Potential and Limitations of a Teleteaching Environment based on H.323 Audio - Visual Communication Systems [1]

Pantelis Balaouras, Ioannis Stavrakakis, Lazaros Merakos
Communication Networks Laboratory (CNL)
Department of Informatics
University of Athens (UoA)
15784, Athens, Greece
{P.Balaouras, istavrak,merakos} @di.uoa.gr

Abstract: The effective support of teleteaching services requires the development of multimedia collaboration systems that are capable of providing real-time and high quality audio-visual communication among distributed instructors and students. In the absence of such specialised systems, technologies tailored to other services are being considered for teleteaching services as well. Such a technology is the H.323 audio-visual communication technology developed to support video communication over IP. Although teleteaching and videoconferencing have similar QoS requirements, teleteaching functional requirements are a superset of those of videoconferencing. The suitability of H.323 technology and its currently available products for supporting teleteaching services is investigated, based on experience gained during a related deployment at the University of Athens.

Keywords: Teleteaching environment; H.323; Multipoint Conferences; IP based teleteaching; Synchronous

1 Introduction

Teleteaching services are expected to enable the synchronous participation in a lecture of a geographically dispersed set of students. The participants would be located in geographically dispersed facilities such as lecture theatres, classrooms, seminar rooms, laboratories, or at home. Teleteaching is expected to improve the quality of teaching especially in cases where an expert instructor is required. Another major fact that drives the deployment of teleteaching services is that the instructor's personality and instructional talent can not be substituted by any virtual environment or asynchronous telelearning service.

Networked multimedia applications and systems are required to implement teleteaching scenarios. Several teleteaching scenarios has been implemented in Europe by employing products based on MJPEG/ATM standards [11,18].

The proliferation of IP technology has provided a widely available networking infrastructure that can easily - and in a cost effective manner - reach potential students. For this reason, it is meaningful to try to develop teleteaching services over IP, particularly now that the QoS provision capabilities of IP are most likely close to materialise.

Two different application suites and architectures exist for multimedia collaborations over IP. The Mbone tools [12] and the H.323 audio-visual communication systems [1]. While the Mbone tools provide for a cost effective, medium quality audio visual communication, mainly for the desktop, they are not mature enough to support a teleteaching environment, especially in a classroom setting. Currently available H.323 products appear to be mature for collaboration services and, thus, could be used to support teleteaching services as well. Such a consideration should be investigated carefully though, since the teleteaching functional requirements can be very different from those of videoconferencing or of a simple collaboration environment.

Presently, the Communication Networks Laboratory (CNL) of University of Athens (UoA) participates in two national projects [16,17], funded by the Greek Ministry of Education and the European Commission, which aim at evaluating and exploring the potential of teleteaching scenarios based on H.323 systems. To this end, the deployment of a teleteaching environment based on H.323 systems in the Athens metropolitan area is in progress.

An overview of the H.323 recommendation [1] is presented in section 2, focusing on the multimedia stream exchange and processing capabilities of H.323 entities. The specification that are currently supported by H.323 products are identified and discussed in section 3.

1 This work and its dissemination efforts have been supported in part by the IST Program of the European Union under contact IST-1999-10160 (VideoGateway).

In section 4, it is discussed how various H.323 components may support several teleteaching scenarios and some design limitations are identified. A set of functional requirements for an effective teleteaching environment are presented in section 5, where the limitations of present H.323 products are discussed and some solutions to overcome these limitations are provided. Finally, some enhancements to present solutions are provided in section 6, in order for the supported teleteaching scenarios to meet the functional requirements of a teleteaching environment.

2 The ITU-T H.323 Recommendations

The International Telecommunications Union (ITU) has developed the H.323 set of standards for audio-visual and data real-time communication over IP networks [1].

The H.323 recommendations define a model that includes four types of H.323 entities, each having a particular role: the *H.323 terminal*, *the H.323 Multipoint Control Unit (MCU)*, the *Gateway* (e.g., H.323 to H.320 Gateway), and the *H.323 Gatekeeper (GK)*. The first three types are characterised as *H.323 end-points*. An H.323 endpoint generates and/or terminates streams and can call other entities and be called. An overview of ITU-T H.323, and a description of the role of each H.323 entity type is presented in [14].

In this section, we are interested in and focus on the capabilities of the H.323 entities for audio, video and data stream exchange and processing, as well as on control issues at the functionality level and device control.

2.1 Capabilities of H.323 terminals

The ITU-T H.323 recommendations define the following capabilities of H.323 terminals.

2.1.1 Audio issues

All H.323 terminals must have an audio codec. An H.323 terminal may optionally send more than one audio channel at the same time and receive more than one audio channel at the same time, which may be necessary for distributed multipoint conferencing. In these cases, the H.323 terminal may need to a) support a multicast communication mode and b) perform an audio mixing function in order to present a composite audio signal to the user.

Each H.323 terminal should be capable of encoding and decoding speech according to Recommendation G.711, and be capable of transmitting and receiving A-law and μ-law coded audio streams. A terminal may *optionally* be capable of encoding and decoding speech using Recommendations G.722, G.728, G.729, MPEG 1 audio, and G.723. Table 1 presents the mandatory and optional capabilities of H.323 terminals.

The H.323 terminal should be capable of asymmetric operation for all audio capabilities it has declared within the same capability set. For example, it should be able to send G.711 and receive G.728 if it is capable of both. The capability set of a terminal includes its encoding and decoding capabilities.

2.1.2 Video issues

The video codec is optional for H.323 terminals. The H.323 terminal may optionally send over more than one video channel at the same time, and receive from more than one video channel at the same time, which may be necessary for distributed multipoint conferencing. In these cases, the H.323 terminal may need to a) support a multicast communication mode and b) perform a video mixing or switching function in order to present the video signal to the user. This function may include presenting the video from more than one terminal to the user.
All H.323 terminals supporting video communications should be capable of encoding and decoding video according to H.261 QCIF [2]. Optionally, a terminal may also be capable of encoding and decoding video according to H.261 CIF or H.263 SQCIF, QCIF, CIF, 4CIF, and 16CIF [3]. If a terminal supports H.263 with CIF or higher resolution, it should also support H.261 CIF. All terminals that support H.263 should also support H.263 QCIF.

H.323 terminals should be capable of operating in asymmetric video bit rates, frame rates and, if more than one picture resolution is supported, picture resolutions. For example, this will allow a CIF capable terminal to transmit QCIF while receiving CIF pictures.

2.1.3 Data issues

One or more data channels are optional. The data channel may be unidirectional or bi-directional depending on the requirements of the data application.

		Audio	**Video**	**Data**
Mand.	#streams	1	0	0
	Coding schemes	A-law, μ-law, G.711	H.261 QCIF	T.120
OptIonal.	#streams	N>1	N>0	N>0
	Coding schemes	G.722, G.723, G.728, G.729, MPEG1 audio.	– H.261 CIF – H.263 SQCIF,QCIF, CIF, 4CIF 16CIF	

Table 1: Coding schemes and number of simultaneously sending and receiving streams.

2.1.4 H.245 procedures

Each H.323 terminal uses the procedures defined in H.245 [7] for:

a) exchanging its decoder capability set (video and audio bit rate, audio and picture format and algorithm options that can be accepted by the decoder), and
b) indicating how many simultaneous audio and video streams it is capable of decoding.

The encoder of a terminal is free to transmit anything that is within the decoder's capability set.

2.2 H.323 Multipoint Conferences

The H.323 Recommendations define *the multipoint controller (MC)* and *processor (MP)* to support multipoint conferences.

2.2.1 Multipoint controller and processor

The MC provides control functions to support multipoint conferences between three or more H.323 endpoints, and point to point conferences that may become a multipoint conference. The MC may provide conference control functions such as *chair control*, *video selection* and *video broadcast*. The MC does not perform mixing or switching of audio, video and data streams.

The MP receives audio, video, and/or data streams from the endpoints involved in a centralized or hybrid multipoint conference. The MP processes these media streams and returns them to the terminals either in *point to point* or in *multicast* mode, provided that the MCU and terminals are capable of communicating in multicast mode.

The MP may perform *video switching* or *mixing*, *audio mixing*, *T.120 data distribution* [5] and *transcoding* between different audio, video and data formats and bit rates in order to support conferences between H.323 terminals with different communication modes.

2.2.2 Types of multipoint conferences

The H.323 standard supports three types of multipoint conferences: centralized, decentralized, and hybrid.

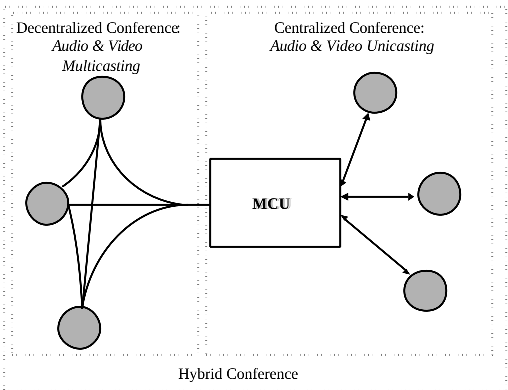

Figure 1:Centralized, decentralized and hybrid conferences

In a *centralized conference*, an MCU is required. The MC and MP are located at the MCU. The MC controls the conference. Each terminal transmits its video, audio, and data streams to the MCU. Then, the MP processes the streams, and distributes them back to the terminal either in unicast (point to point) mode or multicast mode, provided that the MCU and terminals are capable of communicating in this mode.

In a *decentralized conference*, each terminal multicasts its audio and video streams to all the workspaces eliminating the need to engage a central point. The MC provides control functions and may be located at any type of endpoint. Since the terminals receive the multicast audio and video streams, an MP has to be located at the terminals to perform audio mixing and select one or more video streams to be displayed.

A *hybrid conference* combines elements of both the centralized and the decentralized one. In a *hybrid multipoint, centralized-audio conference*, centralized multipoint is used for audio and decentralized multipoint for video. In a *hybrid multipoint, centralized-video conference*, centralized multipoint is used for video and decentralized multipoint for audio.

The capability of supporting centralized multipoint conferences is mandatory for H.323 terminals and MCU. The capability of supporting decentralized and hybrid conferences for H.323 terminals and MCUs are optional. Table 2 summarises the characteristics (capabilities, communication modes and potential locations of MC and MP(s)) for each multipoint conference type.

Conference Type	**Centra -lized**	**Decentra-lized**	**Hybrid-centralized audio**	**Hybrid-centralized video**
Capability *	M**	O**	O	O
Communication modes				
Audio	P2P	Multicast	P2P	Multicast
Video	P2P	Multicast	Multicast	P2P
T.120 Data	P2P	P2P	P2P	P2P
Potential locations of				
MC	MCU	MCU, terminal, GK, Gateway	MCU	MCU
MP	MCU	Terminals	MP-audio: MCU MP-video: Terminals T.120 MP: All endpoints	MP-audio: terminals MP-video: MCU T.120MP: All endpoints
Hints: (*)Capability of Terminals and MCU in supporting the conference type, **(**)**M for mandatory, O for Optional.				

Table 2: Summary of multipoint conference types characteristics

2.3 Capabilities of the H.323 MCU

A typical MCU supports centralized multipoint conferences and consists of an MC and an MP, as indicated earlier. The MP performs video switching or mixing, audio transcoding and mixing, and T.120 data distribution.

Each terminal transmits its audio and video streams to the MCU. The MP performs synchronisation and audio mixing of the received audio streams and distributes back in point to point mode the mixed audio stream to each participant. The MP performs either video mixing, if supported by the MCU, or video switching, activated by voice or a chair control mechanism (**Figure 2**). In the case of video mixing, the mixed stream is transmitted in point to point mode to each participant. In the video switching case, the speaker's video stream (Sa in Figure 2) is transmitted to each participant, while simultaneously the speaker's terminal may be receiving another video stream (Sb in Figure 2).

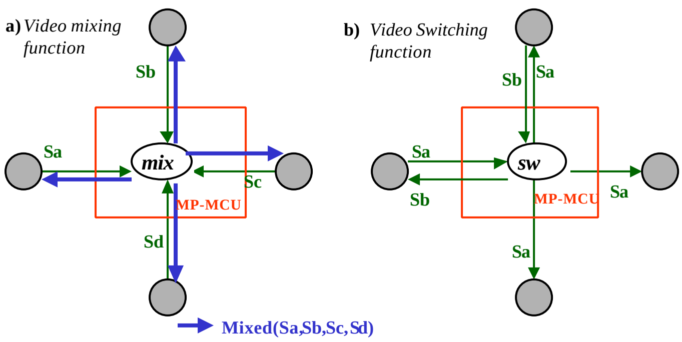

Figure 2: Switching and mixing processing functions performed by a MP located at MCU.

The MC provides conference control functions, which are discussed bellow.

2.4 Control issues

2.4.1 Chair control

The chair control procedure is defined in H.243 [3]. In order for the chair control feature to be used in a multipoint conference, both the H.323 terminals and the MCU should support this option. The chairperson is able to manage the conference (e.g., drop participants) and select which video stream is transmitted from the MCU to each video capable terminal.
More specifically, the chair control terminal in an H.323 conference has the following visualisation-forcing capabilities:

– *video broadcast*, to determine which one of the available video streams will be transmitted to all terminals, except from the instructor's terminal;

– *multipoint visualisation*, to broadcast the chair terminal's video to all participating terminals;

– *video select*, to determine which video stream should be transmitted to the chair terminal.

The chair control terminal receives the floor requests of the other terminals.

2.4.2 Floor control

Floor control is possible only if chair control is activated. If instead of chair control, voice-activated switching or video mixing is used, then the conference is in open conversation mode, in which everyone can speak and their audio will be received by all participants.

2.4.3 Camera control

Camera control is dealt with in the H.281 standard [8]. A moderator, who may be either the instructor or a technician, is able to control both the local and remote cameras that are connected to an H.323 terminal that supports far-end camera control (FECC).

3 Capabilities of present H.323 products.

3.1 H.323 terminals

Today, H.323 terminals are classified by the market as desktop and room systems. Both types are capable of sending and receiving one audio and one video stream, but only in point to point communication mode. Multicast mode is not supported yet.

The H.323 terminals may display the local or receiving video stream in small, medium or full screen window mode. "Picture in picture" (PIP) window mode is supported for a second video signal.

In Table 3, the technical profiles of currently available H.323 room and desktop systems are summarised [19,20,21].

Characteristics	H.323 room terminals	H.323 desktop terminals
Audio & video inputs	1xAudio, 2xS-video, 1xcomposite	1xAudio, 1xS-video or composite, 1xComposite or proprietaty
Selection function in video inputs	Yes	
Audio Coding Schemes	{G.711,G.722,G.723,G.728}	
Video Coding Schemes	{H.261 Sub-QCIF, H.261 QCIF, H.263 Sub-QCIF, H.263 QCIF, H.263}	
Type of sessions	Centralized by H.323 MCU	
Comm. mode	point to point	
Display windows	Small, medium, Full screen, PIP modes	
Outputs	Audio RGB s-video	Audio RGB
Far-end camera control	Y	some
Floor control	only via chair control	no

Chair control	Y	no
Display Devices	TV	PC monitor PC speakers
Data Support	*Yes, by T.120 Application Sharing Services*	
Underlying Network Technology	IP, ISDN, MVIP	IP and some ISDN
Platform	*PC based Windows*	

Table 3: Technical profiles of H.323 room and desktop terminals

For the transmission of the instructor's presentation data (e.g., powerpoint presentation, postscript slides, documents) to the remote terminals, the H.323 terminals provide application sharing services based on the ITU-T T.120 series of Recommendations.

3.2 H.323 MCUs

Today's H.323 MCU products support only the centralized multipoint conference type, which is defined as mandatory by the H.323 Recommendations. The H.323 MCU products support the distribution of audio and video streams only in point to point mode and not in multicast mode, resulting in a waste of network bandwidth. H.323 MCUs are presently supporting voice activated video switching, audio transcoding and mixing, but not video mixing yet. The chair control function is supported as well.

3.3 Quality of multipoint conferences

The quality of a multipoint conference depends on the audio and video coding standards supported by the H.323 terminals, the stream bit rate, and the capabilities of the H.323 MCU to support switching of multiple high bit rate video streams.

3.3.1 Control Issues

Only H.323 room terminals support chair control. In the case of centralized multipoint conference, chair control may be enabled provided that the H.323 MCU supports it.

Camera control is provided by all H.323 room terminals and some of the desktop terminals.

4 Teleteaching Environment

A *teleteaching environment* is expected to enable the synchronous participation in a lecture of a geographically dispersed set of students. The participants could be located in geographically dispersed facilities such as lecture theatres, classrooms, seminar rooms, laboratories or at home. These facilities are referred to as *teleteaching workspaces.*

A teleteaching session i (TSi) will involve a number of workspaces and necessitate certain interactions among them. Let *Wspaces(TSi)* denote the set of these workspaces. The interactions among the workspaces constitute the scenario of the teleteaching activity, referred to as *teleteaching scenario*.

Our aim is to present the set of interactions (i.e., the potential teleteaching scenarios), which can be supported by present H.323 products as well as identify the limitations of these products. For this purpose, a description of the available signals and teaching media in a teleteaching workspace is needed.

In the interest of saving space the standard equipment (capture and display devices, videoconference systems) of a typical teleteaching workspace are not described here. Such a description may be found in [10]. A notation concerning the teaching media sources, in analog (outputs of capture devices) and digital forms is provided in Table 4. This notation is used in this paper.

Nota-tion	Description
Wi	Instructor's workspace
MS	Media Server
Iv(W)	Instructor video signal
Ia(W)	Instructor audio signal
Sv(W)	Student(s) video signal
Sa(W)	Student(s) audio signal
Dp(W)	Data presentations
Dv(W)	Prerecorded Video stream in H.263, MPEG1 or MPEG2 format
Da(W)	Pre-recorded audio stream, e.g., mp3 streams
D(W)	Represents one of Dp, Dv and Da
V(W)	Visual presenter video signal
Mv(W)	VCR video signal
Ma(W)	A tape or CD audio signal

Table 4: Notation; the argument indicates the associated workspace

Figure 3 illustrates the devices associated with the analog and digital teaching media presented in Table 4.

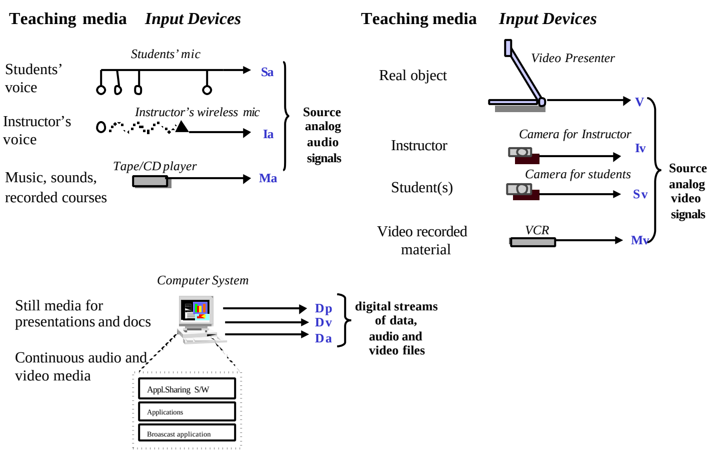

Figure 3: Teaching media and associated devices

4.1 *Required H.323 entities*

It is assumed that only currently available H.323 audio-visual communication systems are used to implement a teleteaching activity.

In addition to the H.323 terminals, an H.323 MCU and an H.323 to H.320 Gateway are required to deploy a teleteaching scenario based on H.323 and H.320 [4] audio visual communication systems (Figure 4). An H.323 Gatekeeper may optionally be used as well.

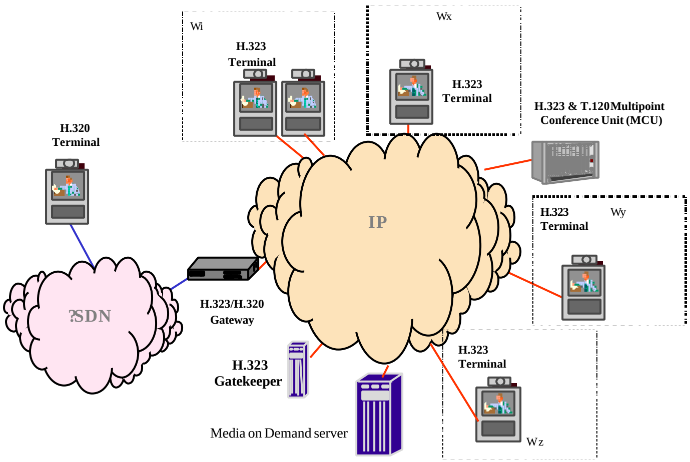

Figure 4: H.323 and H.320 components of a teleteaching environment

A workspace W may be equipped with one or more H.323 terminals depending on the teleteaching scenario.

For instance, multiple H.323 terminals are necessary if multiple streams (audio or video) are expected to be transmitted from workspace W. It is reminded that an H.323 terminal supports the transmission of only one audio and one video stream.

In a typical teleteaching scenario where both the instructor's and the student's figures and voices need to be transmitted from a workspace, two H.323 terminals are usually installed in the instructor's workspace; one for transmitting the instructor's audio and video, and one for transmitting the students' audio and video.

In the case in which only one H.323 room terminal is installed in a workspace, both the instructor's and student's cameras are connected to a single H.323 terminal. A visual presenter may be connected as well, since room terminals are equipped with three video inputs. Only one video input is selected as video source each time. If additional video sources exist, such as VCR, then a video matrix has to be installed in the workspace as well.

4.2 *Potential interactions among workspaces*

In the following paragraphs the potential interactions among the workspaces, which can be supported by currently available H.323 products, are discussed. Figure 5 provides an illustration of stream exchange.

In workspace Wi, the instructor's camera and a wireless microphone are connected to an H.323 room terminal; let Sys1(Wi) denote this terminal. The appropriate viewer for the data presentation(s) should be installed in Sys1(Wi). Instructor's video Iv(Wi), audio Ia(Wi) and data presentations Dp(Wi) streams are transmitted to the H.323 MCU from Sys1(Wi). The V(Wi) visual presenter signal could be transmitted, instead of Iv(Wi). The students' camera and microphone are connected to the second H.323 room terminal; let Sys2(Wi) denote this terminal. Students' video Sv(Wi), and audio Sa(Wi) are transmitted to the H.323 MCU as well.

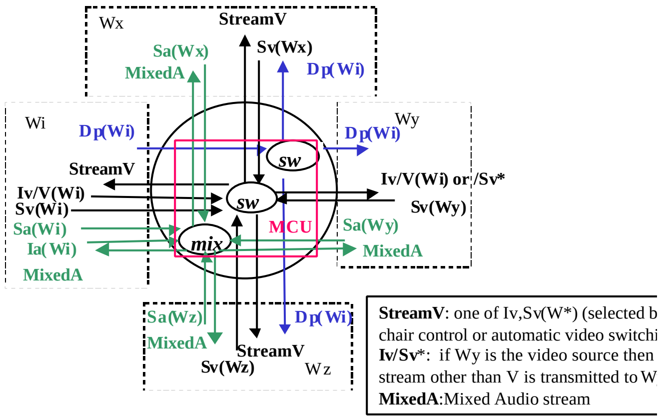

Figure 5: Stream exchange among the workspaces in the H.323 teleteaching scenario.

In each of the remote workspaces Wx, Wy, Wz, a single H.323 terminal is assumed to be installed; let Sys1(W) denote this terminal. The students' camera and microphone are connected to Sys1(W). Students' video Sv(W), and audio Sa(W) are transmitted to the H.323 MCU as well.

Summarising, the H.323 MCU receives the set of video streams Vset={Iv(Wi), Sv(Wi), Sv(Wx), Sv(Wy), Sv(Wz)}, the set of audio streams Aset={Ia(Wi), Sa(Wi), Sa(Wx), Sa(Wy), Sa(Wz)}, and the data stream Dp.

Present H.323 MCU products are capable of performing the following functions:

- *Audio mixing* of audio streams and distributing the produced stream, denoted as MixedA, back to each workspace in point to point mode.
- *Video switching*; one video stream from the set Vset is selected and transmitted back to the H.323 terminals in point to point mode. Let StreamV denote this stream. There are two ways for selecting the stream. An automatic way, in which a voice-activated mechanism selects the speaker's video for transmission. Altenatively, a chairperson selects which video stream should be transmitted to the

remote terminals, provided that H.323 room systems and MCU support chair control.

- *Distribution* of the data stream Dp to all the remote terminals in point to point mode.

Thus, each participating H.323 terminal receives three streams: the audio stream MixedA, the video stream StreamV and the data stream Dp. The stream MixedA is decoded and channelled to the speakers, while StreamV and Dp are displayed on one or two monitors. The processing functions taking place in the MCU, and the streams distributed by the MCU to the terminals of the workspaces are illustrated in Figure 5, for the case of automatic video switching.

5 Functional requirements

In this section, a set of teleteaching functional requirements is presented and the degree of fulfilment by the currently available H.323 products is examined. For the requirements that are partly or not fulfilled, some possible solutions are proposed.

5.1 Lecture delivery phase requirements

The main phase of a teleteaching activity is the lecture delivery phase, in which a remote instructor lectures. The major requirement ***(R1)*** is that the remote students should have high quality audio-visual communication with the instructor and his/her presentation data.

This communication requirement is fulfilled since present H.323 MCU products support the chair control mechanism which controls and enables the constant selection of Iv(Wi), Ia(Wi), and Dp streams for switching to all remote workspaces. The participating H.323 terminals should support chair control. Automatic video switching mode can be used, but in this mode interference from another participant is possible.

To meet the high quality requirement, G.711 audio coding (64 kbps) should be used, and the video codec should support 30 fps CIF at least at 768 kbps.

As far as the display of Dp(Wi) and Iv(Wi) is concerned, it is often required that each of them be displayed on a separate display device ***(R2).*** In the case that a workspace is equipped with a single display device, Dp should be displayed in full screen while Iv(Wi) in PIP window, since it is considered that the data presentation has higher priority than the instructor's video, from an educational perspective. Displaying of video and data presentation in dual monitors is presently supported by many H.323 terminals.

A third requirement ***(R3)*** often met is that instructors should be free to move around as in a conventional classroom. To enable this, the instructor's camera should be a Pan-Tilt-Zoom camera in order to follow him/her as he/she moves. A wireless microphone should also be used by the instructor to facilitate his/her movement. The instructor's camera could support autotracking as well. In this case, no cameraman or technician is required for the camera control.

5.2 Instructor's interaction with students

Instructors interact with students when students ask questions or participate in a conversation with the instructor. In this state of a teleteaching interaction, all remote participants should be able to have audio and visual (less priority in comparison to audio) contact with the person who participates in the dialogue with the instructor ***(R4)***. This requirement is being fulfilled in two ways:

1. By operating the MCU in the automatic video switching mode. Because of the audio mixing, the participants listen to the dialogue immediately. The visual contact is realised after a person's continuous speaking for two seconds. This time is required in order for the MCU to declare who the speaker is, and switch speaker’s video stream to the remote participants.
2. By operating the MCU in the chair control mode. In this mode the chairperson grants the floor to the person's workspace terminal. The audio and video stream is immediately switched to all participants.

A major problem arises when a student/participant wishes to ask for the floor ***(R5).*** Floor control and request is enabled only in the chair control mode. Only the operator of the H.323 room terminal is able to request the floor. Thus, the operator has to request the floor on behalf of a student but this will discourage the students' active participation in the lecture.

A complementary floor control system could also be used. This system should co-operate with both the H.323 terminal and the camera system. The floor control system should permit the student to submit and cancel floor requests to the terminal. The chairperson should be able to view the request and decide to whom to grant the floor. When the floor is granted to a student, the camera system should tilt and zoom to that student.

If chair control is not supported by the terminals then social rules, e.g., hand raising, have to be employed.

5.3 Open conversation

In certain circumstances, the support of open conversation among all the participants is required ***(R6).*** The audio mixing function permits the open audio conversation of the participants. The MCU selects for switching the video stream of the participant who speaks for more than two seconds.

This requirement is fulfilled by using the microphone system in “momentary on” mode. In this mode each speaker should touch the mic’s button while speaking. When the button is released the microphone is muted and the audio stream does not contain audio samples. Thus the MCU recognises the current speaker and selects its video.

5.4 Interference of an external person.

The teleteaching environment should allow for the participation and intervention of an external person in the teaching process ***(R7).*** This requirement can also be fulfilled. The person should use either an H.323 or H.320 terminal and an H.323/H.320 Gateway, and participate just as another workspace.

5.5 Projection of the remote workspaces' signals

A requirement ***(R8)*** often set by pedagogical experts is the following one: the students should be able to attend the reactions and gestures of the remote classroom students in order to avoid sensory deprivation [9].
This requirement could be fulfilled by displaying the remote workspaces' video in one of the permanent, periodic or instant displaying modes. The permanent display requires the installation of a display device in each workspace, dedicated to permanent displaying of the remote workspaces. The periodic and instant display of remote workspaces could be done in a single monitor. In all modes, either one workspace each time, or a panorama of all or some of the workspaces could be displayed.

Today, no simultaneous view of the instructor's video and anyone's remote workspace is feasible either at the main or dual monitor. This is because neither the video mixing function is supported by present H.323 MCU products nor receiving of multiple video streams capability is supported by present H.323 terminals. For these reasons, only the periodic or instant display of remote students can be implemented today with current H.323 technology. The chair control capability is used for this purpose. The chairperson, periodically or instantly, selects a specific workspace's video stream. For the implementation of permanent display, other technology has to be used. For example the Mbone tools, such as vic [12].

In the future, H.323 products could support this teleteaching requirement by enhancing the MCUs and terminals. There are two alternatives, depending on whether a single or multiple workspaces' video(s) should be displayed simultaneously:

A. Multiple workspaces

1. The H.323 MCUs have to *concurrently* support

– video switching of the current speaker's video (e.g., instructor's);
– video mixing of the receiced streams and switching of the produced stream to the H.323 terminals.

2. The H.323 terminals have to be capable of

– receiving two video streams (speaker’s and workspaces’);
– displaying each video stream on a separate monitor.

B. Single workspace

1. The H.323 MCUs have to *concurrently* support selection and switching of two video streams: the speaker's and a workspace's.
2. The H.323 terminals have to be capable of

– receiving the two video streams;
– displaying each video stream on separate monitor.

5.6 Projection of the remote classroom on the instructor's small size monitor(s).

It is sometimes desirable that the instructor be able to attend permanently the students located at remote workspaces by displaying each workspace's video signal to a small size monitor ***(R9).*** This requirement is similar to the previous one.

5.7 Lecture broadcasting

A lecture should be possible to be broadcast to a large number of students who should be able to receive the lecture but not interact (passive students) ***(R10).*** Presently, this requirement is not fulfilled. A limited number of users - each equipped with H.323 terminal - are permitted to be connected to an MCU. Lecture broadcasting service could be provided by MCUs capable of multicasting the audio, video, and data streams. The passive users should be able to use an H.323 terminal capable of joining the multicast session. If the MCU transcodes the video by employing a low quality coding scheme (e.g., Sub-QCIF), then the H.323 terminal could be only a software entity.

An alternative is to implement a broadcasting entity, which participates in an H.323 multipoint centralized conference as an H.323 terminal, and multicasts the receiving streams to the passive participants. The participants may use either the MBONE applications [13] or a commercial product. There exists an H.323 product that follows this approach but communicates only with a specific commercial product.

5.8 Session recording

A teleteaching session should be recorded to be available for latter asynchronous access by the student(s) or the instructor ***(R11).*** This is not supported by the H.323 recommendation. A complementary system should be used, possibly in conjunction with the broadcasting system.

5.9 Session management

A teleteaching environment should provide a way for session announcement, scheduling and management ***(R12).*** The active participants should be invited by a session control entity, while passive participants should be able to join the session. The requirement for an authentication mechanism is optional. Statistics should be kept in a database for billing and accounting.

Presently, the administrator of an H.323 MCU is able to schedule a conference by inserting the IP addresses of the potential participants in the MCU. H.323 terminals whose IP address is not included in the IP addresses list can not join the conference. No announcement mechanism is provided by the H.323 recommendations for announcing the session. Usually, e-mail communication is used for announcing the schedule to the potential participants.

In present multipoint centralized conferences, the participants have to call the MCU. No invitation procedure from the MCU to the participants is supported. Statistics are not kept for billing, accounting or statistics reasons.

6 Conclusions

Present H.323 products are mature for deploying point to point and multipoint videoconference services and collaboration over IP networks in business environments. Regarding to teleteaching services is concerned, present H.323 products can fulfil only partly the teleteaching requirements.

H.323 products need to be enhanced in order to support high quality teleteaching services. More precisely H.323 MCUs should be enhanced to:

- support the multicast communication mode for stream distribution;
- control the terminals so that they not transmit their video stream when they are not selected by the H.323 MCU for switching (currently not always available);
- support the video mixing function;
- support the video transcoding function;
- allow for video switching of multiple streams (at least two video streams);
- enable concurrent video mixing and switching .

The H.323 terminals should be enhanced to:

- support the multicast communication mode;
- be capable of receiving and decoding at least two video streams, and displaying them in dual monitor;
- support chair control (desktop terminals).

All the above capabilities are defined by the H.323 Recommendation. An important enhancement, which is not defined by the H.323 recommendation, is the definition of the procedures that will permit the co-operation of room systems with microphones and camera systems. A microphone system should provide capabilities for submitting and cancelling floor request. The camera system should be able to tilt and zoom to the person who is granted the floor.

Assuming that H.323 products will support these enhancements, the question that remains to be answered is how cost-effective and flexible is the deployment of teleteaching services based on H.323 audio visual systems in comparison with alternative technologies and architectures. Such technologies are the MBONE tools and scalable and composable architectures for distributed multimedia collaboration such as the MASH architecture [15].

References

[1] ITU-T Rec. H.323, "Visual Telephone Systems and Terminal Equipment for Local Area Networks which Provide a Non-Guaranteed Quality of Service", 1996.

[2] ITU-T Rec. H. 261, "Video CODEC for Audiovisual Services at p x 64 kbit/s", 1993.s

[3] ITU-T Rec. H 263, " Video CODEC for Narrow Telecommunications Channels at < 64 kbit/s", 1995.

[4] ITU-T Rec. H 320, " Narrowband ISDN Visual Telephone Systems and Terminal Equipment," 1995

[5] ITU-T Rec. T.120 “Data protocols for multimedia conferencing”, 1996

[6] ITU-T Rec. H.243, "Procedures for establishing communication between three or more audiovisual terminals using digital channels up to 1920 kbit/s", 1996.

[7] ITU-T Rec. H.245, "Control protocol for multimedia communication", 1997.

[8] ITU-T Rec. H.281, "A far end camera control protocol for videoconferences using h.224", 1995.

[9] J. Bourdeau, M. Quallet, R. Gauthier. "Interactivity in Videoconference based Telepresentations." ,ED-MEDIA &ED-TELECOM 98.

[10] Robert Grebner, "Use of instructional material in universal teleteaching environments", Computer Networks and ISDN Systems 29 (1997) 1787-1797. Proceedings of the TERENA Networking Conference '97.

[11] Arno Klein, "Teleteaching scenarios for high bandwidth networks", Computer Networks and ISDN Systems 30 (1998) 1707-1716. Proceedings of the TERENA Networking Conference '98.

[12] S. McCanne, V. Jacobson, 'vic: A Flexible Framework for Packet Video", ACM Multimedia - Nov. 1995.

[13] S, McCanne, "Scalable Multimedia Communication Using IP Multicast and Lightweight Sessions", IEEE Internet Computing, March-April 1999.

[14] G.A Thom, "H.323: The Multimedia Communications Standard for Local Area Networks" IEEE Communication Magazine, pp 52-56, Dec. 1996

[15] http://mash.cs.berkeley.edu/mash/index.html

[16] http://www.gunet.gr/pilots/

[17] http://www.cnl.di.uoa.gr/Projects/tuatm.html

[18] T.Walter, H.Haenni, "Telepoly - A Teleteaching Scenarion supported by high Speed Networks", http://www.tik.ee.ethz.ch/~walter/Telepoly/Telepoly.html
[19] http://www.vcon.com
[20] http:// www.picturetel.com
[21] http:// www.intel.com/proshare

Vitae

Pantelis Balaouras received the Diploma in Computer Engineering and Informatics from the University of Patras, Greece, in 1992. From 1992 to 1995, he was a Research Scientist at the Computer Technology Institute, Patras, Greece. Since 1996, he has been working as a Research Engineer at the Communication Networks Laboratory (CNL) of the University of Athens. He is a PhD candidate in the Department of Informatics, University of Athens, and his research work focuses on the design, development and deployment of advanced multimedia services. He has been involved in a number of European and national networking projects. Also, he is a technical consultant of the Greek Universities Network (GUnet) and a member of the technical committee of the Greek National Research Network (GRnet).

Ioannis Stavrakakis received his Diploma in Electrical Engineering, Aristotelian University of Thessaloniki, (Greece), 1983; Ph.D. in EE, University of Virginia (USA), 1988; assistant professor in CSEE, University of Vermont (USA), 1988-1994; associate professor of ECE, Northeastern University, Boston (USA), 1994-1999; currently, associate professor of Informatics, University of Athens (Greece). His teaching and research interests are focused on resource allocation protocols and traffic management for communication networks. His past research has been published in over 90 scientific journals and conference proceedings. His research has been funded by NSF, DARPA, GTE, BBN and Motorola (USA) as well as Greek and European Union Funding agencies. He has served in NSF research proposal review panels and involved in the organization of numerous conferences sponsored by IEEE, ACM, ITC and IFIP societies. He is a senior member of IEEE, a member of the IEEE Technical Committee on Computer Communications (TCCC) and a member of IFIP WG6.3. He has served as an elected officer for TCCC, as a co-organizer of the 1996 International Teletraffic Congress (ITC) Mini-Seminar, on ``Performance Modeling and Design of Wireless/PCS Networks, as the organizer of the 1999 IFIP WG6.3 workshop and a technical program co-chair for the IFIP Networking'2000 conference. He is an associate editor of the ACM/Baltzer Wireless Networks Journal and the Computer Networks Journal.

Lazaros Merakos received the Diploma in electrical and mechanical engineering from the National Technical University of Athens, Greece, in 1978, and the M.S. and Ph.D. degrees in electrical engineering from the State University of New York, Buffalo, in 1981 and 1984, respectively. From 1983 to 1986, he was on the faculty of Electrical Engineering and Computer Science at the University of Connecticut, Storrs. From 1986 to 1994 he was on the faculty of the Electrical and Computer Engineering Department at Northeastern University, Boston, MA. During the period 1993-1994 he served as Director of the Communications and Digital Processing Research Center at Northeastern University. During the summers of 1990 and 1991, he was a Visiting Scientist at the IBM T. J. Watson Research Center, Yorktown Heights, NY. In 1994, he joined the faculty of the University of Athens, Athens, Greece, where he is presently a Professor in the Department of Informatics, and Director of the Communication Networks Laboratory (CNL) and the Networks Management Center. His research interests are in the design and performance analysis of broadband networks, mobile networks, and advanced video communication technologies and services. He is the author of over 110 papers in the above areas. From 1995 to 1999 he lead the R&D activities of the CNL in the area of mobile communications, in the framework of the Advanced Communication Technologies & Services (ACTS) programme funded by the European Union (projects RAINBOW and Magic WAND). Since 1996, he is responsible for the deployment of the Greek Universities network (GUnet), the national academic network offering high speed connections and advanced networking services to the 18 Universities and 14 Technological Education Institutes (TEI) of Greece. In 1994, he received the Guanella Award for the Best Paper presented at the International Zurich Seminar on Mobile Communications.