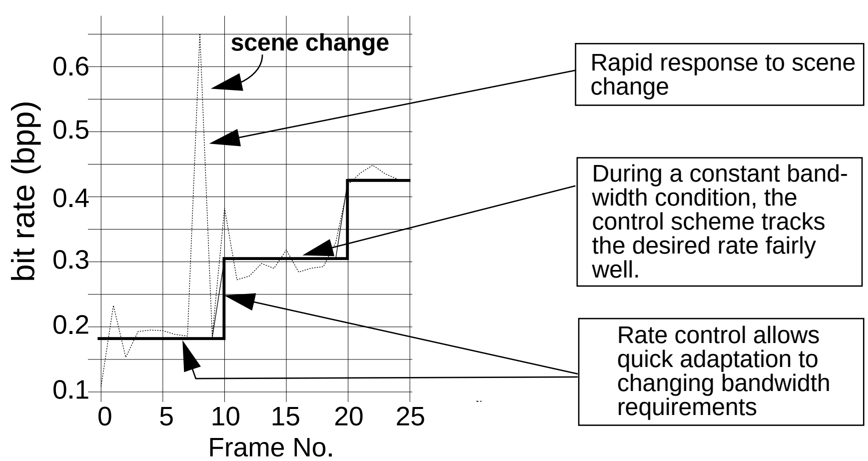

Figure 2: Video rate control performance. Solid horizontal lines are the desired output rate.

Figure 3: 450 kbps 256x256 video, 15 frames/s, intra-frame coding at 0.46 bpp, channel BER 0.003. Unequal error protection of image subbands with RCPC codes.

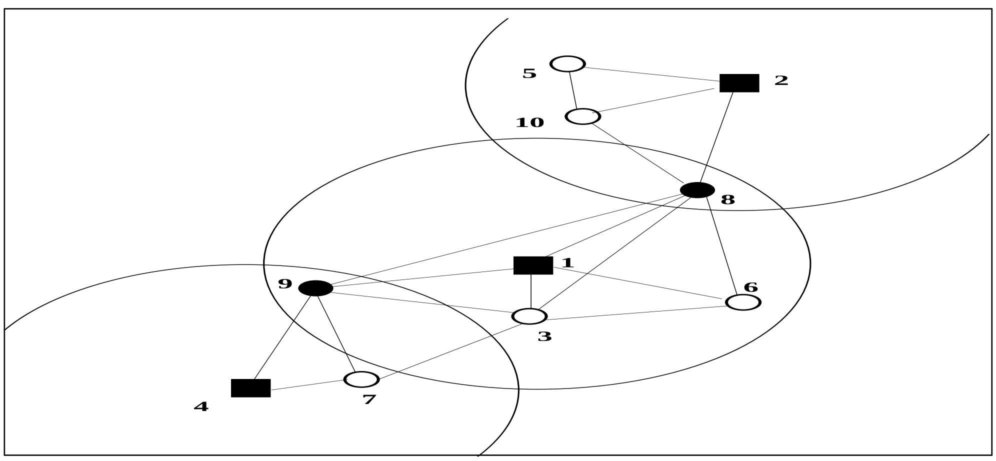

Figure 4: Example of cluster formation (lowest-ID)

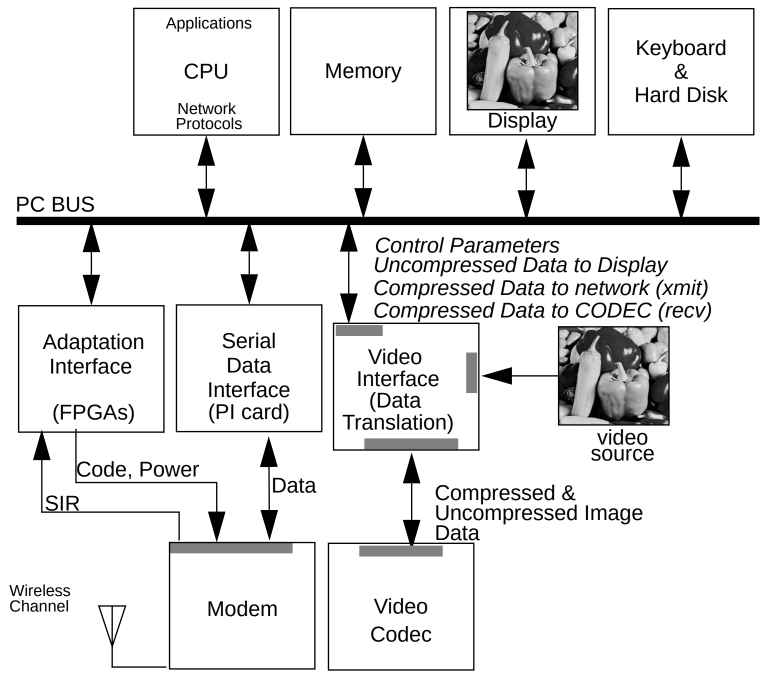

Figure 5: PC-notebook-based multimedia wireless node architecture

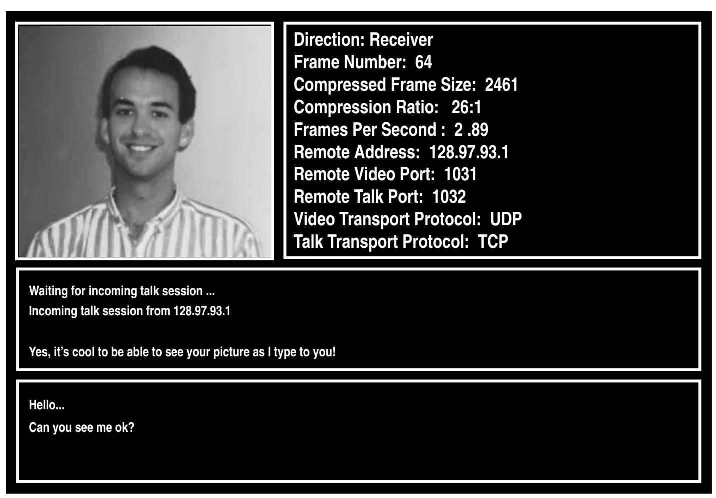

Figure 6: VTalk application