An ATM–Based Local Communication System for Telesurgery

Elmar Holler
Kernforschungszentrum Karlsruhe GmbH, Institut für Angewandte Informatik,
P.O.B. 3640, D-76021 Karlsruhe · e-mail: holler@iai.kfk.de

Torsten Neck
Kernforschungszentrum Karlsruhe GmbH, Institut für Angewandte Informatik,
P.O.B. 3640, D-76021 Karlsruhe · e-mail: neck@iai.kfk.de

Abstract. The MONSUN (Manipulator Control System Utilizing Network Technology) concept allows to control telemanipulators via Local Area Networks (LAN). This concept has been implemented and successfully tested for industrial and medical applications and is currently used at KfK as a basis for the ARTEMIS project (Advanced Robotics and TElemanipulation System for MIS), a telepresence system for minimal invasive surgery being under development. The shortcomings of the LAN-based version of MONSUN — 10 Mbit/s Ethernet is currently being used — are due to the fact that the LAN can be used to close the control loop, but transmission of video and audio signals has to rely on additional transmission media. ATM (Asynchronous Transfer Mode) based high speed communication systems promise an "all-in-one" solution, at least for in-house applications. The system concept of an ATM-based local communication system for telesurgery is presented and the resulting performance characteristics are discussed.

1 Introduction

It was the idea to profit from the accumulated knowledge and experience gained during various research and development projects at the Nuclear Research Center at Karlsruhe (KfK) in telepresence, telerobotics and telemanipulation for technical applications such as plant inspection and maintenance, when the Advanced Robotics and Telemanipulation System for Minimal Invasive Surgery project ARTEMIS was started in 1992.

System and control concepts for MIS telemanipulators are not basically different from conventional, technical telemanipulation systems. In fact one of the main parameters changing in minimal invasive surgery is the decrease in scale with respect to conventional telemanipulation systems. Since KfK has a major research activity in the development of microsystems based on the LIGA technique, the development of MIS telemanipulation systems is expected to benefit from this situation.

2 The MONSUN Concept

While in classical approaches to telemanipulation control device and work unit have the same kinematics, in MONSUN (**M**anipulator **Con**trol **S**ystem **U**tilizing **N**etwork Technology) the so called "universal master concept" is used, which allows to combine application specific work unit kinematics with suitable MMI designs. MONSUN introduces by means of standardized interfaces an "open systems concept"; the micro-computer based controllers of the "work unit" (i. e. the surgical manipulator) and the force reflecting input device are located at their respective sites and communicate via standard LAN-interfaces by means of exchanging standardized control packets; all reference coordinates are Cartesian. The feedback control loop between the input device and the work unit is closed via the local area network. Multiple and possibly different work units and input devices can be connected to the same network and may be associated as required. An existing system can be "easily" extended by means of adding additional work units and input devices, if their controllers provide the required standard interfaces.

The first MONSUN prototype was built in 1990 [1]. It served to interconnect a force reflecting master manipulator to different industrial robots for programming the robots and for direct manipulation. The information that is presented to the operator through feedback consists of a 3-D video overview of the remotely manipulated scene, a close-up video of the robot tool, audio information captured by microphones at the remote site and force reflection. All control information is transferred via Ethernet that offers a shared bandwidth of 10 Mbit/s, but which does not provide for isochronous transmission. Thus video and audio information are transmitted separately using fiber optical links [2]. In the last year this system was extended by two MIS specific work units [3]

- an endoscopic positioning unit and
- a work unit for laparoscopic surgery.

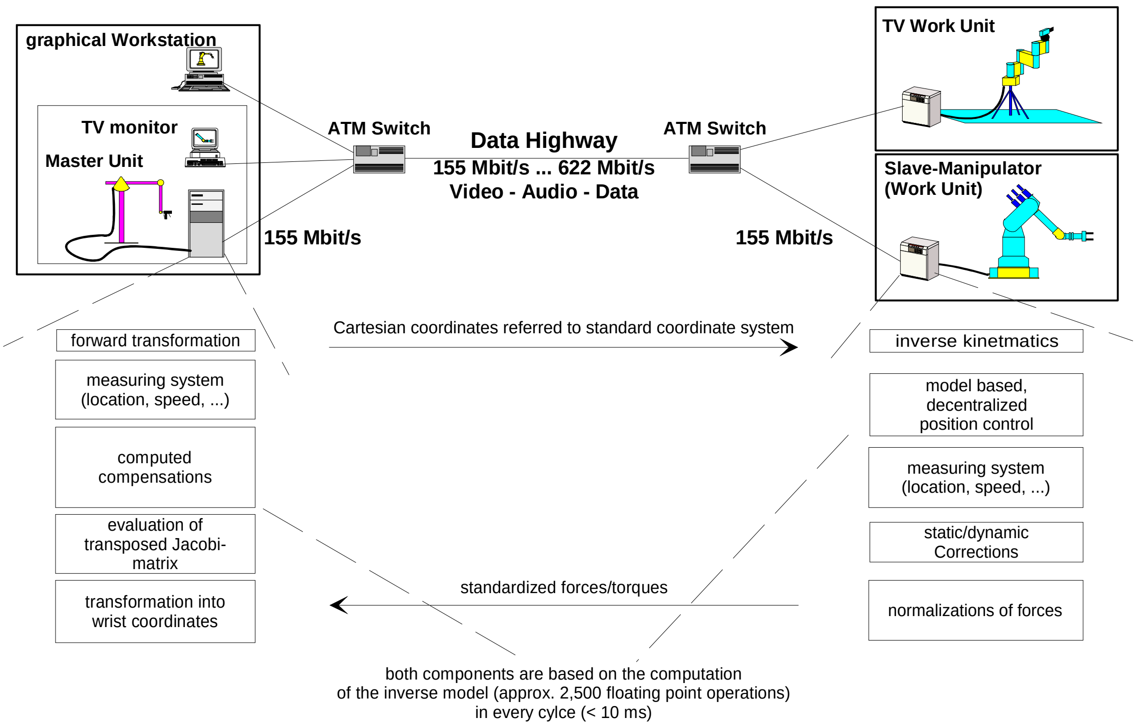

Fig. 1: A Typical Work Unit – MMI Configuration in MONSUN

A typical MONSUN scenario is shown in figure 1, where two remote work units, one of them being a camera positioning unit (TV work unit, TWU), are controlled from a control room equipped with monitoring and master units via a (local) ATM network based on two ATM switches.

Within ARTEMIS (figure 2) MONSUN will allow the surgeon to do his work from a micro-computer based workstation that runs a particular man/machine interface (MMI), thus integrating him in the feedback control.

The MMI has to be designed to put the surgeon into a rather similar scenario as he is used to from direct surgery. This is done by providing multimedia information consisting of

- real-time 3-D video with corresponding audio from the endoscope camera as well as from surrounding cameras that for example give an overview of the operation room or show the patient close up,
- audio signals from the supervising instruments, e. g. the EKG or the breath pump,
- diverse digital data arising from sensors like thermometers or sphygmomanometers and
- special sensor based force, torque and pressure reflection.

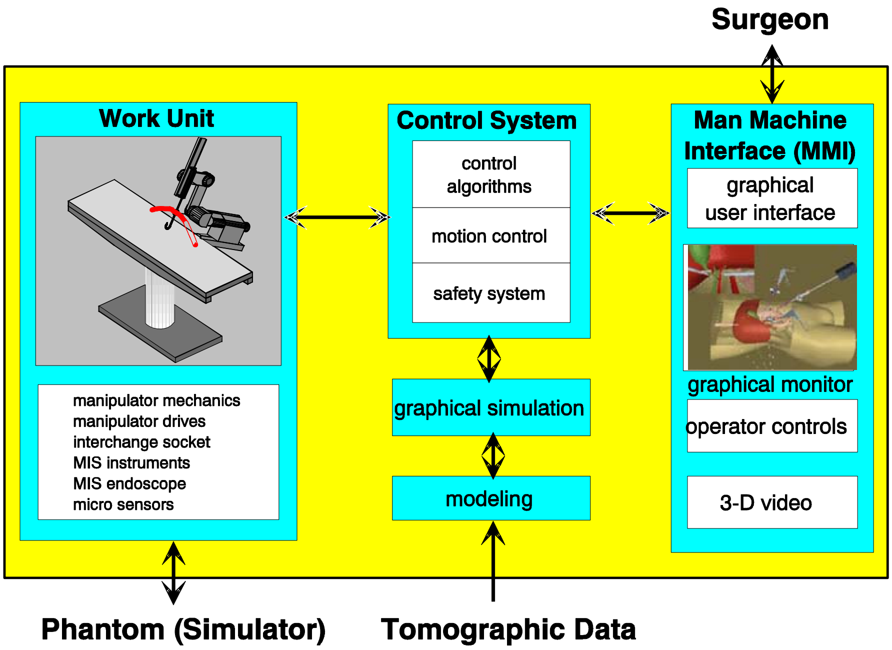

Fig. 2: System Concept of ARTEMIS and functional overview

The effector and the surgical work unit are controlled by means of an input device or master unit (MU) which allows the surgeon to use similar gestures for input as he would use on the surgery tools themselves. The endoscope carrier or guidance system is another work unit controlled via MONSUN; the surgeon may choose either automatic endoscope tracking or control via MMI inputs.

All work units are controlled in Cartesian coordinates, and the returned position, velocity and acceleration data refer to Cartesian coordinates as well. The necessary transformations must be provided by the WU controllers and the MU controller.

The interconnection between any MMI device including master manipulator, control devices, monitoring devices etc. and the remotely controlled work unit equipment – slave manipulators, camera positioning systems etc. – is built up using standard LAN adapters for transmission of control and monitoring information.

3 Bandwidth Requirements for Multimedia Components in MIS Telepresence Systems

A MONSUN based MIS telepresence system as proposed above requires a rather large transmission bandwidth if control, video and audio information are to be transmitted via the same communication medium. Transmission channels have to be provided for

- 3-D video data streams having their sources in the 3-D video endoscope and in supervising cameras at different locations in the operating room,
- digital audio data coming along with the video data of the supervising cameras and from separate microphones
- special digital audio data generated by medical supervising instruments like EKG writers,
- digital data (set point values and measured variables) for the control loop and force reflection data originating from special sensors at the surgical effector for sensible, scaleable feedback to the master manipulator
- digital data arising from other sensors (temperature, pressure, etc.).

3.1 3-D Video Data

In principle the 3-D video developed by KfK uses conventional TV technology:

There are two video cameras that have their lenses arranged in a natural manner referring to the angel of the optical axes thus producing two separate video channels for the left and the right eye. The camera electronic generates data at 25 frames per second on each channel. A special logic time-multiplexes the two video streams (left and right) into one conventional TV channel that is transmitted with an additional left-right-synchronization signal. On the receptors side the video signal is fed via a special video controller to a 100 Hz monitor thus displaying each of the two channels in an effective refresh rate of 50 frames per second. The additional synchronization signal is now used to remotely control high speed LCD shutter spectacles and to close the right glass during the display phase of the left video channel and the left glass during the display phase of the right video channel. Doing the multiplexing at the source of the video streams keeps the transmission relatively simple, i. e. standard video transmission technique can be used.

3.2 TV and Video Bandwidth Requirements

The resulting bandwidth requirements can thus be computed corresponding to non compressed TV bandwidth:

A normal CCD video sensor has a resolution of 403 lines with 512 pixels. Each pixel is represented by a 21 bit combined luminance/chrominance value. The CCD sensor produces 25 pictures per second. So the required digital bandwidth is

$$403 \text{ lines} \cdot 512 \text{ pixel/line} \cdot 21 \text{ bit/pixel} \cdot 25 \text{ Hz} = 103.31 \text{ Mbit/s}.$$

TV of today has a ratio of 4:3 in width to height. In Europe it contains 625 lines in height which leads to an approximate number of 833 pixel in each line. Frames are generated at 25 Hz which means 64 μs duration for each of the 625 lines. The sampling rate is 13.5 MHz. The required video bandwidth that can be computed from this values is 6.75 MHz and the resulting transmission bandwidth 13.5 MHz. The BAS signal actually uses a video bandwidth of 5 MHz and the common FBAS signal 4.43 MHz [4].

These are values for analogue TV transmission; for transmission via a digital communication network we have to digitize the information. This can be done in a closed code, where all the components building a signal are converted into one digital signal. Calculating with the values of the signal described above we get the digital bandwidth

$$2 \cdot 6.75\ \text{MHz} \cdot 8\ \text{bit} = 103\ \text{Mbit/s}.$$

With the closed digital coding of the FBAS signal the sampling frequency used is four times the frequency of the color carrier which leads to

$$4 \cdot 4.43\ \text{MHz} \cdot 8\ \text{bit} = 135.19\ \text{Mbit/s}.$$

The values that are published recently vary in narrow intervals for PAL but give a common rate for High Definition TV (HDTV): Tetzner [5] tells a PAL closed coded data rate of 216 Mbit/s whilst Großmann and Müller [6] evaluate 108 Mbit/s for PAL. The rate computed for HDTV is 1.2 Gbit/s.

3.3 Video Data Compression

The technique used to reduce the required bandwidth is data compression. Compression methods vary in their purpose – compression of single pictures or compression of scenes of moving pictures – and in the mode – symmetric methods that have the same expenditure when compressing as when decompressing or asymmetrical methods that have more expenditure at compression than at decompression. Common standards are the JPEG (Joint Photographic Experts Group) method for single pictures and its pendant for motion pictures MPEG (Moving Pictures Experts Group) whereof the first is a symmetrical and the latter is a asymmetrical method [7]. A symmetrical standard for motion pictures is the H.120/H.261 proposal of the former CCITT.

Using compression the required digital bandwidths can be reduced for PAL from the range of (108 … 216) Mbit/s to about (2 … 8) Mbit/s and from 1.2 Gbit/s to 140 Mbit/s for studio needs or 34 Mbit/s for distribution needs in the HDTV sector.

In our special application – transmitting 3-D video data – compression techniques must be used carefully. The reason is that the multiplexed 3-D information results in slight differences between consecutive frames, which alternately rise from the left camera and the right camera. All compression methods that base only on coding changes of scenes will fail. Compression methods like JPEG that take into account exactly one frame with no reference to the motion scene will be acceptable.

3.4 Digital Audio and Control Data

Compared to the bandwidth requested by the video component all other data do not change the transmission profile basically:

Audio data is transmitted with a bandwidth of 112 kbit/s in digital terrestrial broadcast applications which is sufficient for a CD like quality. However, any channel used for audio transmission must provide for isochronous transmission.

A bandwidth of 1 Mbit/s is enough to carry all digital control and telemetry data, sensor data resulting from sensors like force/torque data, temperatures or pressure data. These data are currently transmitted via the Ethernet in the existing system, which offers a shared bandwidth of 10 Mbit/s.

4 Demonstration of Telemanipulation via a "Kopernikus" Satellite Link

A demonstration of a multimedial telemanipulation via digital high speed communication channels has been arranged during the International Symposium of the Society for Minimal Invasive Therapy (SMIT) in Berlin in October, 1994 [3].

A computer controlled endoscope positioning unit located at Berlin was remotely controlled by means of an input device located at Heidelberg. The link was established via the *"Kopernikus"* satellite and provided a video transfer rate of 34 Mbit/s and a control data transfer rate of 2 Mbit/s.

The master controller for the endoscope positioning system consisted of a SGI workstation running the VR system KISMET (KfK-development). Using KISMET a graphical simulation showed the model of the endoscope positioning system in real-time in its working position over a patient's body. In addition there was a video monitor connected to the reverse video channel which showed the real system in Berlin in real-time via the satellite link. The scenario is shown in figure 3.

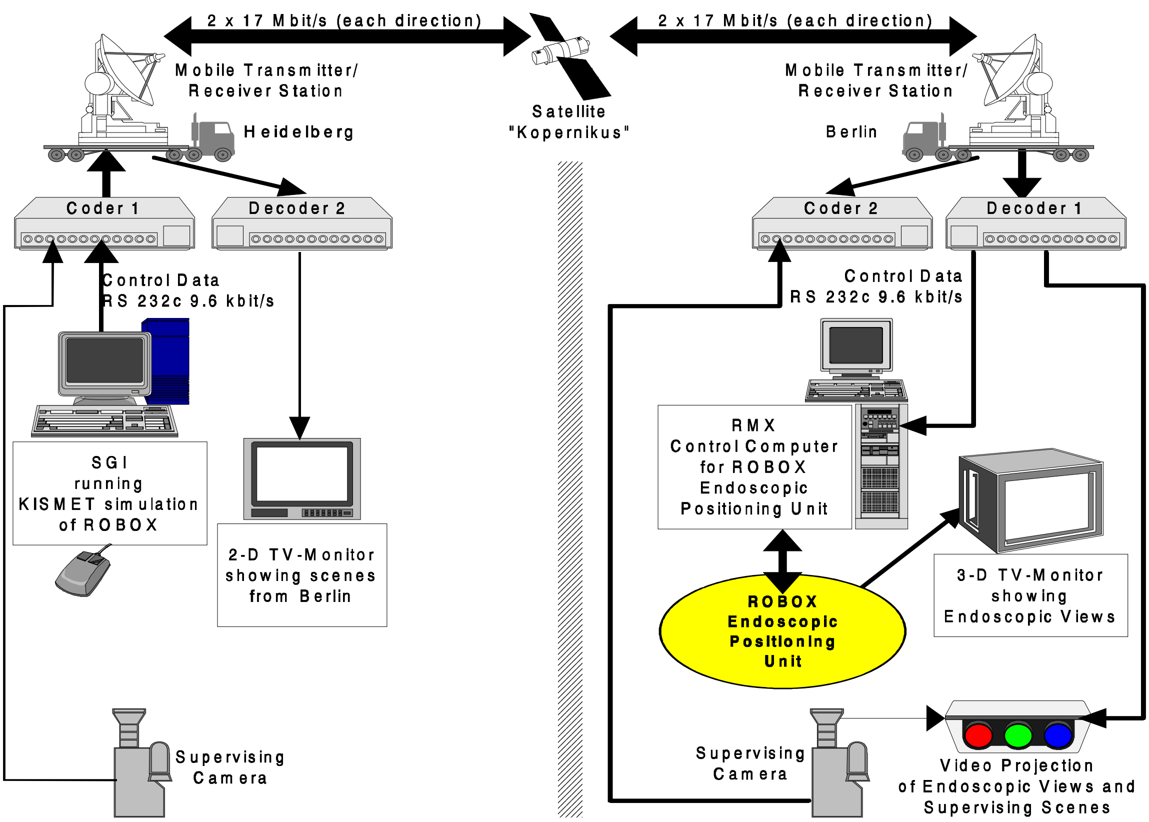

Fig. 3: Demonstration Scenario of the Endoscope Positioning Control via "Kopernikus" Satellite Link during the SMIT 1994 conference

This demonstration showed the essential importance of video transmission bandwidth and the influence of transmission delays: the satellite link introduces a basic transfer delay of 150 ms in each direction which is needed to transport the signal over the approx. 36,000 km from sender to the satellite and back from the satellite to the receiving station. In addition to this fixed delay a variable delay of up to 1 s resulting from compression/decompression and the switching and multiplexing characteristics of the satellite.

To guarantee a reliable real-time video stream that can be used as feedback to the operator in the control loop time delays between an actual event and the time this event gets displayed on the monitoring screen must be minimized. As experiments at KfK with a force reflecting telemanipulator have shown, this delay must not exceed 200 ms if delicate manipulation tasks have to be performed.

If we take into account the influence of this delay limit on the high speed transmission network implementation of MONSUN we obtain an important requirements specification which must be fulfilled in case where compression methods have to be applied: if video frames are generated at a rate of 25/s the mean frame time is 0.04 s; this is equivalent to 5 frames corresponding to the maximum delay of 0.2 s.

5 ATM – Asynchronous Transfer Mode – for the Integrated Transport System

5.1 ATM Overview

The Asynchronous Transfer Mode (ATM) is the standardized transmission method for the IBCN, the Integrated Broadband Communication Network.

It is intended to replace all the different techniques used on LAN, metropolitan (MAN) or wide area (WAN) networks and the Global Area Networks (GAN) in the further future [8].

The principle in ATM is to offer switched high speed carriers and connection oriented links at negotiated bandwidths and qualities of service (QOS) to a wide range of users for different purposes.

The link between the users or the client and a server is built up in a connecting phase via a recently standardized signaling protocol. In this signaling phase the client negotiates the QOS of the link by requesting link parameters which both the server and the network will have to accept or to refuse, offering different values themselves which the client might accept or refuse as well. The main parameters to negotiate are the exclusive bandwidth for the connection and the class of connection that is to be established. Those service classes are standardized by the ITU, the former CCITT, as it is shown in Table 1 below. The titles of the presented classes derive from the ATM reference model where a specific layer has to adapt application data to ATM data formats. This layer is called "ATM Adaption Layer" (AAL) and is built up of the five classes presented in the table.

service class	AAL 1	AAL 2	AAL 3	AAL 4	AAL 5
time correlation btw. partners	**required**		**not required**		
bit rate	**constant**	**variable**			
connection mode	**connection oriented**			**connectionless**	**connection oriented or connectionless**
examples	video transmission, telephone	variable audio transmission	connection oriented data transfer	connectionless data transfer	high speed data transfer

Table 1: ATM Service Classes

The connection established is called a *virtual channel* (VC) that offers an exclusive communication to the two connected partners with the appropriate QOS and bandwidth. The QOS and bandwidth are reliable after the link has been established as long as it exists. All the needed resources are reserved until the termination of the connection.

ATM is quite different from conventional LAN techniques where the existing medium is granted completely to one station at a time, leaving the others waiting. The shared media variants thus can not offer a computable maximum delay which is essential for real-time applications and the bandwidth available to a station depends on the total number of attached stations.

5.2 ATM Network Structure

The structure of an ATM network has typical centralized components. The users are connected to the network via *ATM switches* that offer the transports and form the network by being interwired to other ATM switches. Through protocol converters existing conventional networks can be attached to ATM switches as well thus offering a attractive migration path from LAN to ATM.

5.2 Physical Access to ATM

In the User Network Interface (UNI) definition for ATM the different physical access variants are presented. There are electrical and optical joinings at various speeds and transfer modes. They can all be offered in one switch and may change according to requests by users and applications. The main attachment variants are [9]:

- TAXI 4B/5B (100 Mbit/s, 140 Mbit/s) for multimode fiber (MMF) using the FDDI transfer mode for ATM
- STS-3c (155.52 Mbit/s) optical interface for singlemode or multimode optical fiber (SMF, MMF) also known as SONET interface
- DS3 (44.736 Mbit/s) electrical interface for shielded twisted pair cable (STP)
- E3 (34.368 Mbit/s) and E4 (139.264 Mbit/s) European electrical interfaces for coax cable
- UTP (unshielded twisted pair copper cable) Category 3 attachment at 52 Mbit/s and Category 5 attachment at 155 Mbit/s.

5.3 The ATM Cell Structure

An essential aspect of ATM is that data portions are cut into *cells* of fixed length.

The advantages of this method are:

- the passing time of the cells through the switch are small and constant,
- due to the cell technique the mediation can be realized in fast hardware,
- the small extent of the cells (53 Byte overall, 48 Byte payload) allows to keep the buffers of the transit switches small and fast,
- as there is an negotiated connection established during the link establishing phase all the cells only have to get examined in their header fields (5 Byte).

A packet oriented system like an Ethernet bridge or router has to collect all of an incoming packet and has to write it into a buffer. This lasts 1,210 µs for a maximum packet length of 1,512 Bytes at Ethernet rate of 10 Mbit/s. Thereafter the packet gets analyzed in another 80 µs in average. This makes a passing time of 1,290 µs.

The cell oriented ATM system instead only has to collect the incoming destination address of a cell and then immediately starts delivering the cell data over the appropriate port. The collecting phase lasts 10 µs at maximum and the evaluation phase in average will last 40 µs. This makes an overall passing time of 50 µs.

Cells are routed via an 8 bit *virtual path identifier* (VPI) and a 16 bit *virtual channel identifier* (VCI) in the cell header, which also contains a 1 bit *cell loss priority* (CLP) that indicates that the whole cell might be removed under heavy load depending on the negotiated QOS. As mentioned above the main part of a cell is the 384 bit payload.

5.4 The Projected Test Configuration at KfK

ATM offers a widely scaleable, reliable transport of isochronous data as well as of synchronous or asynchronous data and has computable delays that allow real-time applications being controlled via the integrated network.

So KfK decided to test a small ATM switch and to port the existing MONSUN implementation onto that switch.

The overall capacity of the switch to be applied is 2.5 Gbit/s which will be sufficient to carry uncompressed video data at approximately 155 Mbit/s over a dedicated but integrated channel – an ATM VC – and the other multimedial data that are currently transferred via Ethernet. The main effort that has to be done is porting standard MONSUN interface of the existing system to ATM.

The current interface implementation uses an OSI/ISO protocol stack. The application uses OSI layer 4 – the transport layer – as a socket. ATM switches usually are integrated in a TCP (Transmission Control Protocol) or TCP/IP (Transmission Control Protocol/Internet Protocol) protocol hierarchy which despite a general analogy has significant differences from ISO/OSI in detail.

Two ATM components are available today which provide standard video single cast or multicast delivery through *permanent virtual channels* (PVC) via an ATM network by simply plugging in the video camera to a coder on one end and plugging in a monitor at the corresponding decoder or using the screen of an attached workstation as the output on the other side. First tests with this equipment have been made showing a noticeable delay and loss of frames when displaying the video on the monitor of the workstation. However, it could be clearly shown that frame droppings were not caused by the ATM switch, since the same deficiencies occurred when bypassing the ATM switch and directly connecting the video coder to the ATM adapter of the workstation. These tests will be extended to determine the bottleneck.

6 Conclusions

At a first glance ATM based high speed communication seems to be an ideal basis to meet the communication system performance requirements of telepresence systems for medical applications like MIS via telemanipulation. However these applications introduce a fourth dimension into multimedia communications: the dimension of interaction in real-time. Time critical interaction requires very low turnaround delay times (< 200 ms) for all vital information conveyed to the interacting “man in the loop”, the surgeon. Obviously, ATM based in-house networks offer sufficiently low signal transmission and switching times for closed loop control and audio signal transmission. But video transmission via ATM based networks in telepresence systems for MIS seems only feasible at present, if uncompressed data transfer via an allocated channel with sufficiently high bandwidth is possible. The influence of state of the art compression techniques on video transmission in telepresence applications need further investigations.

References

[1] Helmut Breitwieser; Übersicht über das Universalmaster-Steuerungssystem MONSUN, Primärbericht. Kernforschungszentrum Karlsruhe GmbH, 1991.

[2] Elmar Holler, Helmut Breitwieser; Telepresence Systems for Application in Minimally Invasive Surgery. Proceedings of Medicine Meets Virtual Reality II. San Diego, 1993.

[3] Elmar Holler, Hermann Rininsland; Trends in Manipulator and Robotic Development. Proceedings of the Medtech '94 and Satellite Symposium of the VI. international Meeting of the SMIT. Berlin, 1994.

[4] Ralf Steinmetz; Multimedia-Technologie, Einführung und Grundlagen. Springer Verlag, Berlin, 1993. ISBN 3-540-56724-0.

[5] Karl Tetzner; Das Digital-Fernsehen als Einstieg in Multimedia. dialog, Juli/August 1994, pp. 4–6.

[6] Paul Müller, Hans Peter Großmann; Anforderungen multimedialer Kommunikation. Praxis der Informationsverarbeitung und Kommunikation (PIK) 1/94, pp. 13–19. ISSN 0930-5157.

[7] Georg R. Hofmann; Imaging, Bildverarbeitung und Bildkommunikation. Springer Verlag, Berlin, 1993. ISBN 3-540-56373-3.

[8] Martin de Prycker; Asynchronous Transfer Mode, Die Lösung für Breitband-ISDN. Prentice Hall Verlag GmbH, München, 1994. ISBN 3-930436-03-5.

[9] Klaus Eppele, Hans Lackner; Der Weg zu ATM. Arbeitspapier der Firma HiLAN GmbH, Karlsruhe, 1994.