Video Conferencing in Distance Education

Sami Andberg

Helsinki 12.05.2008

Pro Gradu Thesis

UNIVERSITY OF HELSINKI

Department of Computer Science

HELSINGIN YLIOPISTO – HELSINGFORS UNIVERSITET

Tiedekunta - Fakultet	Laitos – Institution
Matemaattis-luonnontieteellinen tiedekunta	Tietojenkäsittelytieteen laitos

Tekijä – Författare
Sami Andberg

Työn nimi – Arbetets title
Video Conferencing in Distance Education

Oppiaine – Läroämne
Tietojenkäsittelytiede

Työn laji – Arbetets art	Aika – Datum	Sivumäärä – Sidoantal
Pro gradu –tutkielma	12.5.2008	45 + 5

Tiivistelmä – Referat

Tässä työssä käsitellään videoneuvottelua eri näkökulmista; videoneuvottelun tekniikkaa, laitteistoa ja sen käyttöä, videoneuvotteluun liittyvää tutkimusta ja videoneuvottelutekniikan valintaa.

Työssä esitellään myös videoneuvottelutyökalujen valintaan organisaatiotasolla sopiva SECTIONS-viitekehys ja sen pohjalta työssä luotu uusi opettajakeskeinen IDEAS malli.

CCS-luokitus: H.4.3, K.3.1

Avainsanat – Nyckelord
Videoneuvottelu, etäkokous, verkkoluento, etäopiskelu

Säilytyspaikka – Förvaringställe
Kumpulan tiedekirjasto, sarjanumero C-

Muita tietoja – Övriga uppgifter

UNIVERSITY OF HELSINKI

Faculty	Department
Faculty of Science	Department of Computer Science

Author
Sami Andberg

Title
Video Conferencing in Distance Education

Study subject
Computer Science

Study Category	Date	Pages
Pro gradu –thesis	12.5.2008	45 + 5

Abstract

This thesis focuses on video conferencing from different perspectives; possibilities of the technology, using of the technology and implications of the technology.

The SECTIONS framework, for selecting video conference tools, is introduced, and a new, teacher-centric IDEAS model is developed to assist in the selection process.

CCS-classification: H.4.3, K.3.1

Keywords
Video conference, web conference, distance education

Säilytyspaikka – Förvaringställe
Kumpula scientific library , serial number C-

Other information

CORRECTION SLIP 7/29/2010

NOTICE!

There is a wrong reference on the thesis, page 1, reference [NYT24]. The correct article is [NYT27], details below:

To replace:

NYT24 Pictures By Wire Sent With Success for the First Time. New York Times, May 20, 1924

Replace with:

NYT27 Far-off Speakers Seen As Well As Heard Here In A Test Of Television. New York Times, Apr 8, 1927

(Article available online from the New York Times archive for a fee).

The author apologizes for the wrong reference ending up into the official thesis version.

This page has been added to the thesis on July 29th 2010.

The actual thesis document (following pages) is unaltered in the original form – including the wrong reference.

Contents

Glossary

3G	3rd generation mobile phone technologies
Avatar	Computer-generated presentation of a person in a virtual environment
Codec	Coder-Decoder unit or software
G.xxx	A protocol belonging to the ITU G-family
H.xxx	A protocol belonging to the ITU H-family
HD	High definition
HTML	Hypertext Markup Language
ISDN	Integrated Services Digital Network
ISO	International Organisation for Standardisation
ITU	International Telecommunications Union
MCU	Multipoint Control Unit
NAT	Network Address Translation
PSTN	Public Switched Telephone Network
Settop box	A standalone videoconferencing unit with an integrated camera
SIP	Session Initiation Protocol
UMTS	Universal Mobile Telephone System, one of many 3G standards
VCR	Video Cassette Recorder
VOD	Video on Demand
VR	Virtual Reality

1 Introduction

Video conferencing *per se* is not a recent invention – it was already envisioned by the inventor of the telephone, Alexander Graham Bell, who was quoted in a 1924 article in the New York Times as saying "the day would come when the man at the telephone would be able to see the distant person to whom he was speaking" [NYT24]. Experimental videophone installations and successful calls have been made since the late 1920's. The first commercial video conferencing equipment, AT&T's Picturephone Mod I, was introduced with wide press coverage in 1964 [ScW04].

Despite of the seemingly long history, video conferencing has only recently became increasingly popular and disperse in the wake of faster and cheaper connections and better technologies. Modern standalone video conferencing units provide advanced video and audio quality due to more efficient compression and can function over normal broadband internet connections. Growing processing power and cheaper accessories, such as webcams, have also made it possible to participate into a video conference using dedicated software on a normal personal computer without any expensive special hardware. In short, the quality has gone up and the price has come down and the number of different technologies has multiplied. Thus, in the last decade, there has been a noticeable growth in both the number of people and institutions using video conferencing and the software and hardware solutions available.

The aim of this thesis is to give a teacher a broad view on different aspects related to the video conferencing techniques and their importance in choosing and operating a video conferencing solution. In this thesis, the term *video conferencing* is recognised in a broad sense, encompassing, standalone video conferencing units and related infrastructure equipment, software based desktop conferencing systems,

web conferencing systems with video and audio functionality, and video calls from mobile phones.

In Chapter 2, the basics of a video conference are introduced and a common video conference setup with some possible accessories are reviewed. Secondly, the management of a video conferencing session is focused on. Thirdly, the most common extensions for video conferencing are presented. In Chapter 3 the technology including standards and protocols closely related to video conferencing are described in detail.

In Chapter 4, the SECTIONS framework for selecting video conferencing tools is introduced and a new IDEAS model is developed. Secondly, other aspects and today's best practises related to choosing video conferencing system are discussed. Finally, in Chapter 5, the conclusions are presented.

In Appendix 1, the characteristics of general categories of videoconferencing are outlined. In Appendix 2, detailed technical information on a number of popular video conferencing solutions are presented.

2 Utilising video conference

In this chapter, the basics of video conferencing and the most common video conferencing equipment are explained and the typical conference types introduced. Secondly, attention is paid to the management of a video conference and finally some possible extensions for a video conferencing session are reviewed.

2.1 Introduction to video conferencing

A video conference, sometimes referred to as a videoteleconference or VTC, is a means of communication where the sound (audio) is accompanied by a live picture (video). A video conference can be relayed over the Internet or it can utilise ISDN telephone lines, satellite links or wireless networks, even cell phones.

A video conference can be between two *sites*, i.e. locations which are connected to each other via the video conference, or the conference can connect multiple locations. The communication can take place in a special video conferencing studio, or on a normal home computer equipped with a webcam – even a videocall on a modern 3rd generation mobile phone falls into this scope. As can be seen, the term encompasses a wide range of possibilities.

While video conference can never fully substitute for a real experience, research findings indicate, that, for example, a group of people playing the 'mafia' game over a video conferencing connection experienced the same levels of fun and satisfaction than the other group playing the game face-to-face [BHN07]. It should be noted, though, that the group using video conferencing had some complaints about not seeing all of the facial expressions correctly or not hearing everyone properly – known limitations of many video conferencing solutions, but the situation has recently been getting better with the increases in audio and video quality in the modern video conferencing systems.

2.2 Video conferencing equipment

Basic equipment needed for a video conference session include a camera, microphone, a video conferencing unit or a video conferencing software, and a display. In this subsection, the technical setup for video conferencing is laid out and attention paid to specific aspects contributing to a successful conferencing environment.

Camera

Regardless whether the video conference is set up in a dedicated conference room, large auditorium or if a person is participating from a personal computer, there are certain issues that should be taken into account. One issue is the location of the camera. It has been noted, that while realistic eye-contact might be impossible to accomplish with many video conference setups, a reasonable camera position can help participants to spontaneously learn when the remote participant is looking 'straight' at their image [GrM03]. This kind of gaze awareness is important for taking turns and helps in the overall social aspect of the video conference situation.

The best place for camera is directly above and as close to the image of the remote site as possible. This provides the best gaze direction awareness [GrM03]. There is evidence that in case the actual picture of remote site is large compared to the distance between the image and participants, the forthcoming problems with gaze direction can have an impact on the experience of trust and the social atmosphere between participants [NgC07].

Another important issue is the quality and functionality of the camera. Better quality cameras are able to provide a sharper, more colourful image, with less visual noise. Some cameras also offer a wireless remote control to pan, tilt and zoom the image – a functionality that is recommendable in a group setting. Thus,

cameras can have a clear effect on the overall visual appearance of a video conference solution. This holds true even to web conferencing or messenger environment, as webcams also come with different sizes, qualities and functionalities.

Sound

There exists a common agreement between many scholars, that audio quality is the most important aspect in video conferencing [TaI92, Alb06, p.7]. If audio quality is degraded the communication is hard to understand, which usually makes the whole conference quite meaningless. Studies also suggest that there might be a critical level for the audio delay between sites – Tang et al. report that delays longer than 0.5 sec. tend to induce collisions in speech turns and cause problems reading body language, which in turn greatly discourages spontaneous interaction [TaI92], while Dudman refers to Ahuja that the delay should not exceed 0.15 sec or the natural flow of communication starts to suffer [Dud06]. However, some studies using proprietary video-enabled learning environments have reported delays experienced between participants to be up to 60 seconds, which make normal interaction relatively impossible [HeH05].

Fischer and Tenbrink address the problem of turn-taking in video conference mediated communication. Evaluating the use of a basic ISDN-based video conferencing system with compressed picture, half-duplex audio (one way at any time) and almost a 0.5 sec. delay, the researchers found that the participants initially encountered huge problems in changing the active speaker spontaneously [FiT02]. In their observation the participants soon figured out a visual cue – to raise a hand to signal a wish for a turn. This is an expression of a visual cue that seems to come quite naturally in video conferencing and has been encountered in a number of instances.

In reality, it is often the case that the audio has to be picked up by a single tabletop microphone originally included in the video conference package. In this case the positioning of the microphone is of utmost importance. The microphone should be close to the main presenter, away from any disturbances such as projector or laptop fans, presenter's papers, etc. Still, it probably would not cover the far corners of the room (depending on room size). Thus the use of multiple microphones when possible is recommendable.

The minimum requirement to achieve reasonable sound quality is to have the video conferencing room free from external distractions. Also the microphone(s) should be of some quality. In an optimum scenario, there would be a single (possibly wireless) microphone for the presenter and group microphones for the participants. Alberta Education noted this to be the best solution in their review of the video conferencing facilities at different school districts [Alb06, p.7]. Utilising an audio mixer with noise gate and echo cancellation helps the quality – this functionality is nowadays inbuilt to many standalone video conferencing endpoints and even in some web conferencing solutions.

To extend the range of single microphones, it is possible to mount them to the roof of the conference room. Due to the usually automatic gain applied to the microphones they will pick up sound from a larger area – but as they do so they will also be more sensitive to pick up any unwanted sounds [Alb06, p.80], so the benefit is always a compromise which depends on the space and the usage. Wireless microphones should be used when applicable – they have proven to be especially useful to pick up for example the voice of a presenter strolling around on a stage.

Lighting and background

Lighting is an easy way to improve picture quality. If the room is not specially built or equipped for video conferencing, it is probable that there are not enough lights to

provide the optimum quality for the video conference cameras. The result is a flickering visual noise seen especially when the cameras are zoomed in. Another result is a lack of colour saturation. Thus proper lightning is an easy way to improve video quality.

It should be noted though, that often the participants are not keen to have extensive lighting directed at them, possibly making it hard for them to see the screen. Bright ambient lightning usually works fine without disturbing the participants. Attention should also be paid to avoid mixing incandescence lights with fluorescent lighting which would result in a distorted colour reproduction.

One relatively easy and cheap way to reduce the echo in the room and to enhance the video encoding efficiency is to organise an ambient, static background for the image. This can be done for example by adding thick dark blue curtains next to the walls – the thickness helps to absorb background noise thereby enhancing the clarity of the sound and the big static blue area is less consuming for video encoding algorithms resulting in a more stable image with more details for the actual persons especially when utilising low bandwidth connections.

2.3 Related equipment

The basic video conferencing setup can be accompanied by other equipment such as a document camera or a DVD player thus extending the possibilities and versatility of the system. In this subsection other equipment, related to video conferencing, are presented.

Document camera

A document camera, or a visualiser, can be attached to some video conferencing units. This is an useful addition as it enables the speaker to show material and objects in a convenient way, either by switching between the presenter and the material or sending both video feeds simultaneously.

The material to show from a document camera can be printouts as well photographs, physical objects such as rock or scissors, or even a mobile phone screen etc. As the document camera usually has a zoom function, it is possible to show even quite small objects and still get vivid details. It is also possible to write on a paper on the document camera and thus make notes or sketches that are viewed live from all sites.

In distance education, document cameras enable students to show their completed assignments to the teacher just by placing them on the document camera. In fact, document cameras connected to data projectors are becoming increasingly popular also in regular classrooms, replacing old visual overhead projectors - just another sign of their versatility.

PC

One of the most popular presentation tools these days is the personal computer. Attaching a PC to a video conferencing unit using a VGA cable – or utilising a video conferencing-software on a PC – usually enables the presenters to share their desktop image – that is to be able to show slides, web pages, demonstrations, etc.

However, the limitations vary hugely based on the kind of equipment and protocols used – some video conference systems are able to share the desktop image in unaltered full resolution but with a reduced frame rate. Newer units can even do this in addition to showing the video of the presenter. At the other end of the spectrum, older equipment can only show either the presenter or the desktop image in a greatly reduced resolution (resulting to, for example, the normal size text becoming unreadable and thus making it useless to show any web pages or documents that are not specifically prepared for the medium).

DVD, VCR

It is also possible to present pre-recorded videos from a VCR or a DVD player through a video conference. However, according to many video conferencing practitioners, the aforementioned use of pre-recorded material should be utilised only for a small number of short video clips. If a longer video playback is needed, a better way is to send the actual media (for example the video file over the internet or on a DVD) to the remote site in advance to be presented locally which ensures better quality and less potential problems.

2.4 Video conference session types

Generally speaking, all video conferencing events can be divided into two categories depending on the number of locations participating in a video conference session: one type being a session between two locations, called point-to-point connections, and the other a session with more participating locations, thus called a multipoint session. The next subsections map out the differences and characteristics of these two scenarios. Finally, attention is paid to the specific implications of attending method – going to specific video conferencing studios vs. utilising a personal computer.

Point-to-point 'video call' versus multipoint video conference

The most common usage of video conference is the so called point-to-point connection. As the term implies, it is a straight connection between two video conferencing units (or between two computers with video conferencing software, or any combination of those). The audio and video feeds are simultaneously sent from each site to the other, so that both participants see and hear each other. Depending on the characteristics of the actual protocol used, the data could be relayed through a server, a gateway or a gatekeeper proxy, but basically the data flow is bidirectional.

A multipoint video conference differs from the point-to-point connection by encompassing more participating locations – anything from two sites up to hundreds. The actual implementation of a multipoint conference varies hugely depending of the technology and protocols used. For example, in the ITU-T video conferencing family (H.320, H.323, etc.) multipoint conferencing is usually arranged by having all participants contacting a central server, a MCU (Multipoint Control Unit), which handles the selecting of images and mixing of audio and sends the resulting audio and video feed back to every participant – in a way the connections between an endpoint and the MCU can be seen resembling very much a 'point-to-point' connection with single audiovisual feed going to both directions.

On the other hand, endpoints in a SIP based conference are often able to function without a dedicated multipoint control unit, as a multipoint call is produced by sending individual audio and video streams from all participants to each other. This in turn results in a linear growth of usage in the bandwidth for every extra participant (unless the meeting is on a multicast enabled network).

Desktop conferencing

Although desktop conferencing can be point-to-point (such as a messenger call) or multipoint (such as attending a video enabled web conference meeting), it is important to notice some fundamental differences between attending a video meeting from a personal workstation compared to going to a dedicated video conferencing room. As Tang and Isaacs pointed out, by using personal workstation for video conferencing the participants are in their familiar surroundings and have easy access to their own papers, documents and environment [TaI92]. On the other hand, they are more easily distracted by phone calls and other people than when attending in video conference rooms which tend to be more isolated and thus more distraction free.

Due to the cheaper price and easy installation, desktop video conferencing can be utilised almost everywhere where is a computer with a webcam and a broadband internet connection. In a study from 2004, Wang describes using Microsoft NetMeeting as a real time conferencing tool between language learners and the tutor [Wan04]. In this case, most of the students utilised a relatively slow dial-up modem connection to the internet and still were able to get somewhat satisfactory results. The study concludes that broadband is generally needed to provide better quality and reduced delay.

There is also a difference in the accessibility between desktop systems and video conferencing rooms. Dedicated video conferencing facilities can often accommodate a larger number of people, but the rooms usually have to be booked in advance, whereas in desktop conferencing each participant usually attends to the meeting using their own computer and thus conferencing can be initiated much more spontaneously [TaI92].

The combination of video conference rooms with desktop conferencing has also it's advantages – in some school districts in Alberta, Canada the video conference rooms at schools are also equipped with a desktop conferencing unit which could be used simultaneously with the actual set-top video conference connection [Alb06, p.26]. This enables the possibility for one-on-one instruction for students or remote student collaboration simultaneously while the 'public' conference was still utilising the normal video conference channel, thus creating a possibility to have 'private' conversations between some people at different sites.

2.5 Managing video conferencing

During the course of a video conference session, a number of things can affect on the outcome on the event – the microphone might pick up unwanted sounds or the presenter might walk out of the camera image. In this subsection some simple approaches to enhance the quality of the conferencing experience are described.

Microphone muting

The basic rule of the thumb could be described as following: if you are not the speaker at any given moment, mute your microphone. This is especially important in a multipoint video conference, as if the microphone is open, any sound from the participating site (talking to a co-worker, coughing, telephone ringing or someone knocking on a door) will be heard by all members of the conference and this kind of background noise will make it increasingly hard for anyone to hear the person speaking at the moment. To make things worse, some multipoint control units use automatic algorithms to select which speaker is displayed on the screen, and usually this selection is done on the basis of whichever site is making the loudest noise – which could be for example someone tapping the table where a microphone is situated on. Thus, closing microphones, when not speaking, is essential.

It should be noted, that some video conferencing solutions automatically treat microphones as muted unless the person on the site in question is holding down a 'talk' button. This is one way to force the muting by making it the default situation. However, in most solutions this is not the case and the individual users should take care of their local microphones.

Camera panning

As video conferencing is an audiovisual medium, one critical aspect of a successful conferencing session is what is shown on the video channel. This is especially important for the site which is active at any given moment; it is in a human nature that we would prefer to see the speaker. Thus, if the conferencing equipment is equipped with movable camera (by remote control or manually) there should be someone trusted upon pointing the camera on whoever is speaking at the local site (or zooming out to show the whole of participants when just following). On the other hand, if the camera position and angle is fixed, then the arrangement of participants should be such that everyone is in the picture all the time (often used in small groups) or that the one speaking for the moment would walk closer to the

camera and microphone to state his or her opinion or question (often used in big groups). It has also been suggested that a lecturer preparing for a video conference session could use the storyboarding technique to plan the session in advance [Sar98].

Good etiquette

Vide.Net video conferencing initiative has published a video conferencing cookbook that is a recommendable reference for anyone interested in managing a video conferencing session [Vid05]. The cookbook includes a best practises and etiquette chapter. Many national video conferencing organisations have also produced similar tutorials in various languages.

According to ViDe.Net, the best practises and video conferencing etiquette include [Vid05]: testing in advance to make sure everything is ready well before the actual session starts; leaving well enough alone – if it is not broken, do not fix it; paying attention (in video conference it's easy to start reading emails, etc.); keeping microphones muted and talking on turn; minimising distractions – keeping local specialities out of sight for the conference; and remembering, that a video meeting IS a meeting – preparing and behaving accordingly.

Many of the most important issues related to a successful conference experience have been mentioned above (muting microphone when not speaking, panning camera to show participants, having adequate lighting and sound quality, etc.) but the etiquette expands this list for example by stating that if there is a coffee and bagels arranged for the participants in one site, they should be available on other sites as well – or they should be consumed off the camera. This kind of little advices can have a strong positive impact on to the overall experience of the conference.

As an anecdote it can be said, that on many occasions the operating task of microphone muting and camera panning (and zooming) has been successfully given

to a member of the local audience, for example, a student or a technical support person. This designated operator can then take control of the aforementioned tasks with a relatively brief Training and thus relieve the presenter to focus on the actual subject matter together with the social and pedagogical aspects of the session.

2.6 Extensions for video conferencing

With many video conferencing solutions it is possible to add extra functionality either by purchasing specialised equipment or software licences. These functionalities include, for example, the possibilities to stream the conferences live to the internet or to record them for later viewing.

Live streaming

Nowadays, it can be relatively easy to expand the audience of a video conferencing session by allowing non-participants to view the conference using media player software. This expands the number of people able to follow the conference as the bandwidth and software limitations are not usually an issue for one-way streaming as a much lower quality and bandwidth can be used. However, due to the one directional nature of streaming, the participants following the stream are not able to interact directly within the conference but an alternative feedback route has to be used. According to the experiences of the author, it is quite common to utilise dedicated internet chat rooms, mobile phones or discussion boards for interaction from viewers.

Streaming a video conference live for internet based audience is well suited, for example, for broadcasting lectures or panels, where the average participant interaction with the actual event is scarce - in these cases, the internet streaming might even be the top priority, and the video conference connection could serve as one feedback channel among others for participants who have the equipment needed. There are positive experiences of using conference streaming even in team

meetings; for example, if some participants are not able to attend the meeting in any of the dedicated conference rooms, they could follow the stream on almost any pc or mobile device, and participate using a phone or chat to express their views when needed.

There are many ways to come up with the streaming of video conferences, some of them easier than others. The simplest way is to use a video conference software, codec or MCU that has a built in streaming capability – in this case the stream is usually enabled just by switching "streaming on" in the menu. On the other hand, it is possible to connect the audio and video outputs of a video conferencing unit to a computer and use any streaming software to broadcast the event.

Recording of conferences and video-on-demand

Video conferences can be recorded in many different ways. In the 1990's the most popular was to record the session to a VCR. Nowadays, the easiest method depends on the equipment or software used, but in some systems it is as easy a clicking a recording button on the interface and the conference is then stored to a local media file.

The recordings from a video conference can be used as a standalone media file, for example, to be uploaded to a web based learning environment or to a web page, or the file can be used with a media server to provide video-on-demand streaming service. Video-on-demand, VOD, is basically almost the same than live video streaming, with the exception that the source is not a live video feed but a media file stored on a media server. The biggest difference between users accessing a VOD stream, or downloading a media file, is that VOD streams usually start playing the video a few seconds after starting, as when downloading a file one usually has to wait until the whole file has arrived before it is possible to be accessed.

Recording a video conference and providing the file as standalone or on video-on-demand can extend the usability of the video conference session by allowing

subsequent access to the material. For example, a video conferencing lecture or a panel can be recorded into a file to be put onto a media server and thus be accessible on the internet for a desired group or everyone to access. It should be noted, though, that in this way the original interactivity is lost.

In many newer video conference recording systems it is possible to record both the audio/video feed of the conference and the content feed simultaneously. Thus the person watching the recording later on sees two video windows, one with the presenter and other with the content: slides, computer screen, document camera image or whatever was used in the conference.

Gibbs and Larson describe the evaluation and results of student feedback on partly live vs. fully VOD (video-on-demand) courses [GiL07]. Although they did not have the actual video conferencing aspect (no real time remote participants), the resulting lecture is similar enough to VOD recordings of a video conference for their results to be comparable. Their findings state, that when the system was utilised on a VOD only course (students had to access all lectures and examples from the media server or local files) the students were much more critical for any occurring problems. On the other hand, the students who had access both to live lectures and the VOD material were not so concerned with the slight glitches with the recorded material.

Berner and Adams researched the difference of content plus audio and content plus audio plus video presentations of a pre-recorded lecture, and came to conclusion that there were no significant differences in learning or student satisfaction between the two groups. However, they point out that the situation might be different depending on the subject area, the personality and style of lecturer, cognitive styles of the learners etc. [BeA04]. Although the material used by Berner and Adams was not acquired through a video conference, the resulting recording can be assessed similar enough to assume that their findings might be applicable also to a video conference -originated recording.

There is also evidence that video image is especially important when the communication is in foreign language. In this case the visual cues of other speakers can help to convey the meaning of the message, and also provide visual feedback to the speaker that his message is understood [Wan04].

Future possibilities in video conferencing

One interesting aspect, considering the future of communication over distance, is the development of virtual reality environments. The VR environments, such as Second Life [Sec08], offer new educational possibilities compared to a videoconferencing setup, for example, by making it possible for participants to mingle as *avatars* and hold 'hallway discussions' before or after the official program – a possibility which has been referred as an important part of human communication which is normally a bit problematic in regular videoconferencing. Sallnäs has noted that adding video communication to a virtual environment made participants to spend more time discussing a task and have more extensive dialogues [Sal02]. In the educational setting this makes virtual environments especially well suited for dynamic group collaboration and peer interaction.

While virtual environments are still gaining popularity and momentum, everyday video conferencing seems to be heading to its next evolutionary leap; on one hand the desktop based video conferencing utilising either messengers or web conferencing have finally brought video conferencing to the masses and the easy accessibility keeps the usage on steady ascent – on the other hand the new group conferencing systems are capable of handling HD quality video and good quality audio thus greatly enhancing the video conferencing experience to a more life like scenario. At the high end are the special *telepresence rooms* with multiple HD-quality conferencing screens, accommodating life sized participant images and standardised lighting, layout and furniture which have lately been getting more popular especially in corporate settings.

To sum it up, it can be stated that video conferencing is evolving and spreading to new areas. Combined with the ever developing software solutions, increasing processing speed and network capacity, and growing environmental awareness this suggest that the golden age of video conferencing might finally be at the verge of realisation.

3 Video conference technologies

In this chapter, the focus is on the most important standards and protocols related to video conferencing field. This is the most technical part of this thesis and a less technologically oriented reader might be more interested in jumping directly to chapter 4 and utilising this chapter as a reference to the technologies when needed. However, skimming through the headers of the following subsections is encouraged to get a feel on the acronyms as they play an important role when judging on the quality and interoperability issues between different video conferencing solutions. It also forms the basis for understanding the technical details presented in the comparison in Appendix 2.

In the following subsections, all G.xxx and H.yyy standards refer to the corresponding specifications of the International Telecommunications Union [ITU08]. Each section also contains an example of a proprietary codec popularly used in the area in question.

3.1 Signalling

Signalling means the different kinds of standards used in video conferencing systems to control communications between systems. As a rule of the thumb it can be said, that basically video conferencing is possible only between systems using the same standard. However, there are specific gateways which can be used to interconnect endpoints using different standards, which will be discussed separately. In this subsection, the most common video conferencing signalling standards are introduced.

H.320

The ITU H.320 standard describes a video conference connection over ISDN networks. The method gained momentum after the introduction of ISDN telephone

lines and thus brought standardised video conferencing to a larger audience. On the upside, it was possible to use a basically universal digital phone network for calling sites around the world (although differences in the national ISDN implementations caused some problems). Although this was much cheaper than leasing a special fibre optic or satellite connection, the expenses could still be noticeable: holding a one hour international video conference at 384kbps speed would have six times the cost the caller the international dialling fee (six parallel ISDN B-channels needed for the 384kbps connection) multiplied by 60 minutes call duration. So, while cheaper than travelling, the H.320 video conferencing was not a cheap option – the equipment was expensive to acquire, and the operating cost including the telephone line fees did cast another overhead as well. Also, the interoperability between systems from different vendors is often considered problematic.

Despite the problems stated above, H.320 video conferencing is still being utilised around the globe and it is especially important in areas where broadband internet connections are not yet available or where they fail to provide a reasonably reliable bandwidth. Lately the quality of H.320 conferencing has gone up thanks to the new, more efficient audio and video compression techniques.

H.323

The ITU H.323 standard from 1996 mapped out the way for video conferencing systems to connect over the Internet. Compared to the older H.320 standard, the telecommunications fee was now considerably lower – no billing per minute depending on the destination; instead, the video conference system used the normal internet connection of the facility.

Video conferencing over the internet brought also many new problems. One was the bandwidth; to have a reliable 384kbps video conference connection it is generally suggested that one should have a double of the bandwidth in their disposal to compensate for possible network congestion (in this case 768kbps

upstream and 768kbps downstream would be recommended although 512kbps for both directions might be sufficient). One thing to notice here is, that the aforementioned internet speeds are needed for the video conference connection alone – and as the units often share the internet connection with all other traffic from the organisation, this other traffic can easily make momentarily bottlenecks for the video conference connection unless special techniques such as bandwidth throttling or QoS (quality of service) settings are used on the network layer.

Nowadays broadband internet connections are generally fast enough for video conferencing, but still especially in rural areas the access to fast broadband can be problematic and thus limit the call speed and the quality of the conference. International video conferencing calls also seem to present a problem in some countries and sometimes a guaranteed bandwidth from a network operator might be needed.

Another problem related especially to H.323 video conferencing standard is in the definition of the protocol. Besides using readily agreed ports, such as port 1720 for connection signalling, the ports for audio and video are dynamically selected for each session in the range of 1024 to 65535. This causes understandable problems with many firewalls or NAT (Network Address Translation) implementations. There are some workarounds and a new tunnelling standard H.460, but these solutions usually include setting up special software and/or hardware to handle the connection between the local endpoint and the public internet.

H.324 and H.324m

The ITU H.324 is a standard for video conferencing over 'normal' PSTN – the Public Switched Telephone Network. In a way, it's the modern standard for picture phones, but the technology is not widely adopted.

A newer technology based on the H.324 specification is the ITU H.324m, which is probably more familiar to people as the video call capability on the new 3rd

generation (3G or UMTS) mobile phones. Usually, a dedicated gateway is needed for interoperability between H.324m and H.320/H.323, but some newer H.323-endpoints are capable of handling H.324m calls independently.

SIP

The IETF standard SIP (Session Initiation Protocol) differs in many ways from the H.323 video conference standard. Modelled after the ideas in HTML and SMTP, SIP uses simple text-based signalling and allows much freedom to the implementation used in the actual session [ScR98]. Because of the differences in implementation and functionality SIP and H.323 are not directly interconnectable.

In the recent years, SIP has become quite popular in VoIP (Voice over Internet Protocol) products and video call implementations. SIP forms the basis for signalling in many familiar applications such as iChat and Live Messenger (formerly MSN Messenger), etc. SIP is generally seen as a lightweight protocol with more flexibility and extensibility than H.323 and thus it is considered as a strong advocate for the future implementations in the field [KPP04].

Last years have seen SIP becoming a more common feature even in standalone video conference units and MCUs (usually accompanying the H.323 as an alternative connection standard), but so far the SIP implementation in the video conferencing endpoints has usually been on a basic level, not yet utilising many of the possibilities in the standard. Still, this direction of integrating multiple connection technologies into single endpoints is an encouraging sign suggesting increase in the interoperability between different systems.

3.2 Video codecs

Video codecs are essential for video conferencing as they are the technology used to compress the video signal into a series of data packets relayed over the network,

to be decompressed at the receiving site to reform the video image. In this subsection the most popular video compression codecs related to the video conferencing field and implications of the quality they provide are described.

H.261

ITU H.261 is an older video coded, implemented in almost all video conferencing systems. It was originally introduced in 1988 and later revised in 1990 and 1993. Because of the long history it is still an important codec for interoperability as nearly all video conferencing solutions support it.

The picture quality of H.261 is not very good and the colour information is reduced, but the video stream still requires a relatively large amount of bandwidth (compared to the newer video codecs). Thus, with the limited bandwidth of 1990-s, the experienced image quality of older video conference systems was often relatively degraded.

H.263

The ITU H.263 is a more developed video codec. Compared to earlier H.261, the newer codec improved compression efficiency and offered a better colour dynamic. The codec also got multiple extensions, e.g. the so called H.263+ (a.k.a. H.263v2) and H.263++ (a.k.a. H.263v3) which introduced for example customised image sizes and improved decoder memory management. Old Flash players utilised a proprietary Sorenson Spark codec which was based on H.263 [Tab07].

H.264

ITU H.264 (also ratified as the ISO's MPEG 4 AVC) is a standard for efficient picture compression. Relatively new and complex, it requires much more processing power than its predecessors and thus cannot often be implemented onto

old hardware. Besides better compression, H.264 also includes advanced functionalities such as an *in-loop deblocking filter* which effectively eliminates possible packing artefacts from the image [WSB03], which were often considered as a noticeable picture quality problem when using older video codecs.

The efficient compression of H.264 resulting in a reduced need for bandwidth enabled the possibility to use bigger video resolutions at reasonable connection speeds, which in turn has largely contributed to the growing use of HD (High Definition) resolutions in video conferencing. H.264 is also an integral part of many other high definition video implementations such as Blu-Ray, HD-DVD and DVB-T2 (Digital TV Broadcasting) [Kau07]. It is also emerging as a new de-facto standard of streaming internet video (supported by many popular media players such as QuickTime, VLC and even by the newest Adobe Flash Player 9) [Tab07].

3.3 Audio codecs

Audio codecs are protocols used to compress audio signals for network transport. In this subsection, the most common audio codecs utilised in video conferencing are described with some technical detail.

G.711

The old G.711 sound codec (ratified in 1988) is widely used around the world. This codec is probably familiar to the reader as the 'telephone quality' as it is the codec actually used to encode voice in landlines and mobile phones.

The efficiency of G.711's PCM encoding is left far behind in comparison to the newer codecs as it uses either 48kbps, 56kbps or 64kbps bandwidth but is still only able to provide 3.1KHz audio band (300 Hz – 3400Hz) which cuts out both low and high sounds. Still, as a de-facto standard it's very important especially regarding backwards compatibility between different video conferencing solutions.

The G.711 can be used for basic speech, but loses a lot of the dynamic and clarity of the voice, and the narrow frequency bandwidth does not function very well with other sound types such as music. However, it is still used in everyday telephone discussions all the time, and this will probably also be the case for many years to come.

G.722

The G.722 sound codec, also ratified in 1988, uses the same bandwidths than G.711 (48, 56 or 64kbps), but is more advanced by offering a wider sound spectrum (7kHz, 50 Hz – 7000Hz) and thus making the sound clearer and more understandable. As G.711, the G.722 is a widely implemented protocol and forms a good basis for interconnectivity, but provides clearly better quality.

Other widely used audio codecs

The G.722.1 offers the same 7KHz audio band as G.722, but thanks to better compression it uses less bandwidth, namely 24 or 32kbps. This makes G.722.1 a better option when bandwidth is an issue, although there are also codecs, such as G.723.1, which function with even less bandwidth.

The G.722.1 annex C, Standardised from the originally proprietary Polycom Siren 14 codec, offers a more advanced 14 kHz (50 Hz – 14400Hz) audio band allowing more high frequencies and thus enhancing the clarity of sound. The difference is very clear for example when using music or other sound sources than speech.

The G.723.1, ratified in 1996, is a very low bitrate audio codec, which offers a 8 kHz audio band and uses only 5.3 or 6.3 kbps of network bandwidth, thus making it ideal for conferencing when bandwidth is an issue.

The G.728 audio coded from 1992 uses a 16 kbps bitrate to produce 3.3Khz audio band (50Hz – 3400Hz). This codec is also quite widely implemented and thus offers good interconnectivity. However the quality is poor compared to some of the newer codecs.

AAC

A recent addition to the video conferencing field, the MPEG 4 AAC codec provides much more advanced sound quality than the G-codecs and is thus the most recommendable codec in case audio quality is especially important – for example when using music or showing movie clips over the video conferencing connection. The newest versions of AAC, such as HE-AAC v2 (AAC with SBR and PS), have been found out to provide best sound quality for the network bandwidth used, when compared to the other popular audio codecs [MeM06], and as a superset of AAC these implementations are also getting more popular.

AAC implementations in video conferencing unit offers for example 14 kHz, 20 kHz, 22 kHz and 44 kHz audio bands (mono or stereo) using 38, 56, 64, 96, 128 or 192 kbps of network bandwidth. Interesting aspect to notice here is that equipment from different vendors can usually handle only a subset of the bandwidths and/or audio bands mentioned above, and thus interoperability issues might arise even both are capable of handling AAC sound.

Besides using different bitrates, some of the video conferencing equipment use different versions of the standard: AAC-LC (Low Complexity) and AAC-LD (Low Delay). The benefit for using AAC-LD is that the audio compression delay stays constant at 20ms (0.02 sec), while using other AAC implementations the delay gets bigger as the used bitrate gets smaller – AAC with 32kbps can accumulate to a compression delay of over 250ms (0.25 sec), thus making the delay too long for real-time conferencing.

3.4 Proprietary methods

There exist a number of systems that utilise other than standardised methods either for call signalling or audio or video encoding, or both. The reasons for this can be versatile – to avoid licensing costs, or to overcome limitations ingrained in existing standardised solutions.

Systems using proprietary signalling are usually not interoperable with others, but can only connect to other similar systems – however it is possible to construct special gateways to provide interconnectivity. One example of system, utilising speciality signalling, is Skype, which even uses encryption to cipher the messaging on the network [BMM07].

The VP6 video codec from On2 Technologies is probably one of the most utilised proprietary video codecs in the video conferencing field – especially when it comes to desktop video conferencing. The VP6 codec is widely used in applications such as Skype, AOL AIM, etc. [On207]. The codec is also built in to Adobe Flash Player 8 and 9 and is usable on almost any computer with a modern web browser – thus the success of the codec in web based video conferencing solutions.

The proprietary codecs iLBR (internet Low Bit Rate codec) and iSAC from Global IP Systems are widely used in everyday internet applications such as Skype, Yahoo Messenger and AOL AIM as well as in net2phone, Marratech, WebEx, Gizmo, Google Talk, etc. The iLBR offers 4 kHz audio band at 15.2 or 13.3 kbps while iSAC offers better 8 kHz audio band with adjustable network bandwidth between 10 to 32 kbps [GIP07]. As technologies specially developed for internet use, they have a standard response to delay and jitter thus smoothing out the occasional problems with the network connection.

3.5 Other related standards

Other standards related to video conferencing include H.239 for additional media – usually utilised for sharing computer screen – as well as other internet and H.xxx family related standards. In the following these are briefly introduced.

H.239

The ITU H.239 is not a video conference standard in a sense of H.320 or H.323, but it's a way of including a second video stream within a H.323 (or H.320) video conference connection. Basically, this means that it is possible for example to send images of both the speaker and his slides in a single video conference connection so that the receiving end sees both streams simultaneously (for example on two different screens).

Previously, most of the manufacturers had their own proprietary 'dual stream' implementations, such as Tandberg's DuoVideo and Polycom's People+Content, but as a standard H.239 has levelled the field and now the interoperability between different manufacturers is much easier. Many new multipoint control units are also capable of handling H.239 in a multipoint session, which greatly expands the usability of the standard.

Internet standards

On the quite standardised field of video conferencing, there are a huge number of other protocols and standards. They work mostly in the background and are not as important for the normal user to understand, and thus they will not be described in as much detail.

IP – Internet Protocol – the layer transporting traffic on the internet
TCP – Transmit Control protocol – used for example for signalling in H.323
UDP – User Datagram protocol – used often for audio and video feeds
RTP – Real-time Transport protocol – used for audio & video data transport over UDP
RTCP – Real-time Transport Control protocol – used to control RTP traffic
IP Multicast – used to efficiently transport audio & video in multicast enabled networks

Table 1: Internet standards related to video conferencing

H.233, H.234, H.235 – security and encryption in H.323, H.320
H.245 – multimedia control in H.323
H.450 – generic added functionality for H.323, like call transfer (H.450.2) and call diversion (H.450.3)
H.460.18/19 firewall traversal for H.323 signalling/media

Table 2: ITU standards related to H.320 and H.323 video conferencing

The IP-video conferencing field (ie. video conferencing over internet) builds on a number of Internet related protocols and standards, which are not evaluated within a scope of this thesis. The most important internet standards are briefly introduced in Table 1. There exist also a number of ITU ratified standards that function within the H.320 and H.323 video conferencing. A selection of the most important standards is introduced in Table 2.

Video resolutions

Besides video codecs, there is another variable having a strong impact on the video quality – video resolution. The resolution can be understood as the actual size of the

video image (while even a small video image can be zoomed in to fill the whole screen, it would appear to be pixelated and thus of low quality).

Name	Resolution (pixels)	As a digital image
SQCIF	128 x 96	0,012 mpix
QCIF	176 x 144	0,025 mpix
QVGA	320 x 240	0,076 mpix
CIF	352 x 288	0,1 mpix
VGA	640 x 480	0.3 mpix
4CIF (“television”)	704 x 576	0,4 mpix
720p (“HD ready”)	1280 x 720	1 mpix
1080p (“Full HD”)	1920 x 1080	2 mpix

Table 3: Popular video resolutions

The most popular video resolutions are listed in Table 3. The formats QCIF and CIF are defined in the specifications of H.261, while SQCIF and 4CIF are specified in H.263 [ITU08]. VGA and QVGA resolutions are commonly utilised in computer-based communications solutions alongside varying codecs. High definition is normally used with H.264 encoding.

4 Selecting tools for video conferencing

Selecting which video conference tools to use can be a daunting task. There are many variables and focus groups to consider plus the technology is always on the move. In this chapter, two methods for comparing video conference tools are presented and general guidelines and best practises given.

4.1 Popular video conferencing solutions

The recent decade has seen an emergence of multiple video conferencing solutions. Some of the solutions are based on new, advanced standards while others are proprietary. In the following, two examples of the most popular solutions are referenced briefly. Details of the standards and codecs mentioned can be found in the previous chapter. A general technical comparison of the popular technologies can be found in Appendix 1, while a more detailed comparison chart evaluating a number of popular solutions can be found in Appendix 2.

Standard video conferencing

When people generally address *attending a video conference*, it often denotes an ITU standards based video conference, usually utilising the H.320 (over ISDN) or H.323 (over IP) technology. There are multiple vendors manufacturing these products, and different products can have different capabilities – the newer units tend to be able to handle a good quality stereophonic sound such as AAC and H.264 encoded HD video, while older units might be only able to process low quality mono sound of G.711 and a low quality image of H.261. Due to this versatility the standards emphasise on backwards compatibility and thus even the new units are still interconnectable with the older models, although with reduced quality and functionality.

Flash-based web conferencing solutions

Partly due to the wide penetration of Adobe's Flash player, the Flash-based solutions, such as Flashmeeting and Adobe Acrobat Connect Pro (formerly Macromedia Breeze), have lately gained much popularity. Another reason for their popularity might be the ease of use - as the web conference works directly within a web browser without any specific installations. The image size and quality of these web-based solutions are still much reduced from that provided by the new H.323 systems, but web conferencing solutions often offer enhanced interaction capabilities such as embedded multipoint and conference recording, text chat, voting, material sharing and whiteboards, etc.

Other popular solutions

Today there exists multiple instant messengers and internet phone solutions with video functionality. Many of these are based on the SIP (Ekiga, iChat) or Jabber (Google Talk, Gizmo) signalling protocols, while others utilise proprietary solutions (Skype). Some systems have a selection on signalling capabilities. Interoperatibility is only possible between systems supporting the same signalling protocol, and still issues can arise if solutions do not support the same protocols and codecs for audio and video feeds.

4.2 Comparing video conferencing systems

Video conference systems can be compared in many ways. One can compare their technical specifications or ponder on the financial aspects. For successful outcome it is important also to pay attention to the educational and operational aspects.

The SECTIONS -framework

The SECTIONS framework, originally called ACTIONS framework, was presented in 1995 by Tony Bates [Bat95, p.35-59]. The acronym implies a seven axis

comparison framework which takes into account the many different aspects related to distance education. In the newer version from 2003 Bates updates the ACTIONS to SECTIONS framework [BaP03, p.79-80]. The components of the SECTIONS framework are *Students*, *Ease of Use*, *Costs*, *Teaching*, *Interactivity*, *Organisational issues*, *Novelty* and *Speed* [Bat95, p.2-12, BaP03, p.79-80]

As stated by Tony Bates, the SECTIONS framework is a tool for an decision-makers to utilise, for example, in determining which video conferencing solutions to provide support and infrastructure to. An example of using the framework to evaluate Flash media can be found in [BDR05].

The IDEAS -model

The needs of an organisation might not be similar to the needs of an individual teacher. The IDEAS model was developed by the author especially for this thesis. The IDEAS model can be viewed as a subsection of the SECTIONS framework with added emphasis on the unique characteristics of the context and the environment where it is to be implemented.

The main difference between IDEAS model and SECTIONS framework is that *the IDEAS model is intended to be used by an individual with a specific need for a video conference solution* – for example, a teacher looking for a tool to use on a specific course. Thus it should be noted, that, compared to SECTIONS framework, the IDEAS model is even more subjective and context dependant.

As the model, intended to be used with a specific context, need and focus group in mind, the resulting subjectivity and context dependence should not be seen as a problem but rather as a benefit of the model. Thus, the model should be seen more as a practical tool, than a universal theoretical framework for rating the overall excellence of different video conferencing solutions.

The components of the IDEAS model are *Interactivity*, *Details*, *Economics*, *Accessibility* and *Support*. In the following subsections the main dimensions of the IDEAS model are explained in more detail.

Interactivity

The interactivity dimension emphasises the interactive nature of video conferencing. As stated earlier in this thesis, video conferencing is inherently a social and collaborative medium and it should also be used accordingly – sometimes it is justifiable to use video conferencing to conduct unidirectional sessions, but usually this is not the case.

Concerning video conferencing mediated education, it has been noted in many articles, that the one critical aspect for student satisfaction in a video conference learning session is the interaction between the presenter and the students [Ami03, LüL05]. It has also been suggested that especially for K-12 education, i.e. elementary, middle and high schools, the best use of video conferencing could be to utilise it to have 'expert sessions' or 'virtual tours' – that is not to have long one-way lessons for students to observe, but to have shorter interactive discussions and collaborations between students and remote presenters about subjects already introduced in the local classroom [Ami03, Mer05].

Interactivity covers such methods as multipoint audio, multipoint video, ability to show desktop, share files, show a document camera image, or to utilise text-based chat or interactive whiteboard. Usually, not all of these are used in a single session (or even in a single course or project) so the selection of tools might be different for different needs and different pedagogical approaches. Thus, knowing the context and the preferred pedagogical approach is valuable in weighting the importance of different interaction methods.

The use of accompanying asynchronous interaction tools, such as web based discussion forums, could also be considered; for example, in bigger conferences this enables much larger group of participants to be able to interact and express their questions or contributions than what is reasonable to do through the video connection. In the case of using a recording of a video conference session as a material to a new group of participants, asynchronous modes of interaction should be greatly encouraged.

Details

The details dimension emphasises the focus on the immediate context, environment and restrictions related to the planned need. For example, the requirements for a video conferencing solution are quite different in a course where participants would attend the sessions from their home, than on a course where participants are located on two campuses. Also basic details should be mapped out, such as how many participants would be attending, how many sessions there would be, how important is the video quality, is there a need for a recording for later viewing, etc.

A need to have a shared work space to collaborate simultaneously by writing and reviewing a document might be an important detail. The need for a possibility to connect participants from 'out on the field' might pose a different kind of a restrictive detail. It is important to think the details first from the optimal pedagogical and social points of view before submitting them to the restrictions of the possible technologies – sometimes there might even exist a sound solution that the teacher is just unaware of; thus first mapping out the optimal functionality and then consulting an expert in educational technology or media education might reveal new possibilities.

Economics

Video conferencing is often marketed by its cost- and time saving aspects as travelling to other sites can be reduced by utilising the video conferencing

technology. Watson et al. describe the usage of video conferencing in surgical conferences to relay audio, video and content feed to two remote hospitals. The researchers calculate that just the savings in travel time (13 sessions x 1 hour saving per participant x 81 participants) add up to over 110,000$ [WLB07]. However, in this case the researchers do not provide any details on the subjective experiences or on the learning outcomes of the students in the pilot.

The economics dimension is included in the model to take account the financial implications of using a selected technology. Sometimes the economics might look very different depending the role the reviewer has in an organisation; web based conferencing might be free for the end users, but the IT department might see it differently as requiring servers, software licences and administration work which all come at a cost.

According to the philosophy of the IDEAS model the economics of a technology should be seen from the perspective of the individual teacher or meeting organiser (and the related course or project) as that is the concrete level of monetary expense associated with using a specific technology in relation the the planned course or project. Possible end user expenses (such as buying a headset and a web camera or maintaining a broadband internet connection) might not be counted in as an expense, but the amount of participant contribution should also be roughly estimated as there might be vast differences in relation to the requirements of different systems. If new equipment is planned to be purchased, attention should be paid to for both the quality the unit offers (connection methods, audio and video codecs, network bandwidth, etc) and for the possibilities it covers (connecting multiple sites, delivering content in addition to video feed, etc).

Accessibility

The accessibility dimension focuses on the accessibility of the system. For example, if the system is a web based conferencing tool which can be used with any

modern web browser, it is quite safe to assume that many of the potential participants could be able to access the system either from home or from workplace, although probably some of the participants might not have a webcam required for two way video. On the other hand, if the system being analysed is an expensive high quality group video conferencing solution, then it is probably accessible only in some dedicated conferencing rooms – which might not present a problem, if the plan is to use the system within a corporation between two offices.

The participant group and the context where the system is planned to be used plays an important role when outlining the diffusion aspect. In an optimal situation, it would be advisable to map out the systems and technologies already accessible by the focus group, and to ponder if some of the readily available systems could have the other characteristics needed.

Sotillo describes a case study of using Yahoo IM in student-tutor communication in teaching English as a second language. The results in this case report the prevailing problems in the United States with acquiring adequate bandwidth for having reliable audio/video connection which in turn resulted in users dismissing the video channel and using only audio communication [Sot06]. The continuing problems with adequate bandwidth seem to be still an issue both in developing countries as well as some geographically large countries, such as US and Australia, and in the effected areas the bandwidth issues have to be accounted for when deeming the suitability of a system.

Support

It has been noted in the literature, that especially new users of video conferencing do better if adequate support is provided at the beginning [Alb06, p.6, Ami03]. For example, in educational setting teachers should receive video conference -related instruction well in advance. The presence of technical support especially at the very first teaching sessions has been noted very important for the overall success and

acceptance of the technology. Also students are suggested to benefit from *instructional scaffolding* (i.e. cognitive support and guidance), when new to video conferencing.

The support dimension focuses on the scaffolding aspects - is there help readily available? In case the system would be used in a corporate or university setting, this can be seen as if the system is being supported by the organisation, i.e. there is technical maintenance, training and end user support available, etc. The ease of use of the system and the possible prior experiences of the users play a big part in determining how much support might be needed. It should be noted though, that even if the organiser is experienced and does not see support as an important aspect, some other participants might experience much anxiety if the system is new and there is no help available for them in case of problems – which in turn could lead to an increased drop out rate. Thus, the more novices there are among the participants using the system, the more important is the support aspect.

4.3 Selecting most suitable video conferencing tools

The selection of suitable video conferencing methods should start from the need of the course and context. Different needs implicate different solutions, although it is often the existing infrastructure – what systems are already installed at the organisation – that greatly influences the outcome. By utilising the IDEAS model the process of selection can be made more visible and more rational.

The application of the IDEAS model requires some knowledge on the possible solutions and their technical and functional differences. Thus, teaming up with an educational technologist could prove out to be beneficial, but one can also consult colleagues and students, and research the possibilities on the internet.

The selection of possible video conferencing tools available varies from organisation to another, but generally the tools can be divided into the following

categories: H.323 and H.320 systems, SIP or Jabber messengers, Web conferencing platforms, and other solutions. For detailed examples of tools from these categories, see Appendix 2.

I	
	Is there need (and possibility) for data sharing?
	Is there need (and possibility) for collaborative work?
	What possibilities exist for alternate feedback channels?
D	
	How many sites (and participants per site) are estimated?
	How often is the session scheduled?
	How much emphasis is put on audio and video quality?
	What are the specific needs?
E	
	What are the operating expenses (per site)?
	Is there a connection fee (or a multipoint fee)?
	What are the economical constraints?
A	
	What kind of technology is available at the institution?
	What kind of technology is familiar for the participants?
	What kinds of limitations are there (bandwidth, etc.)?
S	
	Is there training and support available for the teacher?
	Is there support readily available for all participants?
	Where to get the help when needed?

Table 4: Example of using the IDEAS model.

The IDEAS model can be utilised, for example, with the help of Table 4, by first answering the questions generally – considering the best of the course – and then specifically for each of the possible solutions in consideration to map out how well they fulfil the expectations.

Examples of current practises

Some popular practises on selecting video conferencing tools can be stated. According to the experiences of the author, *if the session takes place between two (or more) institutions*, the technology utilised is often the standard ITU H.323 - based video conferencing possibly accompanied with infrastructure tools such as a MCU (to connect multiple participants), a gateway (to interconnect H.323 (Internet) participants to H.320 participants i.e. the participants using equipment with the ISDN connectivity), and possibility recording (to save the session for later viewing). The reason for this is the wide deployment and support available for the H.323 technology in the educational setting.

On the other hand, *in case of multiple participants - if most of the participants are attending the meeting from their own workstations* (either at work, at home or at any place where they have the necessary means) the meetings are often held on web-conferencing systems such as Connect Pro, EVO or Flashmeeting. The reason for this is the ease of access - usually no need for special programs, firewall configurations etc. – and these meetings can also be attended by a group as well as on individual (the desktop just might have to be beamed by a projector or screened on a plasma screen). Most of the web-conferencing tools also incorporate the possibility for many kinds of interaction such as shared work space, voting, chat, sharing different types of data and delegation of control of the session.

If the meeting takes place is just between two participants, the most convenient method might be utilising a messenger-style program such as Live messenger or Skype, or a SIP client such as iChat or Ekiga. The reasons encouraging the use of this method is the wide deployment of the necessary software; in some cases most of the possible users might already have some messenger client installed on their computer and they might already be familiar with the program. With reasonable quality these solutions can perform well in many kinds of scenarios from tutoring to collaborating etc. However, these scenarios should be seen only as examples of solutions which can be analysed more thoroughly by utilising the IDEAS model.

5 Conclusions

The main benefit of video conferencing is overcoming the limitations of distance. With different technologies, it is possible to connect geographically dispersed persons or groups to have a meeting or a collaborative work session, and to broadcast the session live for internet audience or to save the session for future use (or both).

As seen, through the course of this thesis, video conferencing is a broad phenomenon covering many technologies and possibilities. A video conference can be between two or multiple locations, it can utilise specialised equipment or run an a normal computer – it can be assisted with content sharing, it can be accompanied with other feedback channels and it can also be streamed live or recorded for later use. There are many solutions and some of them might be better for a specific case than others. The development of the IDEAS model focused on emphasising the planned context in the selection process.

Choosing the most suitable video conferencing tools and utilising them to their best ability can be a confusing task especially for a novice. The IDEAS model suggests, that a sound candidate for a video conferencing solution, can be found by focusing on the *Interactivity*, *Details*, *Economics*, *Accessibility* and *Support* in the corresponding context. By selecting the best tools, and utilising them according to the best practises, it should be possible to hold a successful conference, and thus to empower the participants to take advantage of the new technology by overcoming the limitations of distance.

References

Unless otherwise stated all websites accessed in October 2007.

Links marked with * seem to be offline in May 2008, but are accessible via the Internet Archive's WayBackMachine at http://www.archive.org.

Alb06 Alberta Education, Video-conferencing Research Community of Practice: Research Report. Video Conference Alberta, 2006. http://www.vcalberta.ca/community/Research_Summary_Report_word_version_final.pdf

Ami03 Amirian, S., Pedagogy & Video Conferencing: A Review of Recent Literature. Presentation at the 4th Megaconference, 2002. (Accessed from Internet Archive) http://www.iclassnotes.com/amirian_megacon.pdf *

BHN07 Batcheller, A. L., Hilligoss, B., Nam, K., Rader, E., Rey-Babarro, M. and Zhou, X., Testing the Technology: Playing Games with Video Conferencing. ACM CHI Proceedings, 2007, p.849-852

BaP03 Bates, A. W. and Poole, G., Effective Teaching with Technology in Higher Education. Jossey-Bass, San Fransisco, 2003.

Bat95 Bates, A. W., Technology, Open Learning and Distance Education. Routledge Studies in Distance Education. Routledge, London, 1995.

BeA04 Berner, E. S. and Adams, B., Added Value of Video Compared to Audio Lectures for Distance Learning. International Journal of Medical Informatics, 73, 2004, p.189-193

BMM07 Bonfiglio, D., Mellia, M., Meo, M., Rossi, D. and Tofanelli, P., Revealing Skype Traffic: When Randomness Plays with You. ACM SIGCOMM Computer Communication Review, 37(4), 2007, p.37-48

BDR05 Boyes, J., Dowie, S. and Rumzan, I., Using the SECTIONS Framework to Evaluate Flash Media. Innovate 2 (1), 2005. http://www.innovateonline.info/index.php?view=article&id=55

Dud06 Dudman, J., Voice over IP: What It I, Why People Want It, and Where It Is Going. JISC Technology and Stardards Watch, September 2006 http://www.jisc.ac.uk/media/documents/techwatch/tsw0604.pdf

FiT02 Fischer, K. and Tenbrink, T., Video Conferencing in a Transregional Research Cooperation: Turn−Taking in a New Medium. In Döring, J., Schmitz, H. W. and Schulte, O. (Hrsg.), Connecting Perspectives. Videokonferenz: Beiträge zu ihrer Erforschung und Anwendung. Shaker, Aachen, 2003. http://www.informatik.uni-bremen.de/%7Etenbrink/VKFischerTenbrinkproc.pdf

GiL07 Gibbs, B. and Larson, E.. 2007. Using Video Conferencing in Lecture Classes. Innovate 3 (5), 2007. http://www.innovateonline.info/index.php?view=article&id=388

GIP07 Global IP Solutions, High Quality Codecs. Global IP Solution Website, 2007. http://www.gipscorp.com/high-quality-codecs/index.php

GrM03 Grayson, D. M. and Monk, A. F., Are You Looking At Me? Eye Contact and Desktop Video Conferencing. ACM Transactions on Human-Computer Interaction, 10(3), 2003, p.221-243

HeH05 Heeler, P. and Hardy, C., A Preliminary Report on the Use of Video Technology in Online Courses. JCSC 20(4), 2005, p.127-133

ITU08 International Telecommunications Union, Recommendations. ITU web site 2008. http://www.itu.int/rec/T-REC/en

Kau07 Kauranen, J. Teräväpiirtoformaatit HD DVD ja Blu-Ray (In Finnish). Batchelor's Thesis, TAMK University of Applied Sciences, 2007. https://oa.doria.fi/bitstream/handle/10024/5866/Kauramaa.Janne.pdf

KPP04 Katrinis, K., Parissidis, G. and Plattner, B., A Comparison of Frameworks for Multimedia Conferencing: SIP and H.323. Proceedings of IASTED Internet Multimedia Systems and Applications, Kauai, 2004. http://people.ee.ethz.ch/~katrinis/pubs/Katr_IMSA04.pdf

LüL05 Lück, M. and Laurence, G. M., Innovative teaching: Sharing expertise through videoconferencing. Innovate 2 (1), 2005. http://www.innovateonline.info/index.php?view=article&id=59

MeM06 Meltzer, S. and Moser, G., MPEG-4 HE-AAC v2 – Audio Coding for Today's Digital World. EBU Technical Review (305), Jan 2006. http://www.ebu.ch/en/technical/trev/trev_305-moser.pdf

Mer05 Merrick, S.,Videoconferencing K-12: The State of the Art. Innovate 2 (1), 2005. http://www.innovateonline.info/index.php?view=article&id=24

NgC07 Nguyen, D. and Canny, J., MultiView: Improving Trust in Group Video Conferencing Through Spatial Faithfulness. ACM CHI Proceedings, 2007, p.1465-

NgC07 Nguyen, D. and Canny, J., MultiView: Improving Trust in Group Video Conferencing Through Spatial Faithfulness. ACM CHI Proceedings, 2007, p.1465-1474

NYT24 Pictures By Wire Sent With Success for the First Time. New York Times, May 20, 1924

On207 On2 Technologies, Why Choose On2 Video? On2 Technologies Website, 2007. http://www.on2.com/technology/on2-difference/ *

ScR98 Schulzrinne, H. and Rosenberg, J, A Comparison of SIP and H.323 for Internet Telephony. Proceedings of NOSSDAV, Cambridge, U.K. Jul. 1998. http://www1.cs.columbia.edu/~hgs/papers/Schu9807_Comparison.pdf

ScW04 Schnaars, S. and Wymbs, C., On the Persistence of Lackluster Demand – the History of the Video Telephone. Technological Forecasting & Social Change 71 (3), 2004, p.197-216

Sal02 Sallnäs, E.-L., Collaboration in Multimodal Virtual Worlds: Comparing Touch, Text, Voice and Video. In Schroeder, R. (editor), The Social Life of Avatars, Springer-Verlag, 2002, p.172-187.

Sar98 Sariola, J.,The Planning of an Open Learning Environment and Didactic Media Choice in Teacher Education. In T. Nummi, A. Rönkä and J. Sariola, Virtuality and Digital Nomadism – An Introduction to the LIVE Project (1997–2000). University of Helsinki. Media Education Publications 6, 1998, p.23-49.

Sec08 Second Life Grid, Education and Non-Profit Organisations. Second Life Grid website, 2008. (Accessed 26.4.2008) http://secondlifegrid.net/programs/education

Sot06 Sotillo, S. M., Using Instant Messaging for Collaborative Learning: A Case Study. Innovate 2 (3), 2006. http://www.innovateonline.info/index.php?view=article&id=170

Tab07 Taborga, T., Flash Player 9 Update H.264. Adobe Labs Wiki, 2007. http://labs.adobe.com/wiki/index.php/Flash_Player:9:Update:H.264

TaI92 Tang, J. C. and Isaacs, E. A., Why Do Users Like Video? Studies of Multimedia-Supported Collaboration. Technical report. Sun Microsystems Laboratories 1992. http://www.research.sun.com/research/features/tenyears/volcd/papers/2Tang.pdf

Vid05 Vide.Net Video Conferencing Cookbook. Vide.Net 2005 http://www.vide.net/cookbook/

Wan04 Wang, Y., Supporting Synchronous Distance Language Learning with Desktop Videoconferencing. Language Learning & Technology 8(3), 2004, p.90-121.

WLB07 Watson, A. R., Lee, K., Billiar, T. and Luketich, J.,Interactive Multi-Site Video Teleconferencing as a Means to Increase Surgical Resident and Staff Participation at Core Teaching Conferences. Surgical Forum Abstracts, 205(S3), 2007, p.S85

WSB03 Wiegand, T., Sullivan, G. J., Bjøntegaard, G. and Luthra, A., Overview of the H.264/AVC Video Coding Standard. IEEE Transactions on Circuits and Systems for Video Technology, 13(7), 2003, p.560-576

Appendices

Appendix 1 – Technical comparison

Technology	Network	Codecs	Multipoint	Content	Recording	Chat	Quality
H.320	ISDN	ITU	MCU needed	New systems	Add on	No	VM to HD
H.323	IP	ITU	MCU needed	New systems	Add on	No	VM to HD
SIP	IP	Depends	Varies	Varies	Rarely	Usually	VM to DVD
Flash-based	IP	Flash supported	Usually	Usually	Usually	Usually	VM to VHS
Java-based	IP	Depends	Varies	Varies	Varies	Usually	VM to DVD
Proprietary	IP	Depends	Varies	Varies	Rarely	Usually	VM to DVD

Scale: No, Rarely, Varies, Usually
Quality scale: VM (Video Mail), VHS, DVD, HD

Table 1. General evaluation

The technical comparison is a rough estimate on the possibilities of the various systems. It should be noted, that there is a lot of variation in each of the groups, so more detailed data needs to be used on the basis of the evaluation.

Appendix 2 Technical details

Legend: p-proprietary, de-decode, CP-continous presence, VS-voice switching, €-extra cost

			Price	Standards (max speed per second)				
		Type	endp / srv	H.320	H.323	SIP	Other	Recording
H323-boxes	Aethra Vega X7	SetTop	€	512k €	4 Mb	4 Mb		
	Aethra Vega X5	SetTop	€	2 Mb €	2 Mb (4Mb €)	Mb (4Mb €)		
	Lifesize Room	Codec	€	-	5 Mb	5 Mb		
	Polycom VSX-8000	Codec	€	2 Mb €	2 Mb	2 Mb		
	Polycom HDX 9004	Codec	€	2 Mb €	4 Mb	4 Mb		
	Polycom Viewstation Fx (fr. 2003)	SetTop	€	2 Mb €	2 Mb	-		
	Sony PCS-HQ90	Codec	€	-	8 Mb	-		
	Sony PCS-G70P	Codec	€	2 Mb €	4 Mb	-		
	Sony PCS-1 (from 2003)	SetTop	€	768 k €	2 Mb	-		
	Tandberg 900 MPX	SetTop	€	512k €	2Mb	2Mb		
	Tandberg EDGE 95MXP	Codec	€	512k €	2Mb	2Mb		
	VCON HD5000	Codec	€	-	4 Mb	-		
	VCON HD3000	SetTop	€	-	4 Mb	-		
H.323-software	DyLogic Mirial	Software	€	-	2 Mb	x	€ H.324m	x
	Ekiga	Software	free	-	x	x		
	Emblaze-VCON vPoint	Software	€	-	4 Mb	-		x
	Microsoft Netmeeting	Software	free		435 k	-		
	Polycom PVX	Software	€	-	2 Mb	x		
	VisualNexus Endpoint	Soft+srv	€	-	2 Mb	-		x
	Xenex VisualMeeting	Software	€	?	4 Mb	?		
	Xmeeting	Software	free	-	x	x		
3G	3G Phone	Phone	€	-	-	-	H.324m	
Messengers	iChat	Software	free	-	-	x	Jabber	
	Google Talk	Software	free	-	-	-	Jabber	
	Gizmo	Software	free	-	-	x	Jabber	
	Skype	Software	free	-	-	-	p	
Web-based	Acrobat Connect Pro	Soft+srv	free / €	-	-	-	p	srv
	Flashmeeting	Soft+srv	free / -	-	-	-	p	srv
	Marratech	Soft+srv	free / €	-	€ srv	x srv	p	srv
Streamers	DVTS	Software	free	-	-	-	p	
	VLC	Software	free	-	-	-	many	x

Disclaimer: This data is approximate and presented to the best knowledge of the author at the time when being written

Dual Image / Data Sharing			Audio									
H.239	T.120	Other	G.711	G.722	G.722.1	G.722.1c	G.723.1	G.728	G.729	StereoSur.	AAC	Other
x	x		x	x	x	x	-	x	-	-	LD 48-64	
x	x		x	x	x	x	-	x	-	-	LD 48-64	
x	?		x	x	-	x	-	-	x	-	LC	
x		People+Content	x	x	x	x	-	x	G.729A	x	-	
x		People+Content	x	x	x	x	-	x	G.729A	x	-	
-		People+Content	x	x	x	-	-	x	-	-	-	
?			x	x	-	-	-	x	-	-	LD? 96,192	
x			x	x	x	-	x	x	x	-	LD? 96,192	
x			x	x	x	-	x	x	x	-	?	
x		DuoVideo	x	x	x	-	-	x	-	-	LD 64,128	
x		DuoVideo	x	x	x	-	-	x	-	-	LD 64,128	
x			x	x	x	-	x	x	x	-	LD?	
x			x	x	-	-	x	x	x	-	-	
-	x		x	-	-	x	x	-	-	-	-	
-			x	-	-	-	-	-	-	-	-	iLBR
x	?	DuoVideo	x	x	x	-	x	x	x	-	LD?	
-	x		x	-	-	-	x	-	-	-	-	x
x	x	People+Content	x	x	x	de?	-	x	G.729A			
x	x		x	x	-	-	x	-	-	-	-	x
x	x	?	x	x	-	-	x	x	x	-	LD	
-							x					x
-											LD	
-												iLBR
-			x									iSAC, iLBR
-												iLBR
-		p										?
-		-										?
-		p										iLBR
-			-	-	-	-	-	-	-	-	-	p
-			x								x	many

Sources: Manufacturer and reseller websites, tech blogs and forums, Wikipedia

Video codec							Main video resolution						
H.261	H.263	H.263+	H.263++	H.264	On2 VP6	Other	SQCIF	QCIF	QVGA	CIF	VGA	4CIF	720p
x	x		x	x	-		-	x	-	x	-	x	x
x	x	x	x	x	-		-	x	-	x	-	x	de
-	x	x	-	x	-		?	?	?	?	?	?	x
x			x	x	-		?	x	?	x	?	x	-
x			x	x	-		?	x	?	x	?	x	x
x		x	x	-	-		?	x	?	x	?	x	-
-	-	-	-	x	-		?	?	?	?	?	?	x
x	x	x	x	x	-		de	x	-	x	-	x	-
x	x	x	x	x	-		-	x	-	x	-	-	-
x	x		x	x	-		de	x	-	x	?	x	x
x	x		x	x	-		de	x	-	x	x	x	x
x	x	x	x	x	-		-	x	x	x	-	-	-
					-		-	x	-	x	-	x	-
x	x	x	-	x	-		x	x	-	x	-	x	x
x	-	-	-	-	-		-	x	-	-	-	-	-
x	x	x	x	x	-		-	x	x	x	x	-	de
x	x	-	-	-	-		x	x	-	x	-	-	-
x	x	x	-	x	-		?	?	?	x	x	?	-
x	x	x	-	x	-		-	x	-	x	-	€	-
x	x	x	x	x	-			x		x	x	x	de
x	x	?	?	x	-		?	?		?			
	x							x					
-	-	-	-	x	-				x		x		
-	-	-	-	-	-		-	-	-	-	-	-	-
x	-	-	-	-	-					x			
-	-	-	-	-	x				x		x		
-	-	-	-	-	?	p	?	?					
-	-	-	-	-	?	p	?	?					
x	-	-	-	x						x			
-	-	-	-	-	-	p	-	-	-	-	-	x	-
				x		many		x	x	x	x	x	x

Specifications subject to change as new functionalities are added

Video framerate max	Encryption H243,245 DES	AES	Other	Firewall traversal H.460.18/19	Other	Optional or embedded multipoint
30	-	x				9 @ 256bps - 3 @ 1Mb, CP, H.239
30						
30	-	x		-		4 CP, 6 VS, H.239
30		x		x		6 sites (or 4 with H.264), CP/VS, H.239
30		x		x		4 or 8 sites, CP/VS, H.239, TC, max 6 Mbps
30				-		4 Sites CP/VS
60	-	x		-		4 Sites CP/VS
30	-	x		-		6 sites
30	-	-		-		6 sites
30	x	x		x		4 @ 768, CP
30	x	x		x		4 @ 768, CP/VS, TC, H.239
30	-	x		-		-
30	-	x		-		4 @ 384, 3 @ 768, VS
30	-	-		-		-
15?						-
30	€	x	x		x	-
15?			?	-		-
30	-	x				-
?						-
30	-	x	x		x	-
?			x		x	-
15?			?			
30						4 site
-						
?			p			
30					p	
15?			p		p	Server based
5?					p	Server based
30?	x		x			Server based
30	-	-		-		-
30				-		One to many, altough too much delay for VC