Report Number: WUCS-94-20

1994-01-01

Production Quality Video Over Broadband Networks: A Description of the System and Two Interactive Applications

William D. Richard, Jerome R. Cox Jr., A. Maynard Engebretson, and Jason Fritts and Brian L. Gottlieb and Craig Horn

The Washington University MultiMedia eXplorer (MMX) is a complete, host-independent multimedia system capable of transmitting and receiving JPEG-compressed video, CD-quality audio, and high-resolution radiographic images over the Washington University broadband ATM network. If the host is equipped with an ATM interface card, normal network traffic can be supported via an ATM extension port on the MMX. The major components of the MMX are an ATMizer and three multimedia channels. The ATMizer implements the host interface, the interface to the ATM network, and hte interface to the three multimdeia channels. This paper describes the architecture of the MMX, the software... **Read complete abstract on page 2.**

Production Quality Video Over Broadband Networks: A Description of the System and Two Interactive Applications

William D. Richard, Jerome R. Cox Jr., A. Maynard Engebretson, and Jason Fritts and Brian L. Gottlieb and Craig Horn

Complete Abstract:

The Washington University MultiMedia eXplorer (MMX) is a complete, host-independent multimedia system capable of transmitting and receiving JPEG-compressed video, CD-quality audio, and high-resolution radiographic images over the Washington University broadband ATM network. If the host is equipped with an ATM interface card, normal network traffic can be supported via an ATM extension port on the MMX. The major components of the MMX are an ATMizer and three multimedia channels. The ATMizer implements the host interface, the interface to the ATM network, and hte interface to the three multimdeia channels. This paper describes the architecture of the MMX, the software used with the system, and two applications which have been developed to demonstrate the capability of broadband ATM networks for multimedia applications.

Production Quality Video Over Broadband Networks: A Description of the System and Two Interactive Applications

William D. Richard, Jerome R. Cox, Jr., A. Maynard Engebretson, Jason Fritts, Brian L. Gottlieb and Craig Horn

WUCS-94-20

July 1994

Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
St. Louis MO 63130-4899

This work was supported in part by the National Science Foundation, Ascom Timeplex, Southwestern Bell, SynOptics and the Mallinckrodt Institute of Radiology

Production Quality Video Over Broadband Networks:

A Description of the System and Two Interactive Applications*

WILLIAM D. RICHARD[1], JEROME R. COX, JR.,[2,3], A. MAYNARD ENGEBRETSON[2],
JASON FRITTS[1], BRIAN L. GOTTLIEB[1,2], AND CRAIG HORN[1],

Washington University in St. Louis

[1]Department of Electrical Engineering

[2]Department of Computer Science

[3]Electronic Radiology Laboratory

One Brookings Drive

St. Louis, Missouri 63130

ABSTRACT

The Washington University MultiMedia eXplorer (MMX) is a complete, host-independent multimedia system capable of transmitting and receiving JPEG-compressed video, CD-quality audio, and high-resolution radiographic images over the Washington University broadband ATM network. If the host is equipped with an ATM interface card, normal network traffic can be supported via an ATM extension port on the MMX. The major components of the MMX are an ATMizer and three multimedia channels. The ATMizer implements the host interface, the interface to the ATM network, and the interface to the three multimedia channels. This paper describes the architecture of the MMX, the software used with the system, and two applications which have been developed to demonstrate the capability of broadband ATM networks for multimedia applications.

Key words: Broadband, ATM, JPEG, Compression, Multimedia.

* This work was supported in part by the National Science Foundation, ascom Timeplex, Southwestern Bell, SynOptics, and the Mallinckrodt Institute of Radiology

I. INTRODUCTION

A. The Washington University ATM Network

Over the past six years, the Applied Research Laboratory at Washington University has developed a prototype Asynchronous Transfer Mode (ATM) switch, a set of demonstration network peripherals, and a testbed broadband network [1]. The original network was based on four geographically dispersed, 16-port, 100 Mb/s, prototype ATM switches connected by more than 200 miles of single mode fiber.

Recently, the network has been expanded, and the prototype switches were replaced by 155 Mb/s commercial versions whose designs were derived from our prototype and licensed by Washington University. As part of this upgraded network, we have developed a Department of Computer Science network with more than 20 multimedia workstations. We now plan a broader expansion of the network throughout the campus, initially with several dozen additional multimedia workstations going to hundreds of workstations over the next few years. These workstations have been purchased by our colleagues and collaborators from several different vendors and did not ship with multimedia capability. We were anxious to demonstrate the full capability of our ATM network and have sought a solution that allows the use of existing computing platforms yet provides high-performance multimedia communication, i.e., a platform independent multimedia solution.

We call this network enhancement and expansion Project Zeus, the development, deployment and application of a broadband network on the University campus. Figure 1 shows the plan for Project Zeus including an operational network with OC3c (155 Mb/s) and OC12c (622 Mb/s) switches and a gigabit testbed. The gigabit testbed is to be linked to other gigabit switches in the metropolitan area by an experimental OC192 (9.95 Gb/s) add/drop SONET ring. Presently installed is the Department of Computer Science subnet, eight OC3c switches, and connections to Radiology and Barnes West. Completion dates planned for the operational network, the gigabit testbed, and the SONET ring are 1995, 1996, and 1997, respectively.

Collaborators in Project Zeus throughout the campus who have challenging applications requiring broadband technology will be linked by the network making possible distributed data acquisition, computation and presentation of results. In addition, two applications that have potential for wide use, medical imaging and communication for the deaf, are under development. These applications will be described below in detail.

B. The Washington University MultiMedia eXplorer

The Washington University MultiMedia eXplorer (MMX) is an important component of Project Zeus; it provides a complete multimedia system capable of transmitting and receiving video, audio, and radiographic images over the Washington University broadband ATM network. The MMX is a second-generation multimedia system based on the earlier, host-based Washington University Multimedia System (MMS) [2]. Campus locations where MMXs are to be installed are shown in Fig. 1. A block diagram of a typical MMX configuration is shown in Fig. 2.

An MMX is typically used with an NTSC [3] video camera, microphones, and amplified stereo speakers. The analog NTSC video signal generated by the video channel is usually fed into a "video-in-a-window" card installed in the host computer as shown in Fig. 2. In this configuration, the received video is displayed in a window on the normal workstation display eliminating the need for a separate video monitor. A high-resolution radiological image display is supported for medical applications. Other configurations are possible, including the use of a separate video monitor or the use of a laser disk player or other audio/video source

C. Applications

To reduce medical costs, hospitals are seeking cost saving opportunities through the sharing of information via networks. For example, using an MMX and medical video, a specialist can efficiently support ultrasound and fluoroscopic examinations at a number of satellite locations. The advent of broadband networks promises to remove the barrier to electronic communication for the deaf and hard of hearing. We have demonstrated the effectiveness of the MMX to provide the capability for the deaf to communicate over arbitrary distances as if they were face to face. These two applications foreshadow large-scale usage of broadband networks when metropolitan and wide-area tariffs become sufficiently low.

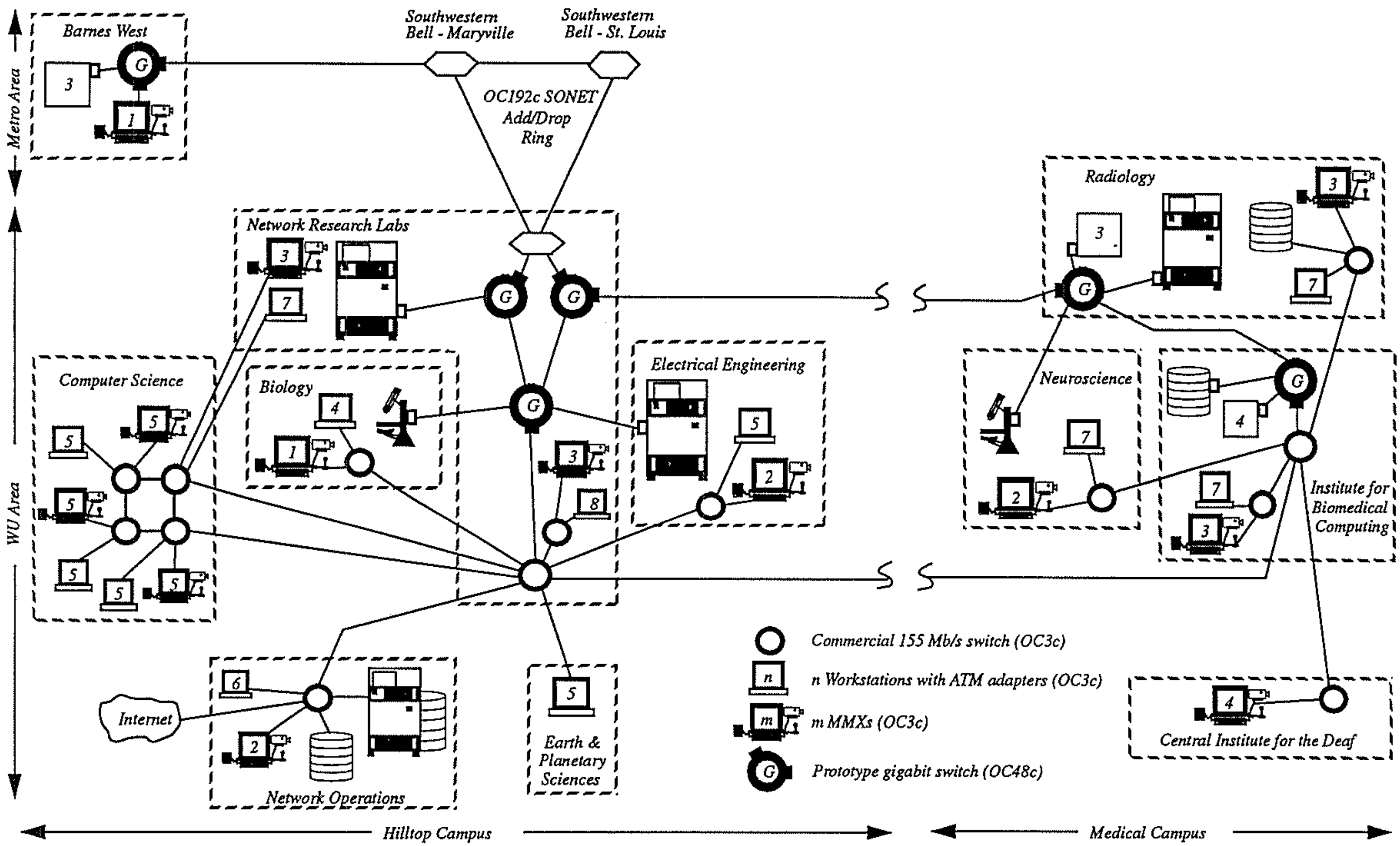


Figure 1: Planned deployment of the Washington University Zeus network from 1994 to 1997.

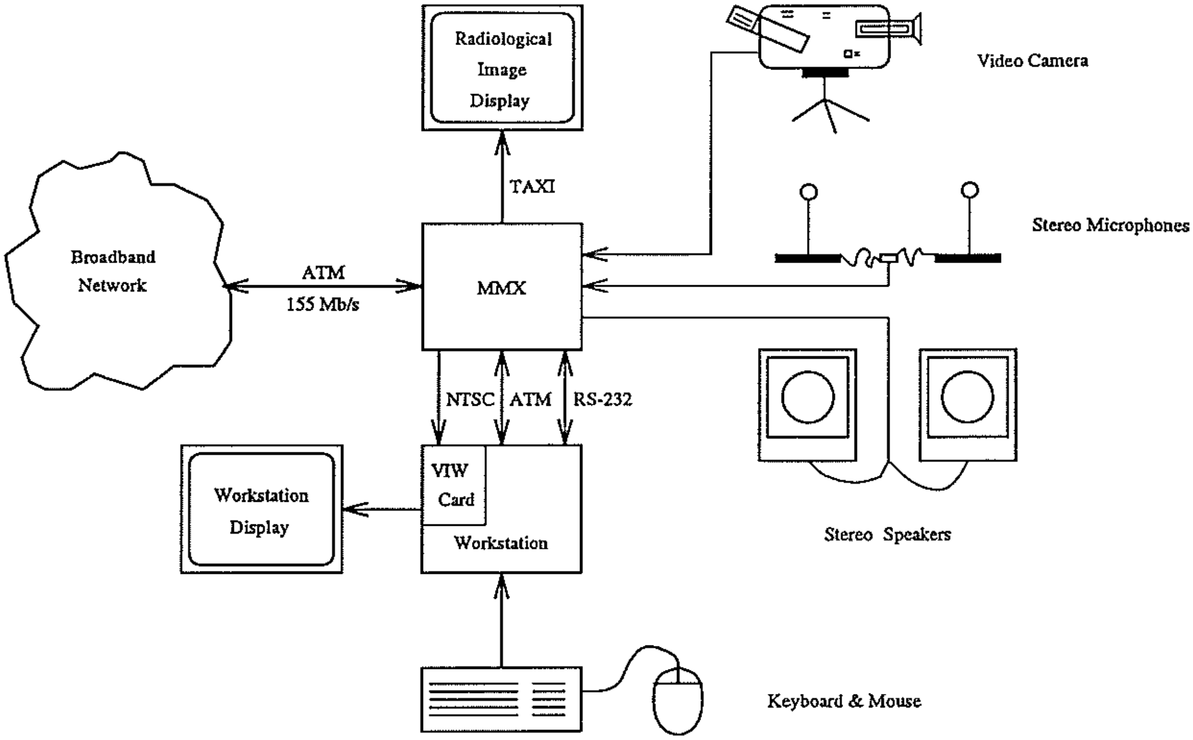


Figure 2: An MMX is typically configured with a video camera, stereo microphones, and amplified stereo speakers. The video output from the MMX is fed into a "video-in-a-window card in the workstation in the configuration shown.

II. THE WASHINGTON UNIVERSITY MULTIMEDIA EXPLORER

A. Overview

An overview block diagram of the MMX system is shown in Fig. 3. The MMX is designed to operate with any host. A standard RS-232C interface [4] is used to control the MMX from an application program running on the host. If the host is equipped with an ATM interface card, it can receive normal network traffic via an extension port on the MMX provided by internal "T" and "Y" connections. The "T" connection copies in-bound ATM cells and transmits them to the host ATM card. The "Y" connection multiplexes cells from the host ATM card with cells originating from the MMX to form a single cell stream. Audio, video, and radiological image delivery functions are performed by the multimedia subsystem.

B. The ATMizer and "T-Y" Connections

As shown in Fig. 4, the ATMizer subsystem consists of a 16 MHz Motorola MC68030 CPU [5] with 4 Megabytes of DRAM, 128 Kilobytes of EEPROM, a Multi-Function Peripheral Chip (MFP) [6], and interfaces to the ATM network and to the multimedia subsystem. The EEPROM holds the embedded control software for the MMX. The DRAM is used for serial I/O buffering and CPU channel ATM cell management. The MFP provides timing and interrupt services, as well as the serial port used to communicate with the host.

Embedded code for the local MC68030 CPU is written in C [7] and currently compiled using the native C compiler on a NeXT host. A library of I/O function calls, e.g., *printf* and *scanf*, was written to replace the stdio C library to facilitate code development and debugging. A low-level monitor program, which is part of the embedded code, performs memory dumps, allows program execution, etc. For debugging or diagnostic purposes, a terminal can be used with the on-board serial port to interact with this monitor.

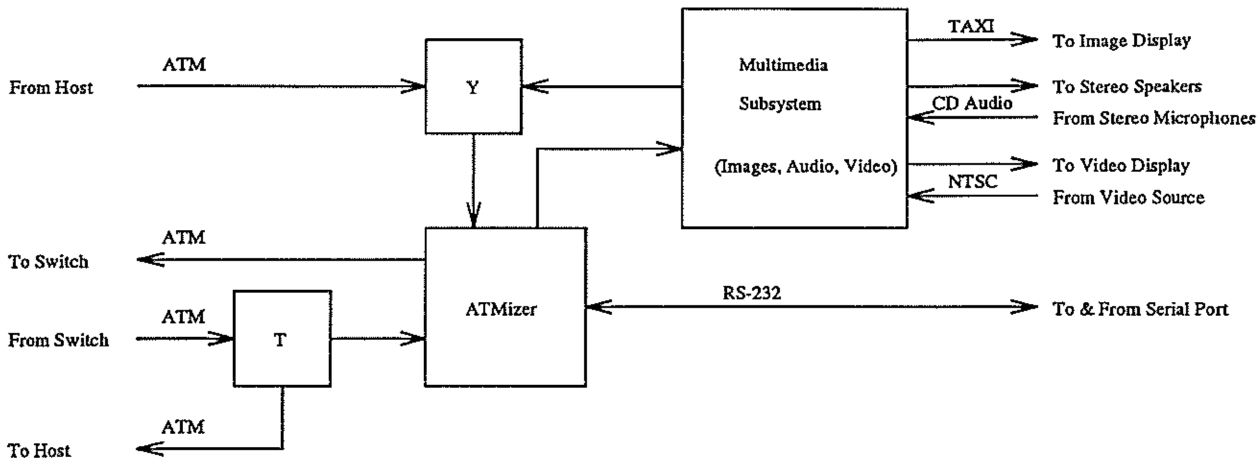


Figure 3: The MMX system block diagram.

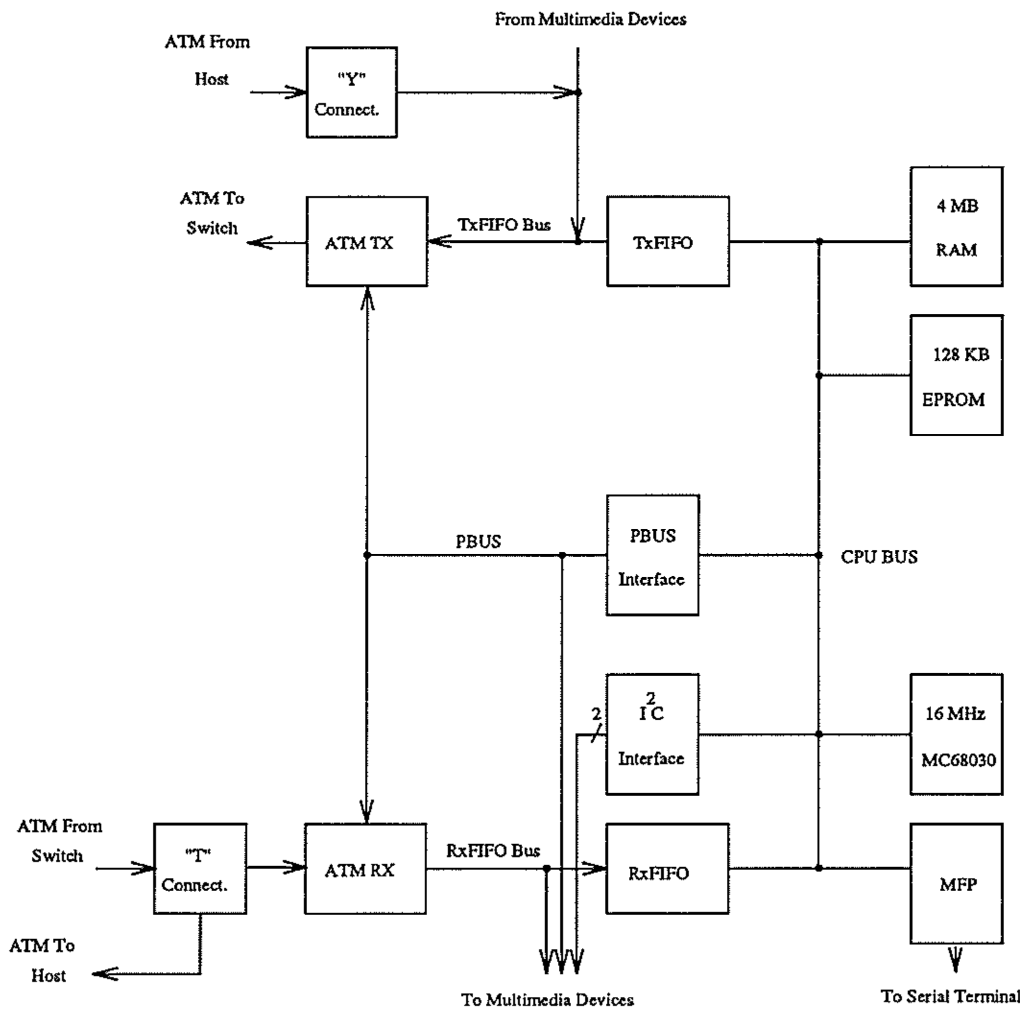


Figure 4: The ATMizer and "T-Y" connections.

Three dedicated buses, the PBUS, the TxFIFO Bus, and the RxFIFO Bus, are used to connect the multimedia subsystem to the ATMizer. The PBUS is a dedicated, 16-bit, multiplexed address/data peripheral bus derived from the CPU bus. The audio and image compression components in the multimedia subsystem are controlled via the PBUS. A fourth dedicated bus, the I^2C Bus [8], is used to control the video components in the multimedia subsystem. Multimedia data streams pass between the multimedia subsystem and the ATM network, via the ATMizer, over the TxFIFO and RxFIFO buses. Local CPU bus bandwidth is not used for multimedia data transfer.

The local CPU interfaces to the ATM subsystem via a pair of FIFOs. These FIFOs are mapped into the address space of the CPU. To transmit an ATM cell, the CPU writes the cell, including header, to the TxFIFO. The output of this FIFO is connected to the TxFIFO Bus. Incoming ATM cells are written into the RxFIFO from the RxFIFO Bus. The CPU can check for incoming cells by either polling the RxFIFO empty flag from this FIFO or by enabling an interrupt service routine. The RxFIFO empty flag is used to generate the required interrupt request.

The interface to the ATM network is implemented using an Advanced Micro Devices TAXIchip transmitter/ receiver pair [9]. The ATMizer is configured to transmit and receive data via either twisted pair or multi-mode fiber using the framing protocol specified by the ATM Forum for ATM systems operating at 155 Mb/s (8B/10B) [10]. The "T" connection, using analog circuits, copies incoming signals and transmits them to both the local TAXI receiver and to the ATM link to the host. The ATM interface on the MMX can be programmed to ignore cells not specifically directed to either the local CPU or one of the multimedia channels. Similarly, the host must ignore cells bound for the MMX.

The TxFIFO Bus is also used by the multimedia channels for data transmission to the ATM network. In addition, this bus is used to multiplex cells from the "Y" connection into the output ATM stream. Each multimedia channel, and the "Y" connection, use a FIFO to supply data to the TAXIchip transmitter. The programmable full flag from each FIFO connected to the TxFIFO Bus is sampled by the ATM Tx block at the beginning of each ATM cell transmission cycle. A priority scheduling scheme is implemented by the ATM Tx block that gives the audio channel the highest priority, the "Y" connection the second-highest priority, the video channel the third-highest priority, and the local CPU the fourth-highest priority.

In general, the multimedia subsystem provides a data stream to the TxFIFO Bus that does not include the appropriate ATM header for transmission. The ATM Tx block inserts the required headers. These headers are stored in tables in the ATM Tx block by the local CPU via the PBUS. The capability also exists to provide headers for a data stream from the TxFIFO and/or the "Y" connection, although this is typically not done. The local CPU and the host normally generate complete cells for transmission that include the appropriate header and header error control (HEC) byte.

The RxFIFO Bus is used to supply data to the RxFIFO and to the multimedia subsystem. As the bytes of a cell are received from the ATM network, they are clocked into a byte-wide shift register where the VPI (Virtual Path Identifier) and VCI (Virtual Channel Identifier) fields are examined to identify the target channel. This function is performed using look-up tables initialized by the local CPU via the PBUS. The ATM Rx block routes each incoming cell to the appropriate channel by asserting the appropriate FIFO select signal to one or more of the FIFOs attached to the RxFIFO Bus. The ATM Rx block also monitors the output of the HEC circuit and can, optionally, disable delivery of corrupted cells.

The tables used by the routing portion of the ATM Rx block to route cells also contain information as to what parts of the incoming cell are to be delivered. For cells received on a given virtual channel/path, any combination of the four-byte header, the header error check byte, or the 48-byte payload can be routed to the target channel. This allows the RxFIFO to receive a complete cell, including the header, while the multimedia subsystem receives only a payload data stream. This functionality also allows the RxFIFO to receive just the 4-byte header and the 48-byte payload or just the header error check byte. This latter option can be useful when measuring data throughput since only one byte need be processed per cell.

C. The Video Channel

There are currently three multimedia channels used with the ATMizer. The video channel, shown in Fig. 5, digitizes an input color NTSC video signal [3], compresses the digital video data stream using a hardware Joint Photographic Experts Group (JPEG) [11] compression engine, stuffs the appropriate framing tags into the data stream, and routes the resulting data stream to the ATMizer via the TxFIFO Bus for transmission. The video channel also accepts a compressed video data stream from the ATMizer via the RxFIFO Bus, strips the framing tags, decompresses the data stream, and produces an output NTSC video signal. A frame buffer is used in conjunction with a novel Phase-Locked-Loop (PLL) scheme to eliminate the synchronization problems caused by separate transmitter and receiver clocks [12]. Two hardware JPEG engines are used by the video channel so that a full 640-by-480 pixel video image can be compressed for transmission and decompressed for display simultaneously at 30 frames per second.

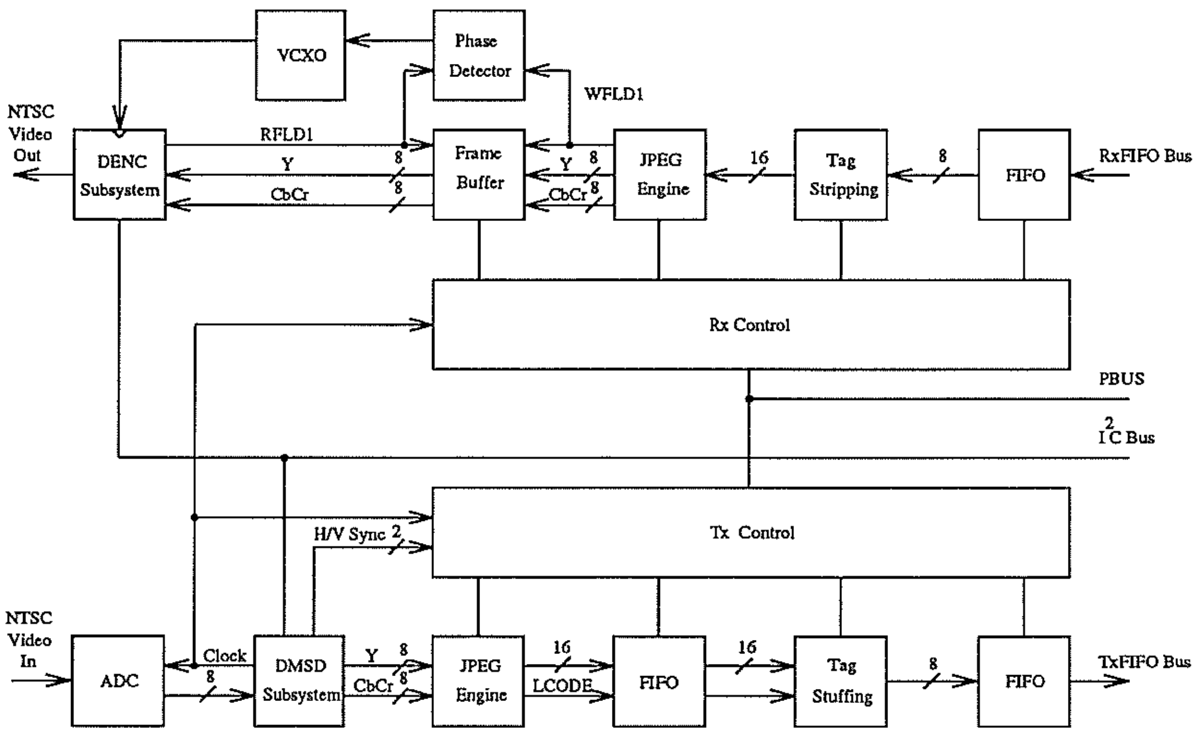


Figure 5: The video channel.

The video channel uses Phillips/Signetics video components [8] to digitize the incoming NTSC color video signal and to regenerate the output video signal. A Digital Multistandard Decoder (DMSD) is used in conjunction with an Analog-to-Digital (A/D) converter to produce a 12.2727 Megabyte per second Y:Cb:Cr (Luminance, Chroma Blue, Chroma Red) data stream during the active portion of each video line. The Y:Cb:Cr data is subsampled and multiplexed in 4:2:2 format.

The output of the DMSD is fed into a JPEG Engine implemented using an LSI Logic JPEG chip set [13]. This three-chip set consists of an L64765 Color and Raster-Block Converter, an L64735 Discrete Cosine Transform (DCT) Processor, and an L64745 JPEG Coder. The Color and Raster-Block Converter is used to convert the digital video data stream, which is produced by the DMSD in raster format, into block format. In the transmission portion of the video channel, the DCT Processor is used to compute the forward DCT over the 8 x 8 data blocks generated by the Color and Raster-Block Converter. The coefficient data stream produced by the DCT Processor is fed to the JPEG Coder. In the transmission portion of the video channel, the JPEG Coder encodes each video frame as specified in the JPEG standard [11]. The Coder performs quantization and has Differential Pulse-Code Modulation (DPCM) coding, run-length coding, and variable length (Huffman) coding capabilities.

The output of the JPEG Coder is buffered using a FIFO. Data from the FIFO are processed by the Tag Stuffing block. This block inserts into the data stream the appropriate JPEG codes for start of field one, start of field two, and end of field. The output of the Tag Stuffing block is again buffered in a FIFO. The output of this FIFO is connected to the TxFIFO Bus. This FIFO's programmable full flag, which is programmed via the PBUS to assert when the FIFO contains 48 bytes, is used to signal the ATMizer that a full cell payload is ready for transmission.

The operation of the transmission portion of the video channel is controlled by the Tx Control block. The operation of this channel, with the exception of the FIFO connected to the TxFIFO bus, is synchronous to the video sample clock created by the DMSD.

In the receive portion of the video channel, a FIFO attached to the RxFIFO Bus is used to buffer the incoming data stream from the ATMizer. Data passes from this FIFO through a Tag Stripping block to the JPEG Engine. The Tag Stripping block strips the start of field one, start of field two, and end of field tags from the data stream. This information is used by the Rx Control block to control the data flow through the JPEG Engine. The JPEG Engine consists of the same three components used in the transmission portion of the video channel. Here, however, the components are used in the decompression mode, rather than the compression mode.

The video data stream produced by the Raster-Block Converter in the JPEG Engine is written into a Frame Buffer constructed using Field Memories (FMEMs) [14]. FMEMs are similar to normal FIFOs in that read and write access my occur asynchronously. Unlike a conventional FIFO, however, read and write address pointers increment independently, possibly passing one another multiple times. Both address pointers are reset to zero on initialization.

The operation of the receive portion of the video channel is controlled by the Rx Control block. The operation of this channel is data-driven and is asynchronous to the video pixel clock generated by the Digital Encoder (DENC) subsystem. Since the data stream received by this channel was created by a remote transmission subsystem with an independent clock, synchronous operation, i.e., operation synchronous with the video sample clock generated in the local transmission subsystem by the DMSD, would be very difficult to implement [12]. Over a period of time, the receive FIFO buffer would either overflow (remote transmitter clock too fast) or underflow (remote transmitter clock too slow). This problem would be compounded by cell-to-cell jitter introduced by the ATM network. The Rx Control block uses a gated clock scheme to operate the JPEG Engine only when data is available. The JPEG Engine is not synchronized with the DENC, and it writes decompressed image data into the Frame Buffer at its peak rate when data are available. The overhead associated with the horizontal and vertical blanking intervals at the transmitter ensures that the JPEG Engine has sufficient bandwidth to eliminate data overflow problems. The frame buffer, which can be read and written independently, is used with the novel PLL scheme described below to minimize the problems caused by asynchronous image decoding/generation.

The DENC produces the output NTSC video signal from the Y:Cb:Cr data in the Frame Buffer. The DENC has all the circuitry necessary to perform the modulation required to convert the digital Y:Cb:Cr data into standard NTSC composite video [3]. The DENC clock input is driven by the output of a Voltage-Controlled Crystal Oscillator (VCXO) driven by a phase detector that compares the odd/even field signals from the JPEG engine and from the DENC. In its current mode of operation, this PLL scheme controls the DENC clock frequency so that the read and write pointers to the Frame Buffer are separated by approximately one-fourth the length of the buffer.

When the remote video source is switched, it is possible for the write and read pointers to be briefly coincident before the PLL adjusts the phase of the read pointer relative to the write pointer. If this happens, the DENC may display portions of two transmitted video frames during one display frame time. Since two video frames typically differ by only a small amount, and since the PLL quickly moves the pointers away from each other, artifacts due to the coincidence of the write and read pointers are generally not perceptible.

If the PLL circuitry is modified so that the write and read pointers are always coincident (within the limits set by network jitter, etc.), a phenomenon we term "waterlines" will be visible. This horizontal "tear" in the displayed image is due to the display of portions of two frames and image object motion between frames. If no PLL circuitry is used, and the DENC clock input is driven by a free-running oscillator, this mode of operation, i.e., where the write and read pointers are coincident, can occur if the frequency and relative phases of the transmitter and receiver subsystems match. In this case, the "waterlines" artifact is visible until the pointers drift far enough apart.

Experiments indicate that the maximum, worst-case frequency deviation of the VCXO from its center frequency of 12.2727 MHz allowable, before artifacts are visible on the display monitor, is approximately 1 kHz. This maximum deviation has been found to be a function of the monitor, with some monitors allowing larger deviations. With a maximum deviation of 1 kHz, it can take a minute or more, worst case, for the PLL to pull the read pointer one-fourth the length of the Frame Buffer from the write pointer. Since a much smaller separation is required to eliminate interaction, which requires only a few seconds, this solution has been found to be very acceptable.

Currently under development is an additional video subsystem that will provide Picture-In-a-Picture (PIP) capability for the MMX, allowing the concurrent display of local and network video. This PIP system will allow a local video signal to be displayed in a small window inset into the network video display. This capability will be particularly useful in allowing a user to view the display from their local camera while in video conference with another network user. The PIP system will utilize Siemens' Picture Insertion Processor chip [15]. Used in conjunction with NTSC decoders and encoders, the PIP will be able to insert a local video window of 1/9 or 1/16 full size anywhere within the video display. Control of the window size and location (and other options) will be performed through an I^2C Bus Interface, i.e., the PIP will be controlled through the same mechanism currently used on the MMX.

D. The Audio Channel

The second multimedia channel, shown in Fig. 6, digitizes an input analog stereo audio signal, i.e., both left and right channels, using a stereo audio codec [16] and produces a CD-quality digital audio data stream for transmission by the ATMizer (via the TxFIFO Bus). The codec oversamples at 64 times the output word rate of 44.1 kHz and generates 16-bit values for both the left and the right inputs. The 48-byte payload of each transmitted cell contains twelve digital audio data words, i.e., twelve 16-bit left and 16-bit right channel sample pairs. With this sampling rate and payload format, one audio cell is transmitted every 272 microseconds. This channel also accepts a digital audio stream from the ATMizer (via the Rx FIFO Bus) and generates an output analog stereo audio signal. Monaural input and monaural output are also supported by the system.

A TMS320C50 Digital Signal Processor (DSP) [17] provides volume control, loopback, left/right mixing, and amplification functions, as well as single-source payloads and multiple-source, monaural payloads. Multiple-source payloads are accepted by the receiving channel on a priority basis and mixed together as a blend of the two audio streams. Source priorities are established when a receiving MMX system is initialized and are stored in global memory. The DSP implements a voice activation switch so that audio sources generate cells only when sounds above a threshold are present. Feedback cancellation filter, also implemented by the DSP, are used to eliminate the direct signal path from the loudspeaker to the microphone. The echo cancellation feature of the audio board can be enabled to eliminate feedback due to having a microphone in close proximity to the speakers. This feature is most useful in a conferencing situation, where multiple conference participants are using microphones and speakers. If not properly controlled, such a configuration can have a severe feedback loop. With the echo cancellation enabled, this loop is broken without the need for precise arrangement of the speakers and microphones.

The functions of the audio channel are controlled via a 1 kilobyte dual-port memory. One port of this memory is connected to the local DSP bus. The other port is connected to the local CPU via the PBUS. The audio volume and sampling rate settings are adjusted via memory locations in the dual-port memory. Eight sampling rates from 7.35 kHz to 44.1 kHz are supported.

Since the audio sampling frequency at the transmitter can be slightly different from the playback frequency at the receiver, a mechanism is needed to compensate for data overflow or underflow. The problem is similar to the synchronization problem described above for the video channel. The solution for single source payloads implemented by the DSP involves the controlled duplication or deletion of audio samples. The required duplication or deletion is based on the amount of audio data contained in the RxFIFO buffer. If this FIFO is between one quarter and three quarters full, as indicated by the FIFO's flags, data is neither duplicated nor deleted as it passes through the DSP. If

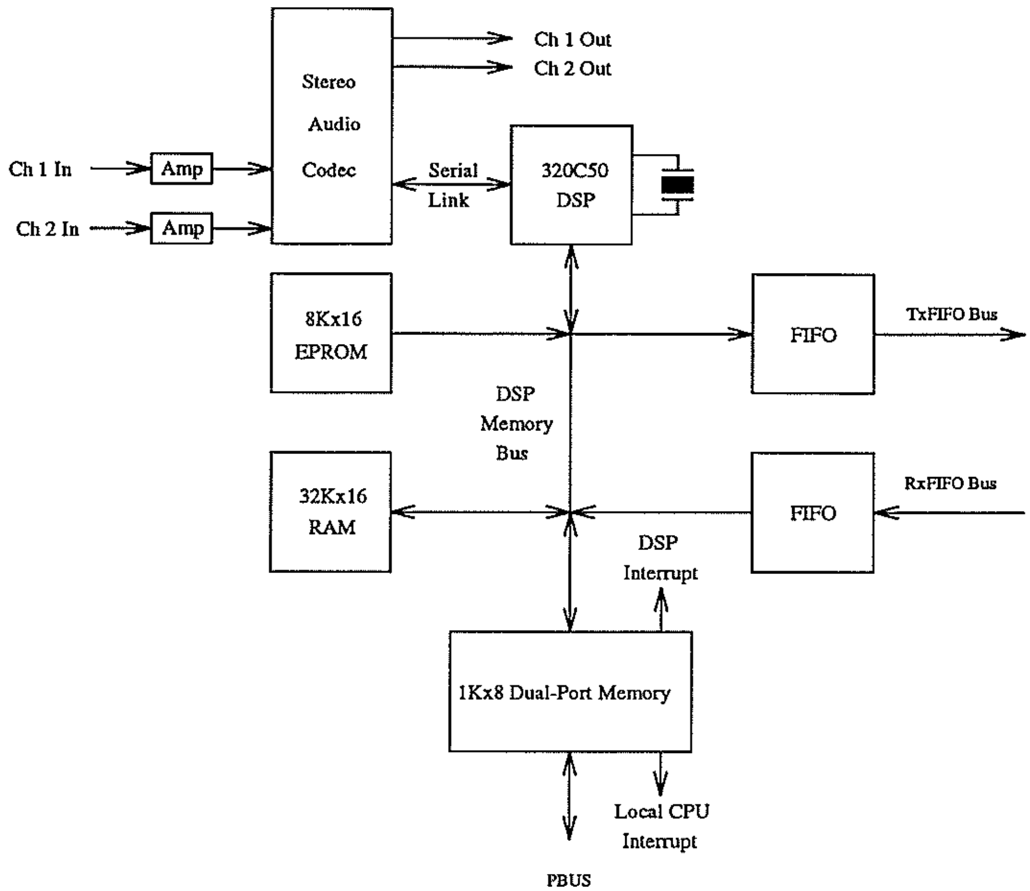


Figure 6: The audio channel.

the FIFO is less than one quarter full, one left-right sample pair of each cell is duplicated as data are passed from the FIFO to the D/A converter. This action results in a build-up of data in the FIFO over time, and, eventually, the FIFO will become one quarter full. If the FIFO is empty, zero sample values are transmitted to the D/A converter until new data arrives.

If the FIFO is more than three quarters full, the DSP deletes the first left-right sample pair of each cell. This action results in a reduction of data in the FIFO over time, and, eventually, the FIFO will become three-quarters full. If the FIFO is full, an error message is sent to the local CPU, and the FIFO is reset. In routine operation, this situation should not occur.

For multiple source payloads, a generalization of the duplication or deletion algorithm described above is used. The DSP must manage multiple pointers, but the scheme for synchronizing the audio from each of several sources is similar to the one described above.

E. The High-Speed Radiological Image Channel

The third multimedia channel accepts high-resolution radiographic images from the ATMizer via the RxFIFO Bus and reformats them for transmission to a high-resolution monochrome display [18]. The radiographic images used with this channel are served to the ATM network via a dedicated "mini-ATMizer" transmitter [19]. This transmitter accepts a data stream in the format required by the display and reformats it for transmission over the ATM network.

III. SOFTWARE

A. Overview

The MMX software suite consists primarily of a low level monitor running on the embedded CPU and a library of C routines for accessing the monitor functions via the serial port. This library has been used to develop several application programs for the Project Zeus testbed on multiple workstation platforms.

B. Embedded Software

The embedded software on the MMX CPU provides control of all devices in the MMX. A summary of a subset of the MMX internal commands is shown in Table 1 below. The functions of these commands are described briefly in the following paragraphs.

Command	Action	Syntax
cl, close	Close a Route	cl *device [rx|tx] header*
dis host	Disable Host Port	dis host
en host	Enable Host Port	en host
i, initialize	Initialize	i *[rx|tx]*
o, open	Open a Route	o *device [rx|tx] parts header*
q	Read Q Factor	q
set q	Set Q Factor	set q *enc dec*
set vs	Set Video Source	set vs *source*
set volume	Set Audio Volume	set vol *aa ab ba bb alb blb*
iic	Send an I^2C Command	iic *addr msg1 msg2 ...*

Table 1: A subset of MMX monitor commands.

For ATM communications, the embedded software provides the ability to filter and route incoming cells to any of the devices in the MMX. The routing is based on the VPI and VCI of the cell. For each VPI/VCI pair, the MMX may be programmed to pass to the destination device any combination of the cell header, HEC byte, and payload. On the transmit side, the outgoing headers for the audio and video channels may be programmed to send the audio and video streams using any VPI/VCI pair. Transmission of all streams is handled automatically whenever a full cell of data is available in the respective transmit FIFO. Programmable FIFO flags, which are programmed by the MMX initialization command, are used to initiate cell transmission.

For the video channel, the library provides routines for adjusting the JPEG Quantization (Q) Factor and for selecting which of three video inputs should be encoded and transmitted. There is also a general I^2C command for manipulating the various video parameters in the DMSD and DENC [8]. The I^2C command is used, for example, to put the DMSD into "VCR mode." This allows the DMSD to better follow the jitter in an analog NTSC signal originating in the VCR transport mechanism.

For the audio channel, there are functions for controlling the volume of the audio output, the mixing of local audio with the incoming ATM audio, the sampling rate, echo cancellation, and the number of incoming ATM audio streams. There are six controls for audio volume. Two are for listening to the local audio. These send the audio from the stereo inputs A and B to their respective outputs, bypassing the ATM link. The remaining four volume controls are for mixing the ATM A and B audio channels at the outputs. One control is for channel A to output A, one is for channel B to output B, one is for channel A to output B, and the last is for channel B to output A. The latter two controls are useful when a system has only one speaker. By using the appropriate controls, the MMX can send both audio channels to the same speaker, converting the stereo signal into a monaural signal.

The audio channel sampling rate command can set the sampling rate to one of eight different settings ranging from 7.35 kHz (voice quality) up to 44.1 kHz (CD quality). In all cases, the samples are 16 bits for each of the two channels.

The commands for manipulating multiple incoming audio streams include a mechanism for prioritizing sources and for source selection. Each incoming stream has an ID associated with it that allows the audio board to separate the incoming streams. Based on the priority given to each stream, the streams are mixed by the DSP and sent to the outputs.

The memory map of the MMX is compatible with that of the Washington University Multimedia System (MMS) [2,21]. Thus, an MMS provides a convenient development environment. Code can be written in C and compiled on a NeXT using a standard UNIX C compiler. The code can be tested by downloading it into the DRAM of an MMS through the NeXTBus and executed there. This allows rapid prototyping without the need to program an EEPROM every time a new feature is added. Once a new version of the code is tested and working, an EEPROM can be programmed and placed in the MMX.

Future additions to the embedded software will include the ability to download and execute code on the embedded CPU and on the DSP on the audio board. These additions will allow the MMX to be easily customized for different applications.

C. MMX Control Library

A library of C routines was written to provide a simple programming interface to the MMX. The library effectively places a wrapper around all of the functions described above. This wrapper allows a programmer to make a simple function call without having to worry about the specifics of the serial control protocol associated with the MMX. This abstraction allows for changes in future versions of the embedded software without the need to rewrite the applications running on the workstations. This library was used to develop the applications described below.

D. MMX Host Software

The software for the host computer demonstrates the capability of the MMX hardware to support a multimedia workstation. One of the central programs in this suite is VideoExchange [22], a video conferencing tool that supports, for example, a multiple party video conference or electronic communication among the deaf and hard of hearing. VideoExchange provides a simple graphical user interface displaying the current state of the conference, including active participants as well as the source viewed by each participant. Users may easily switch their display to view and hear any of the participants. The program VideoExchange can also be used with the Washington University Multimedia System, allowing conferences involving both MMSs and MMXs.

Another application for the MMX is RIVA, the Radiology Image Viewing Application [23,24]. RIVA allows physicians to retrieve medical images, via the high-speed radiographic image channel, along with the appropriate written reports. Medical video used for ultrasound and fluoroscopic examinations and video reports stored on laser disk can also be retrieved via the video channel. As discussed below, this allows physicians at remote locations to participate in a video conference to discuss patients, and, in particular, examine their radiographic images and medical video clips.

IV. APPLICATIONS

The purpose of Project Zeus is to develop networking components, to install a networking testbed, and to develop challenging applications that stress the testbed. Two such applications will be described along with some speculations about the future use of network bandwidth by these applications.

Increasingly, to reduce medical costs, hospitals are merging into systems containing dozens of hospitals and healthcare facilities. These systems provide cost-saving opportunities through the centralization of expertise, the use of automation, and the sharing of information via networks. Such an opportunity exists in connection with fluoroscopic and ultrasonic examinations carried out at satellite clinics and hospitals.

For example, a technologist conducts the examination at the remote site (see Fig. 7). A specialist (gastrointestinal radiologist or obstetrician) located at the medical center provides guidance to the technologist via video conferencing and by monitoring medical video captured from the fluoroscopic or ultrasound equipment. Both the patient and the technologist can interact with the physician despite the distance. It is important to provide a good audio channel for interaction between the Center-of-Expertise and the Remote Site both during video conferencing and during audio conferencing while the technologist and physician are examining medical video sequences.

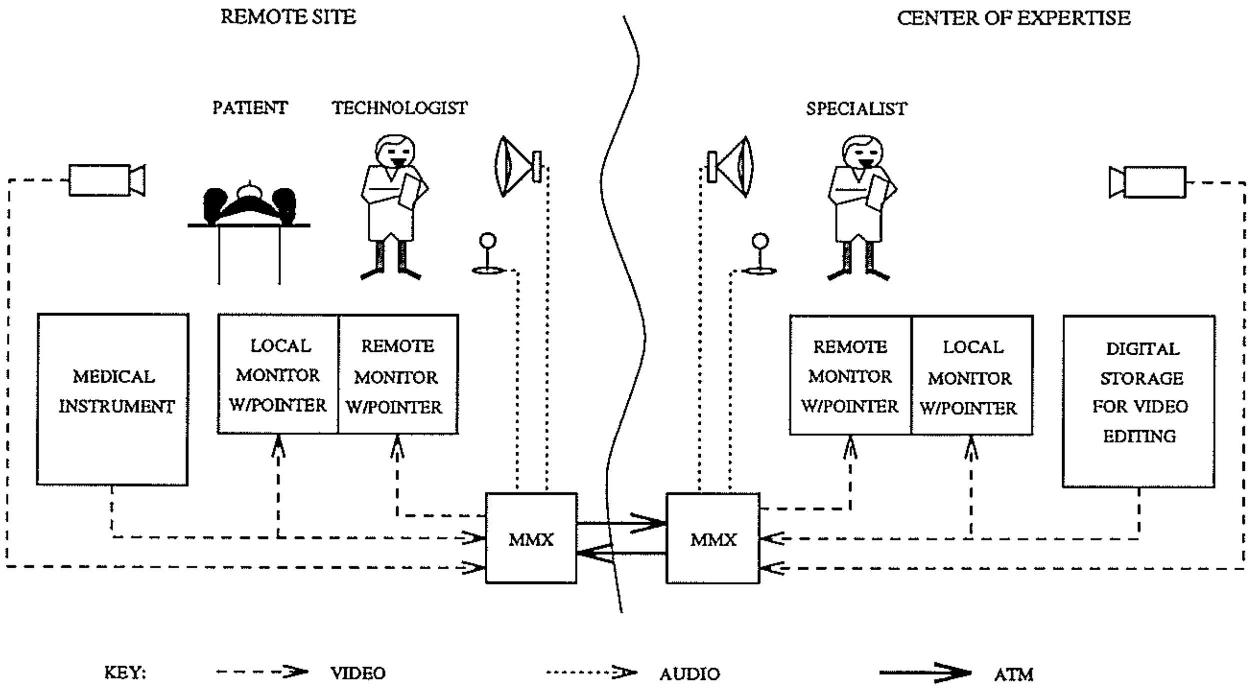


Figure 7: Use of MMX to support remote examinations requiring medical video.

Thus, with this arrangement, the specialist can efficiently support examinations at a number of satellite locations. There is no time lost for travel by the specialist, and a number of satellite locations can be placed throughout the community and in rural areas. For those cases where the equipment is expensive or a technologist cannot be occupied fully at the remote location, a mobile examining facility can be used to increase utilization. Of course, it will be necessary to have broadband access at each of the mobile exam locations, but that access should be readily available within a few years.

Early psychophysical results indicate that compression of the medical video for these specialties can be accomplished using motion JPEG at average data rates of 11 Mb/s or less without producing visually detectable artifacts as a result of image reconstruction. Using a staircase psychovisual protocol, Gohel, et al. [25] used recordings of video examinations both with a variable amount of compression and without compression. These recordings were taken from examinations in the three fields of OB/GYN ultrasound, abdominal ultrasound, and GI/GU fluoroscopy. The threshold of artifact detection for specialists in these three fields of medicine yielded average data rates that ranged from 2.5 to 11 Mb/s.

Several demonstrations have been conducted with MMXs set up as shown in Fig. 7. The physicians participating in these trials have agreed that the quality of the video sequences provides the support needed for supervision of examinations at remote locations.

The deaf and hard of hearing population has been disenfranchised with respect to electronic communication. The hard of hearing have had some ability to communicate with each other and the hearing world by amplified sound of telephone bandwidth, while the deaf have been limited to communications through teletypewriter networks. The advent of broadband networks promises to remove these communication barriers.

The deaf and their hearing interpreters can communicate using video via lip reading, finger spelling, signing, or some combination thereof. For example, two deaf individuals can communicate directly, or a deaf person can communicate with a hearing person indirectly through a hearing interpreter acting as and intermediary (see Fig. 8). The three individuals can be in three different locations: the deaf person and the hearing interpreter at locations with MMXs and the hearing party communicating with the interpreter over a conventional telephone connection.

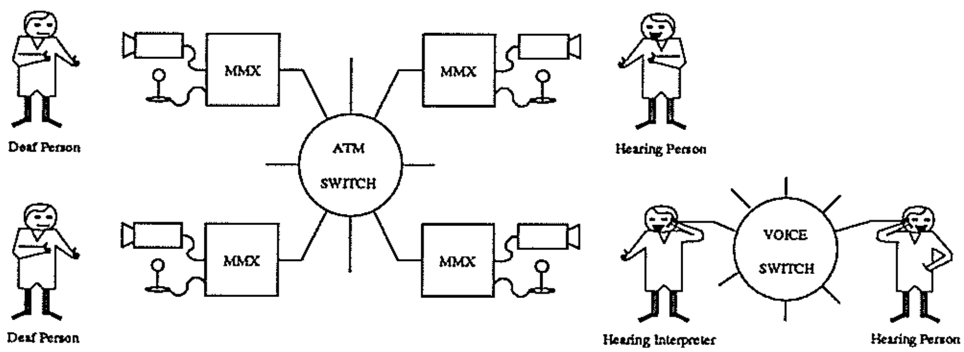


Figure 8: The MMX can be used effectively by the deaf and hard of hearing to communicate between themselves or with the hearing world directly, or indirectly through a hearing interpreter.

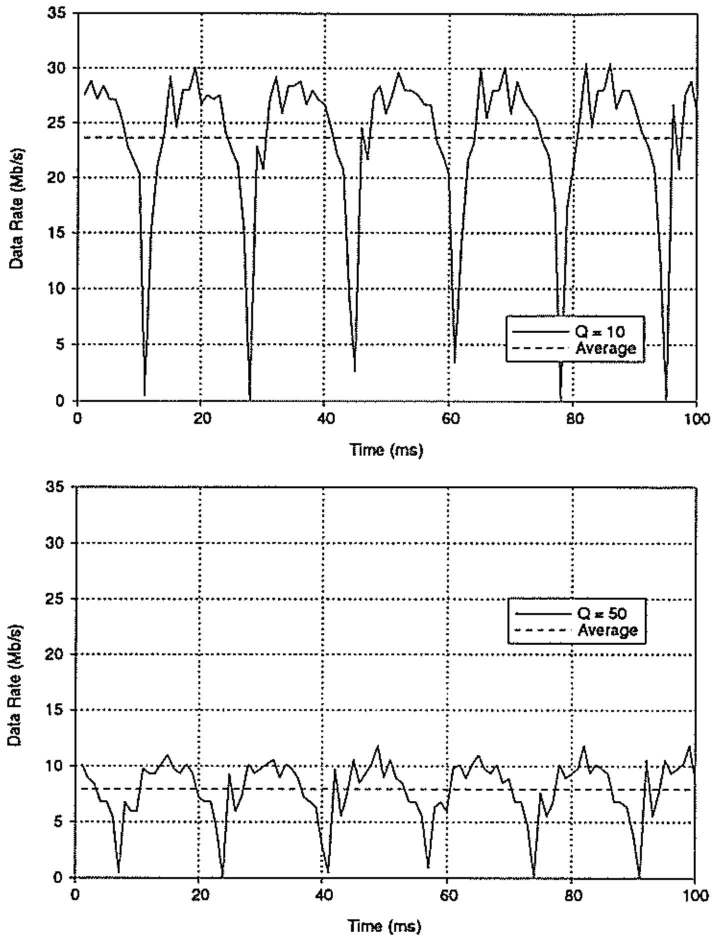
We have demonstrated on several occasions the effectiveness of the MMX to provide communication for the deaf over arbitrary distances as if the parties to the conversation were face to face. Initial studies of the amount of compression tolerable suggests that JPEG compression at 8 Mb/s with modest reconstruction artifacts do not harm the viewer's ability to communicate. However, lip reading was substantially more sensitive to compression artifacts than were finger spelling or sign language [26]

V. RESULTS AND DISCUSSION

The MMX has been operational since mid-1993. Twenty of the systems are currently deployed in the Department of Computer Science ATM network and are being used for demonstration and research purposes. Approximately twenty more will be installed during 1994. The issues associated with both video conferencing and teleradiology in a broadband ATM environment are being studied using the installed base of systems.

Figure 9 illustrates the JPEG data rate versus time for a typical image using Q factors of 10 and 50. This figure shows the data rate averaged over 1 ms intervals. The peak rate shown for Q = 10 is approximately 30 Mb/s, while the average rate is approximately 24 Mb/s. This average data rate corresponds to a coding rate of approximately 2.6 bits per pixel. The data rate plot is quasi-periodic, i.e., the data rate drops to near zero during vertical retrace (it would drop to zero if the effects of the 1 ms averaging and system buffering were removed). The fact that even and odd video fields are different, and therefore generate different data rates, is also evident from Fig. 9.

Figure 9 also illustrates the JPEG data rate versus time for the same image using a Q factor of 50. The peak rate shown is approximately 12 Mb/s, while the average rate is approximately 8 Mb/s. This average data rate corresponds to a coding rate of approximately 0.87 bits per pixel for the particular image sequence from which these data were obtained. Other sequences can produce different data rates.



Overall, the MMX produces production quality video using Q factors between 15 and 30. The resulting data rate is typically between 5 and 15 Mb/s. This corresponds to a coding rate of approximately 0.5 to 1.5 bits per pixel.

Figure 9: ATM data rate of a JPEG stream with Q factors of 10 and 50.

Since the MMX uses motion JPEG, a plot of the data rate is periodic with the period being equal to one field, i.e., 16.7 ms. Since motion JPEG does not use interframe coding, there are no motion artifacts in the transmitted images. Image corruption, due to cell loss or a switch of the source of a video stream, is also limited to a single field. This would not be the case with an MPEG [27] system where the loss of a single cell would corrupt several frames on the average.

The novel PLL scheme used with the MMX effectively eliminates the synchronization problems caused by the independence of video transmitter and receiver clocks. In the audio subsystem, the clock synchronization problem is addressed by duplicating or deleting a single stereo sample in an ATM cell (one out of 12 samples) whenever the audio FIFO is below 1/4 full or above 3/4 full. In the MMX, no audible effect is produced by this procedure. In fact, a simulation of the procedure has demonstrated the continuous duplication or deletion of one out of 12 samples at CD sampling rates (44.1 kHz) is not perceptible.

Bus bandwidth limitations on current workstation-class machines make real-time, bus-based, full-rate NTSC video applications difficult, if not impossible, to implement. In addition, bus-based applications consume CPU cycles that could be used for other purposes. One typical method used to address this limitation is to reduce the spatial and temporal resolution of the transmitted video signal. The Washington University MultiMedia eXplorer achieves real-time, full-rate video, audio, and radiographic image service by bypassing the host bus. This frees the host bus for normal network traffic. Until host bus and CPU speeds increase significantly, the approach used by the MMX can be used to achieve a high-quality multimedia solution.

The two applications described above permit a forecast of possible large-scale usage of broadband networks providing that metropolitan and wide area tariffs become sufficiently low. The MMX and our network of commercial ATM switches demonstrates that today's technology is functionally adequate. The business case for these applications is poor today, but may be stronger as demand grows and tariffs decrease.

An estimate of the ultimate demand for these two applications can be made. Assuming there are several hundred hospital systems in the nation and several dozen satellite clinics in each system, we can expect tens of thousands of medical video sessions at a full-duplex rate of 30Mb/s during a busy part of the day. The aggregate bandwidth required might easily approach 1 Tb/s. The hard of hearing includes up to 10% of the population. Assuming high-quality video conferencing would allow conversations among the deaf and hard of hearing with the same frequency as among the hearing population, one can estimate over 500,000 simultaneous calls at 10 Mb/s corresponding to an instantaneous load of over 5 Tb/s.

Both of the applications discussed require high quality video with little or no motion artifact. In the case of medical video, there is likely to be a future demand for higher resolution, particularly in fluoroscopy examinations. Even today there are 1024-by-1024-pixel-resolution fluoroscopy systems at 30 Hz frame rates that would produce about 40 Mb/s compressed video streams. If the streams from these higher resolution systems were not compressed, the data rate for monochrome video would be about 320 Mb/s. Thus, higher aggregate data rates are likely in the future. Both the additional equipment cost and the network tariffs will drop with time, and such decreases will stimulate additional demand which in turn will further reduce equipment cost and tariffs. However, it is uncertain when these regenerative effects will begin. Will it be in the next year or the next century?

VI. CONCLUSIONS

The Washington University MultiMedia eXplorer is a complete multimedia system developed for Project Zeus capable of transmitting and receiving full-rate, JPEG-compressed, NTSC video and CD-quality audio over the Washington University broadband ATM network. Real-time, full-rate operation is supported through the use of dedicated hardware that does not require the use of host-bus bandwidth for its operation. High-speed host communication is supported in conjunction with the MMX via an ATM extension port. Radiographic images can be received by the MMX and delivered to an auxiliary high-resolution monochrome display. The MultiMedia eXplorer demonstrates in a dramatic manner the capability and appropriateness of broadband ATM networks for multimedia applications. In particular, the medical video and deaf communication applications clearly demonstrate the advantages of such a system.

VII. ACKNOWLEDGEMENTS

The authors would like to thank, G. James Blaine, Tom Chaney, Pierre Costa, Pete Flugstad, Ken Krieger, Jim Miller, Steve Moore, Kurt Smith, and Bob Whitman for their contribution to the effort described here. The MMX development was supported by Project Zeus sponsors SynOptics, Ascom Timeplex, Southwestern Bell, Washington University and the Mallinckrodt Institute of Radiology. The MMX was based upon a design developed in the physician's workstation project with the sponsorship of Southwestern Bell, Washington University and the Mallinckrodt Institute of Radiology.

VIII. REFERENCES

[1] J.R. Cox, Jr., M.E. Gaddis, and J.S. Turner, "Project Zeus," *IEEE Network*, vol. 7, no. 2, pp. 20-30, March 1993.

[2] W.D. Richard, J.R. Cox, Jr., B. Gottlieb, and K. Krieger, "The Washington University Multimedia System," *ACM Multimedia Systems*, vol. 1, pp. 120-131, 1993.

[3] K.B. Benson, ed., *Television Engineering Handbook*. New York: McGraw Hill, 1986.

[4] *Interface Between Data Terminal Equipment and Data Communication Equipment Employing Serial Binary Data Interchange*. Washington, D.C.: Electronic Industries Association, 1969.

[5] *Motorola MC68030 Enhanced 32-Bit Microprocessor User's Manual.* Englewood Cliffs, NJ: Prentice-Hall, 1990.

[6] *MC68901 Multi-Function Peripheral Data Sheet,* Motorola Semiconductor Products, Inc., Jan., 1984.

[7] B.W. Kernigan and D.M. Ritchie, *The C Programming Language*, Englewood Cliffs, NJ: Prentice-Hall, 1988.

[8] *The Video Data Handbook,* Signetics Company, 1991.

[9] *TAXIchip Integrated Circuits Technical Manual,* Advanced Micro Devices, 1989.

[10] *ATM User-Network Specification Version 2.0,* The ATM Forum, June 1, 1992.

[11] G.K. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 31-44, Apr. 1991.

[12] W.D. Richard, "ATM video transmission: Considerations and an implementation," *Proc. 1st Int. Conf. Comp. Commun. and Networks*, San Diego, CA, June 1992, pp. 207-210,

[13] *JPEG Chipset Technical Manual,* LSI Logic Corporation, March, 1993.

[14] *MOS Memory Data Book,* Texas Instruments, Inc., 1991.

[15] *Siemens SDA9288X Picture Insertion Processor Data Sheet,* Siemens Corporation, January, 1993.

[16] *Crystal Semiconductor CS4216 Stereo Audio Codec Data Sheet,* Crystal Semiconductor Corporation, March, 1992.

[17] *TMS320C5X User's Guide*, Texas Instruments, Inc., 1991.

[18] J.R. Cox, Jr., R.G. Jost, T. Monsees, S. Ramamurthy, and M. Karlsson, "An inexpensive electronic viewbox," *SPIE*, vol. 914, pp. 1218-1224, 1988.

[19] D.M. Zar and W.D. Richard, "Mini-ATMizer User's Guide and Technical Manual," WUCS-92-50, Department of Computer Science, Washington University, St. Louis, MO, 1992.

[20] B. Gottlieb, "The Multimedia eXplorer RS-232 Instruction Set," ARL Working Note 94-05, Applied Research Laboratory, Department of Computer Science, Washington University, St. Louis, MO, 1994.

[21] K. Krieger, "An ATM Multimedia Interface for the MDWS," WorkNote_ERL_93_2, Electronic Radiology Laboratory, Washington University, St. Louis, MO, 1993.

[22] B. Gottlieb, "VideoExchange: A Multimedia Conferencing Tool," ARL Working Note 93-10, Applied Research Laboratory, Department of Computer Science, Washington University, St. Louis, MO, 1993.

[23] J.R. Cox, Jr., S.M. Moore, R.A. Whitman, G.J. Blaine, R.G. Jost, L.M. Karlsson, T.L. Monsees, G.L. Hassen, T.C. David, "Rapid Display of Radiographic Images," *SPIE Medical Imaging V: PACS Design and Evaluation,* vol. 1446, pp. 40-51, 1991.

[24] M.S. Wimmer, G.J. Blaine, R.G. Jost, S.M. Moore, and J.W. Studt, "A Distributed Approach to Integrated Inquiry and Display for Radiology," *Journal of Medical Systems*, vol. 15, no. 4, pp. 299-309, 1991.

[25] N.R. Gohel, G.J. Blaine, J.R. Cox, Jr., D.R. Fuhrmann, D.M. Balfe, D.L. Gray, and W.D. Middleton, "Considerations in the transmission and storage of JPEG-encoded medical video," *Proc. IS&T/SPIE Int. Symp. on Electronic Imaging: Science and Technology*, High-Speed Networking and Multimedia Computing Conference, San Jose, CA, February 6-10, 1994, vol. 2188, pp. 243-253.

[26] J.D. Miller, R.M. Uchanski, E.J. Frederick, A.F. Heidbreder, "Effects of video compression on visual language communication," *Proceedings of 1994 IEEE Workshop on Vidual Signal Processing and Communications* (Sept. 19-20, 1994).

[27] LeGall, Didier, "MPEG: A Video Compression Standard for Multimedia Applications," *Commun. ACM*, April 1991, vol. 34, no. 4, pp. 47-58.

BIOGRAPHIES

William D. Richard received the B.S., M.S., and Ph.D degrees in electrical engineering from the University of Missouri-Rolla in 1983, 1985, and 1988, respectively. He joined Department of Electrical Engineering at Washington University in St. Louis as an Assistant Professor in 1988. His research interests include ultrasonic imaging, medical instrumentation, and computer architecture.

Dr. Richard is a member of the Institute of Electrical and Electronics Engineers, the American Society for Engineering Education, and the Society of Automotive Engineers. He is also a member of the Sigma Xi, Phi Kappa Phi, Phi Eta Sigma, Eta Kappa Nu, and Tau Beta Pi honoraries.

Jerome R. Cox, Jr., received the S.B., S.M., and Sc.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1947, 1949, and 1954, respectively.

He is the Welge Professor of the Department of Computer Science at Washington University in St. Louis. Joining the faculty in 1955 as an Assistant Professor of Electrical Engineering, he was appointed Associate Professor in 1958 and Professor in 1961. From 1964 to 1975, he was Director of the Biomedical Computer Laboratory. He also served as Chairman of the Department of Computer Science from 1975 to 1991. He is currently Director of the Applied Research Laboratory and Research Director of the Electronic Radiology Laboratory. His research interests are digital communication, computer systems design, and electronic radiology.

Dr. Cox is a Senior Member of the Institute of Medicine of the National Academy of Sciences and is a Fellow of the Acoustical Society of America, the Institute of Electrical and Electronics Engineers, and the American Medical Informatics Association.

Maynard Engebretson received a BSEE degree from the University of Minnesota in 1958 and MSEE and D.Sc. degrees from Washington University in 1963 and 1970, respectively. His research interests have focussed on the application of computers and digital signal processing methods to problems of hearing and deafness at Central Institute for the Deaf where he was Associate Director of Research from 1987 to 1992. He is a senior research associate in the Department of Computer Science, and Associate Research Professor in the Department of Otolaryngology at Washington University in St. Louis, Missouri. His current interests include multimedia, speech processing, acoustic array processing, and active noise control. He is a member of the Institute of Electrical and Electronics Engineers, the Acoustical Society of America, Sigma Xi, and Eta Kappa Nu.

Jason Fritts received the B.S. degree in electrical engineering from Washington University in St. Louis in 1994. He will begin his graduate study in the fall of 1994 at Princeton University, working towards his doctoral degree in electrical engineering. His work involves hardware development for the MMX multimedia and ATMizer subsystems.

Brian L. Gottlieb received the B.S. and M.S. degrees in computer science from Washington University in St. Louis in 1991 and 1993, respectively. He also received the B.S. and M.S. degrees in electrical engineering from

Washington University in St. Louis in 1991 and 1993, respectively. His graduate work included software development for the MMX system.

Craig Horn received the B.S. degree in Computer Science from Washington University in 1992. He also received the B.S. and M.S. degrees in Electrical Engineering from Washington University in 1992 and 1994, respectively. His graduate work involved hardware development for the MMX multimedia and ATMizer subsystems.