A Multi-Stream Tool to Support Transmission in Surgery Applied to Telemedicine

Julio Silva[1], Anderson Ferreira[1], Elenilson Vieira[1], Marcello Passos[1], Erick Melo[1], Tatiana Tavares[1], Gustavo Motta[2] and Guido Souza Filho[1]

[1] LAViD (Digital Video Applications Lab), Informatics Department, UFPB, João Pessoa, Brazil
{julio, anderson, elenilson, marcello, erick, tatiana, guido}@lavid.ufpb.br

[2] DI (Infomatics Department), UFPB, João Pessoa, Brazil
Gustavo@di.ufpb.br

Abstract. The increasing network bandwidth capacity and the diminishing costs of related services have led to a rising number of applications in the field of Information and Communication Technology. A special case is applications based on video streaming. Telemedicine can be highlighted in some scenarios for applying this technology, such as clinical sessions, second medical opinion, interactive lessons or virtual conferences. These scenarios often imply a dedicated transmission environment. A restriction in such solutions is the inability to handle multiple video streams. Thus, this paper presents a low-cost infrastructure for video collaboration in healthcare and based on open technologies. The proposed infrastructure enables remote management of simultaneous multiple streams. We also discuss results of experiments held in the Lauro Wanderley Academic Hospital, Brazil. One of the results is the contribution for teaching experiences, particularly by allowing students to remotely regard surgical procedures and providing real-time interaction. Finally, we present new prospects for using the developed technology on other applications in Telemedicine and Telepresence.

Keywords: Digital video, Management, Telemedicine, Collaboration.

1 Introduction

The domain of Information and Communication Technology (ICT) has been going through a notable transformation which is characterized by the global connectivity and the increasing use of multimedia devices. These factors have afforded the development of new transmission networks to handle large volumes of data and increasing power transmission [1], as the Internet2 [2]. High power transmission networks enable the development of applications that require a large bandwidth.

Actions directed to Telemedicine and Telehealth is growing around the world at an accelerated pace. Large technology companies like Polycom [3], Tandberg [4] and

Cisco [5] are investing heavily in these areas. Cisco, for example, presented, in 2010, the Cisco HealthPresence, at the Healthcare Information and Management Systems Society (HIMSS) Conference. Cisco HealthPresence is a new technology in advanced Telemedicine that enables remote medical appointments, with features and technologies never used before. All of this combining high-definition video and high quality audio, as well as enabling medical data transmission, which gives the patient the feeling of being in a face-to-face appointment. According to Lima [6], among the several forms of Telemedicine there are videoconferences, which allow real-time integration, by sending and receiving high-quality video and audio along geographically distant points. Thus it is essential to ensure a secure data transmission.

Therefore, this paper presents a tool consisting of a set of components that together allows the easy management of several streams, controlling everything from capture to display in an efficient and intuitive way. We present the main actions related to the development of tools for remote multimedia streams management in order to illustrate the state-of-the-art in this field. The main characteristics obtained from this study compose the basis of the architecture presented in this paper. Finally, we present results and feedback from the use of the tool in real environments.

2 Related Work

In order to characterize the state-of-the-art we highlight in this section two projects and developed tools. The analysis of related projects gives higher importance to the works developed in the institution which were indispensable to the current work. On the other hand, the analysis of the tools allows comparing this work face to other tools which implement similar features, as described in sequence.

2.1 Projects

The Digital Video Workgroup [7] had the main goal which was developing and deploying an infrastructure based on the RNP (National Network for Education and Research) network that offers support to applications involving digital video manipulation. Another responsibility of this workgroup was motivating and providing the necessary conditions to create, store and transmit digital video content around the country – Brazil – by the use of two developed tools: JDLive [8] and JDVoD [9]. The first one is a server for transmitting high and low quality digital video. The latter is a video-on-demand server.

The GingaVR [10] project aimed at developing and demonstrating a platform focused on the creation of immersive virtual reality applications over high speed networks. One challenge of this project was to allow the remote control of a robot, associated to the transmission of several synchronized high quality and real-time video streams, as well as the exhibition of these video streams in a Digital Cavern.

The Digital Media and Arts Workgroup (GTMDA) [11] focused mainly on providing a new advanced way of Human-Computer Interaction, which allows the intermingling of human and synthetic agents in shared and distributed real-time multimedia spaces, all over high speed computer networks with large volumes of data.

The Video Collaboration in Health Workgroup (GTAVCS) [12] proposed an infrastructure based on hardware and software with remote management for capturing and securely distributing multiple simultaneous streams in order to provide support for several scenarios of video collaboration in health.

2.2 Tools

There are several studies in the literature with the main goal of manipulating digital media, specially, digital video. The most relevant are described below.

DICE [13] is a tool that manipulates real-time video, by the use of filters. These filters applied to the videos are nothing more than the manipulation of pixels that, when combined mathematically, creates new images from a real-time video stream. This tool was developed only for Apple Macintosh platform and is limited to the use of filters on local real-time video streams captured from cameras plugged on the computer, i.e. it does not use any transmission system of TCP/IP protocols stack.

Grass Valley 300 [14] is a device whose main purpose is to switch multiple real-time or pre-recorded video sources. These input sources can be capture through composite cables (analogical) or via serial or parallel cables (digital). This device enables capturing a total of 64 different sources and switching them to any output by manipulating buttons and levers. Despite of being a device highly professional, it is limited to local video manipulation. Another setback is its purchase or rental cost.

SuperCollider (SC) [15] is an environment and a programming language for real-time audio synthesis and algorithmic composition. Since version 3, the environment is divided into a server and a client that communicate to each other via sockets by using the Open Sound Control (OSC) 6 (CNMAT). The SC server supports single plugins made in the C programming language, which makes it easier to develop algorithms for efficient manipulation of sounds. All external controls on the server are made via OSC. On the server side the sound generation process is via an optimized executable line command called *scsynth*, which in most cases is controlled by the SuperCollider environment, but which can also be independent. In this environment, insertion and manipulation of sounds over the network occurs in real-time using the TCP/IP protocol. Another point to be highlighted is that SuperCollider is developed under the GNU Generic Public License (GNU GPL), or just free software, which makes it a suitable environment to be added into other free software projects requiring real-time audio manipulation.

INTERACT [16] is a software used by many researchers who need to collect data in observational studies of multimedia environments. It allows the user to interact with their audio/video material, analyzing it by pressing a few keys on the computer or using the mouse. It is possible to manipulate parts of the video, allowing going into actual details of the scene. It can also integrate any external data, such as physiological information or data stored on the hard disk, offering a wide range of possibilities for visualization and analysis such as statistics, reliability analysis and sequential analysis in a time interval.

The TVSL [17] is a tool that captures video and audio content and sends them to Icecast servers. It has a graphical user interface that displays video streaming, allows the application of video effects dynamically and allows setting the parameters of the transmission. In the use of TV Free Software (TVSL), it was identified the requirement of a software that meets the task of capturing audio and video, transcoding and streaming multimedia (stream) to Icecast 2 servers. There are different methods for transmitting, using and combining programs that perform each of the above tasks. So, in this scenario, it has been developing the TVSL as the main program to setup TVSL transmission, allowing real-time capture and encoding of different video and audio inputs from external devices such as firewire cameras, USB webcams, and PCR and DVB cards.

2.3 Comparative Analysis

Table 1 presents a comparative analysis among the tools described in section 2.2. The tools were analyzed regarding their TCP/IP support, audio streaming, video streaming, statistical data generation, and if each one is a distributed system.

For example, in telemedicine is required that a tool allows the connection of remotely distributed users (distributed system), statistical data generation enabling transmission simulation, as well as that is provides TCP/IP support. These requirements were, therefore, used as the basis for developing our tool, described in the next section.

Table 1. Comparative analysis among tools

	TCP/IP Support	Audio Streaming	Video Streaming	Statistical Data Generator	Distributed System
Grass Valley	No	Yes (only locally)	Yes (only locally)	No	No
SupperCollider	Yes (only audio)	Yes	No	No	Yes
DICE	No	No	Yes (only locally)	No	No
INTERACT	No	Yes (only locally)	Yes (only locally)	Yes	No
TVSL	Yes	Yes	Yes	No	Yes

3 The Developed Tool

The main objective of our tool is to remotely manage and coordinate distributed multimedia sources with different media encoding formats. Following, we present its main developing models: conceptual, architectural, component and security models.

3.1 Conceptual Vision

The tool uses any network that gives minimum support for high definition streams (1440x1080) at rates of 25Mbps and standard definition (720x480) at rates of 4Mbps or 8Mbps depending on the compression technique used. These streams can be generated locally or in different geographic locations and they are controlled and managed by the principal component of our tool, the *Articulator*.

3.2 Architectural Vision

The architecture illustrated in Fig 1 shows the components articulator, decoder, encoder, reflector, videoserver, videoroom and webservice. It also shows the possibilities for transmission rates: high quality (HD, SD) and low quality (mainly directed to the Web).

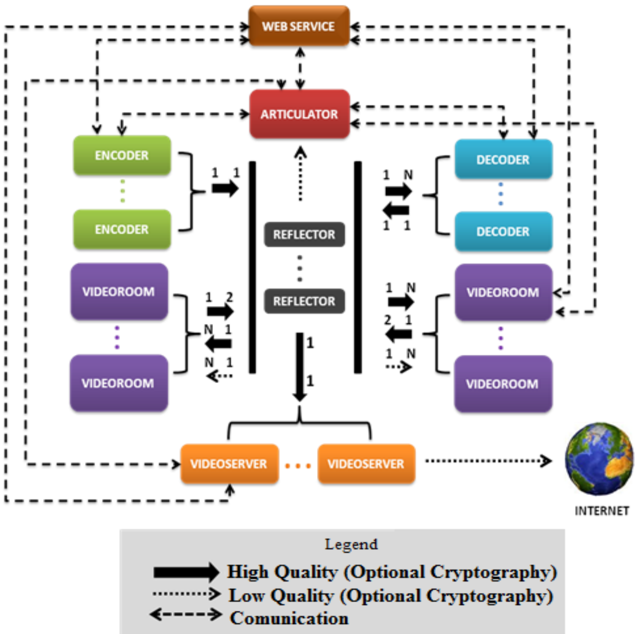

Fig 1. Architectural vision

3.3 Component Vision

It can be seen at the previous topic that the tool uses a distributed architecture based on functional components. The main ones are: encoder, decoder, videoroom, webservice, reflector and articulator.

The *encoder* is responsible for making the media source encoding, which can be either a capture device or files stored in the hard disk – AVI, WMV or TS files. It is also responsible for streaming and sending the captured media to a *reflector* that will distribute the stream to the targets set by the *articulator*.

The decoder has the main feature of capturing a single media stream and decoding it in order to display the media on an appropriate device. The capture of the stream is done via a UDP port, which is automatically combined in advance with the *articulator*.

The *videoserver* has the core functionality of streaming low resolution video to the Web in various formats, specifically popular file formats used nowadays, such as FLV, OGG and H.264, allowing a wider range of options for users viewing the video streamed to the Internet. But to broadcast the media in different formats, it is required a robust machine that can perform the original audiovisual content transcoding.

The *videoroom* encloses the functionalities of *encoder* and *decoder* components, which makes easier the simultaneous communication with multiple clients. The development of this component had the main goal to meet an easier configuration of capture and display devices in a surgery room. Spaces such as surgery rooms are usually limited. Moreover, considering infectious disease control issues in a surgical environment it is advisable to concentrate on a single device the functions of capturing and displaying media.

The *webservice* main features are (a) create/update session, (b) create/update user, (c) insert or remove user from a session, (d) finalize a session. Sessions are composed of encoders, decoders and/or videorooms. In each session, you can isolate a specific configuration of components so that you can forbid access for unauthorized users to a particular audiovisual content. Thus, a single articulator is capable of manage various audiovisual sessions. This component meets requirement of easily manage multiple independent surgeries within the same hospital, for example. Thus, the content of each session (i.e. each surgery) is restricted to authorized users, who may be Medicine students, residents or doctors within the surgery room.

The *reflector* optimizes the distribution of media streams over the network. This component works in two different scenarios: one is the direct send of stream to a *decoder* or a *videoroom*, at the same rate it received; the other scenario is characterized by transcoding the media into a lower rate, in order to send it to the *articulator*.

The *articulator* is the principal and most complex component. It is responsible for remote managing all others components, enclosing much of the functionality offered by the tool. One of its main features is the scheduling of video streams, with which you can program the hour when media streams are sent from *encoders* to *decoders*.

3.4 Security Vision

The strategy developed is based on the authentication of all video sources and destinations, as well as the transmission of encrypted media streams.

One way to address the secure transmission of multiple streams is sharing a symmetric key for each pair of users (source/destination). However, this strategy is not feasible due to the high cost of processing the encrypting the stream for each destination, since the encryption key is different for each pair of users. In order to handle this limitation, the concept of session is an alternative. A session is defined as a space created to group users and share data in a deliberative meeting. A session can be moderated or not, which implies giving permission to access any information to all users or only for guest users invited by the creator. In order to obtain the unique symmetric key of a particular session, the user needs to authenticate in the session. Authentication is accomplished through queries to a database of users located in a centralized server and the communication is done using the *Rivest-Shamir-Adleman* (RSA) algorithm 1024 bits [18].

Among the features to be protected there are the streams sent between users in the session. The streams are encrypted and decrypted using the *Advanced Encryption Standard* (AES) algorithm 128 bits [19] with shared symmetric key of the session.

4 Case Study

The use of the tool in telemedicine involves a proposal for transmitting surgeries in two academic hospitals of RUTE Network: São Paulo Academic Hospital at UNIFESP and Lauro Wanderley Academic Hospital (HULW) at UFPB. In both cases it is required remote management, media capture and secure distribution of multiple simultaneous streams (video, audio, and clinical parameters).

In the first case - São Paulo Academic Hospital, UNIFESP - the goal was transmitting a surgical procedure, namely transapical heart valve implant. The transmission used four simultaneous video streams. In such procedure, a prosthesis is used to replace the problematic valve. The prosthesis can be implanted through a minimum cut, which leads to the chest via a catheter. During the procedure, the heart continues beating and the equipment for extracorporeal circulation is not necessary. This is a high complexity procedure; therefore, students and other professionals in loco have considerable difficulty to visualize the details during the surgery. Fig 2 shows the dynamics of this procedure where we can see the large number of people in the room.

In order to deal with the problem of little free space in the surgery room and to allow a larger number of viewers, the tool was used to provide interaction between the professor/physician in the operating room and other geographically distributed participants.

Fig 2. Surgical procedure illustrating a medical lesson

At the Lauro Wanderley Academic Hospital (UFPB) the tool has been used for transmission operation. In these experiments, multiple streams were transmitted between the surgery room, where the procedure was performed, and the telemedicine room, where students and teachers could interact and follow the procedure in real time. The procedure transmitted was an inguinal hernia surgery using laparoscopy as can be seen in Fig 3.

Fig 3. Schematic view of the surgery

In this experiment, while a surgeon performed the surgery another doctor followed the procedure with their students in the telemedicine room at the HULW as can be seen in Fig 4. Doctors could interact via audio and image at any time of the surgery. Two cameras were used during the experiment: the endocamera (internal view) and another camera responsible to capture the external view of the procedure.

Fig 4. Telemedicine room at HULW during experiment

Fig 5 shows the articulator, module that manages streams captured by encoders and displayed by decoders. We can see the streams that are manipulated during surgery.

The surgery room and the telemedicine room sent and received streams, which allowed interaction among participants of the two rooms. A multimedia stream was captured in the telemedicine room and displayed in the surgery room and two multimedia streams were captured in the surgery room and displayed in the telemedicine room, but only one was displayed at a time. These flows were switched according to the need for participants to see the surgery from different angles. Researchers at LAViD followed the events of both rooms, but no stream was sent to any of the rooms.

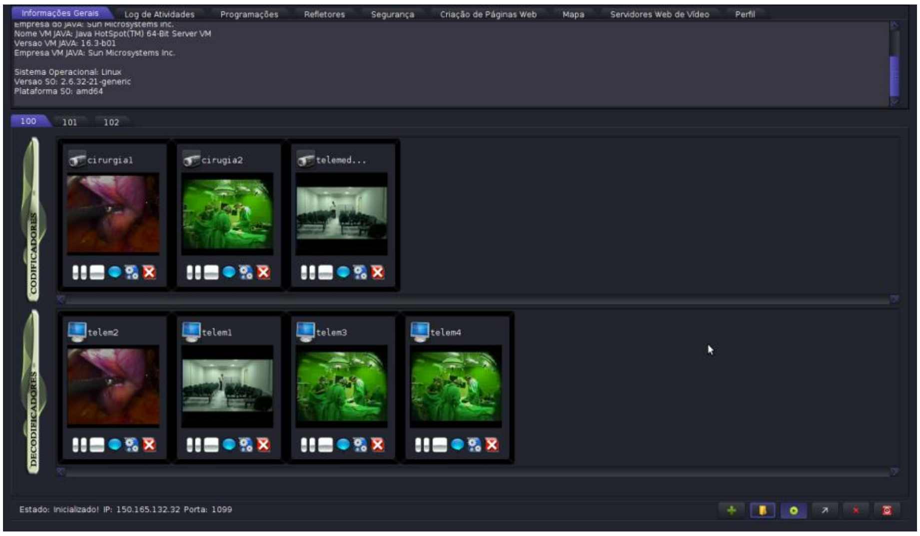

Fig 5. Articulator managing the transmission at the Lauro Wanderley Academic Hospital at UFPB

The experiments had good assessment of physicians, Medicine students and researchers from LAViD. There is an expectation for similar activities and the possibility of adopting the method as part of classes conducted in the hospital.

5 Conclusion

The use of specialized tools for managing streams in artistic or sports events, or experiments such as telemedicine or tele-education, reduces operational complexity and increases the possibilities of transmission. Therefore, the tool presented in the current paper aims to contemplate such scenarios as an integrated solution for managing distributed real-time multimedia streams. The tool provides support to coordinating different multimedia sources in order to facilitate the interaction among virtual and human agents. With only computers, cameras and/or projectors it is possible to any user starting its own transmission in a simple and low-cost way.

Acknowledgments. The current work has been made possible thanks to the financial support provided by CAPES, under the RH-TVD Program for human resources training in Digital TV. We also thank RNP (National Network for Education and Research) for funding the workgroups cited on this work, specially, GTMDA and

GTAVCS. Finally, we thank the Lauro Wanderley Academic Hospital for the indispensable partnership to the conclusion of this project.

References

1. MELO, E. A. G. et al. Arte e tecnologia: Lições aprendidas com a realização de performances artísticas baseadas na distribuição de conteúdo multimídia. In: Conferência Latino Americana de Informática. 2009. Pelotas, RS.
2. RNP (Brasil). Internet2, http://www.rnp.br/redes/internet2.html.
3. POLYCOM. Polycom Healthcare Solutions, http://www.polycom.com/solutions/industry/healthcare.html.
4. TANDBERG. TelePresence Conferencing Infrastructure, http://www.tandberg.com/video-conferencing-multipoint-control.jsp.
5. CISCO. Cisco lança HealthPresence,Tecnologia inovadora para Telemedicina, http://www.cisco.com/web/PT/press/articles/100412.html.
6. LIMA, Claudio Marcio Amaral de Oliveira et al. Videoconferências: sistematização e experiências em telemedicina. Radiol Bras, São Paulo, v. 40, n. 5, out. 2007, http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-39842007000500012&lng=pt&nrm=iso.
7. RNP (Brasil). GT Vídeo digital, http://www.rnp.br/pd/gts2002-2003/gt-vd.html.
8. SOUZA FILHO, G. L.; SILVA, Linconl David Nery e; SOUSA FILHO, Gilberto Farias de. JDLive - Dynavideo Java Live Video Distribution Server. 2004.
9. SOUZA FILHO, G. L.; BATISTA, Carlos Eduardo Coelho Freire. JDVoD - Dynavideo Java Video on Demand Distribution Server. 2004.
10. RNP (Brasil). GIGA VR – Plataformas para o Desenvolvimento de Aplicações de Realidade Virtual Imersiva e Distribuída sobre Redes de Altíssima Velocidade.
11. GTMDA - Grupo de Trabalho de Mídias Digitais e Arte, www.lavid.ufpb.br/gtmda.
12. GTAVCS – Grupo de Trabalho em Ambiente de Vídeocolaboração em Saúde, www.lavid.ufpb.br/gtavcs.
13. THOMPSON, John Henry. JHT Other work, http://www.j4u2.com/jht/newwork.html.
14. GRASS VALLEY (França). GVG 3000 User Manual, www.grassvalley.com/docs/Manuals/DigitalSwitcher/0159_00.PDF.
15. MCCARTNEY, James. SuperCollider, supercollider.sourceforge.net.
16. CENTER FOR NEW MUSIC AND AUDIO TECHNOLOGY (CNMAT). Open Sound Control: Introduction to OSC, http://opensoundcontrol.org/introduction-osc.
17. TV SOFTWARE LIVRE. Projeto SLTV, br.gnome.org/TV/ProjetoSLTV.
18. Rivest-Shamir-Adleman (RSA), http://people.csail.mit.edu/rivest/Rsapaper.pdf.
19. Advanced Encryption Standard (AES), http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf.