A PC-Based Voice-Controlled Front-End of an Endoscopic Video Server in DICOM

G. Bellaire, D. Steines, G. Graschew, A. Thiel*, J. Bernarding*, T. Tolxdorff*, P. M. Schlag

Surgical Research Unit OP 2000, Robert Rössle Klinik am Max-Delbrück-Centrum für Molekulare Medizin, Universitätsklinikum Charité, Medizinische Fakultät der Humboldt Universität zu Berlin,
*Department of Medical Informatics, University Hospital Benjamin Franklin, Free University of Berlin
Email: bellaire@rrk-berlin.de

Abstract. Results of a video-server implementation for endoscopic investigations are presented. A PC-based workplace for video-capturing equipped with a speech-controlled interface for the surgeon is described. The clip size of the video to be stored is determined by clip length, image resolution, frame rate, and compression scheme. Captured clips are converted into DICOM 3.0 and integrated into a digital patient record. Additionally, an approach is presented to analyze the motion of the endoscopic camera for future automatic video-cutting. This approach analyzes compressed MPEG-2 video sequences to detect scenes of diagnostic interest.

Keywords: Video Documentation, Video in DICOM, Video Indexing, Endoscopy, Second Opinion, Voice Control

1 Introduction

The Surgical Research Unit OP 2000 designs, develops, and implements modules of the operating room of the future [9]. One of the key modules of the concept is the telecommunication module. Via telemedicine the advantages of increasing specialization in medicine can be effectively used by integrating the necessary data and diagnostic findings and by involving different medical experts. There are external and in-house telemedical applications. External applications are often point-to-point, require security mechanisms using public networks, and include expert consultations. Internal applications are mostly based on the access to various data archives and online data sources.

This paper describes the PC-based front-end of an endoscopic application. A prototype for the endoscopic system has been constructed. It has an interface to our surgical HIS and soon it will be used to supplement endoscopic investigations in our clinic.

The use of public networks such as the Internet for transmitting patient-specific data must comply with local legislation on data security and integrity. The typical heterogeneous hardware and software infrastructure within a hospital requires a standardized data exchange format and protocol like that provided by the DICOM standard in radiology. Unfortunately, DICOM does not yet include inherent data security concepts. The goal of the MedSec project that is implemented by Benjamin Franklin University Hospital, is to evaluate the secure high-speed transfer of very large amounts of data within LAN/WAN networks linking project partners using DICOM [2, 10].

Key issues of the video server are intelligent and standardized access using a database, the DICOM protocol for data security demands, and video compression [1].

2 A Video Interface for DICOM

Our concept for a video server combining a commercial database (INFORMIX Dynamic Server) [8] with the Mallinckrodt CTN DICOM server [6] opens new perspectives for medical documentation. Second opinions from internal or external medical experts can be obtained online (live) and offline (e.g., diagnostic data). The offline, external second-opinion procedure for diagnosis can be simplified by secure Internet access.

The following service parameters (SP) are required for archiving medical video adequately:

- *Objective* SP
  patient data (name, address, age, insurance)
- *Application-dependent, subjective* SP
  finding (primary, secondary, etc.)
  technical specification of the video (e.g., image source, compression, image size, frame rate, etc.)
- *Content-based* SP
  features of images or image segments
  classification of images or image segments

To be able to exchange video data with the Mallinckrodt DICOM software, we developed a definition for a new generic DICOM object containing video data. The Mallinckrodt software was modified accordingly so that it can handle the new objects and recognize the additional attributes.

3 Video Grabbing

Video can be captured using a PC equipped with an Aspro Invader32 video digitizer adapter with hardware Motion-JPEG compression. The analog PAL output signal from the endoscope is connected to the video adapter. The SDK shipped with the adapter provides the necessary function calls for initialization and capture control.

3.1 DICOM conversion

The dialog for video capture control (see Figure 1) written in C++ is integrated into a JAVA application that provides the DICOM functionality. Patient information can be entered using the GUI or be sent to the application from the GUSTAV surgical documentation system. The video clips are recorded into .AVI files on hard disk. After concluding the examination, all recorded video clips are wrapped with a DICOM header containing necessary patient data and examination parameters. These newly created DICOM objects can afterwards be sent to the DICOM server (see Figure 2).

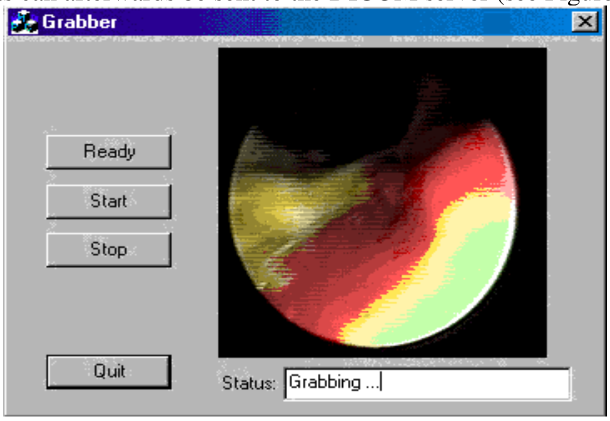

Figure 1: Speech-controlled Front-End for video-grabbing

3.2 Speech control

To achieve speech control, we implemented the script language of the SmartWord Professional 3.0 (Dragon) system using a commercial PC with sound blaster 16V audio adapter. The workplace requires 5 commands for speech control. The system has a very high recognition rate concerning false positive and false negative results. Even without an adaptation of the speaker's voice to the system, the recognition rate is tolerable. In our experience the most sensitive parameter of the speech recognition is the relative position of the microphone to the speaker's mouth. We are currently experimenting with different microphone configurations for use in the operating room (e.g., under the surgeon's mask, positioned on the operating light, etc.).

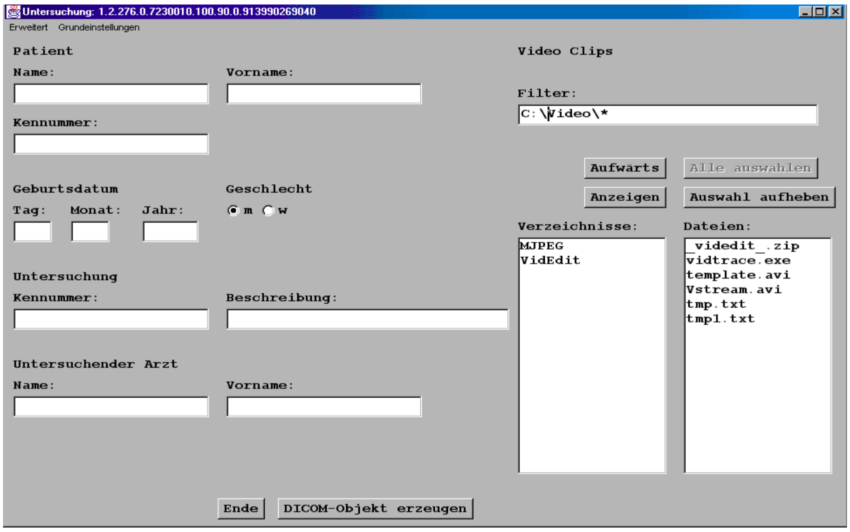

Figure 2: Front-End for DIOCM conversion.

3.3 Viewer

There are three different viewing options for the system. Our clinic uses the documentation system GUSTAV, which is based on NEXTStep™ technology. To integrate video sequences into GUSTAV, a link to a video file is generated and stored on a NFS-mounted hard disk. In GUSTAV there is a list of all surgical procedures available for a given patient, including endoscopic intervention. If a video documentation of an endoscopic intervention has been produced, a keyframe of the video is shown in the patient file. For video playing the NEXTStep™ Quicktime™ player is used. Currently doctors in our clinic have access to GUSTAV from 15 workplaces.

Additionally, a JAVA-based DICOM video client has been developed. This client can connect with a DICOM video server. A patient can be selected using his/her unique patient ID. Findings are ordered by (DICOM) studies and by the video acquisition date.

For external use, an improved JAVA video client with security facilities is being constructed that is based on developments in the MedSec project [10]. Performance tests have been conducted using SSL within DICOM to encrypt image and/or patient data transferred between Robert-Rössle-Klinik and Benjamin Franklin hospital (two Berlin university hospitals) [3].

4 Results

Table 1 shows the parameters of the system. To get an impression of the viewing performance using our 10 Mbit/s in-house network, consider the following example.

An image with a resolution of 352x288 contains 101376 Pixel. The underlying color format YCbCr 4:2:0 implies 1.5 Bytes/pixel. The resultant 152 Kbytes/frame times 25 frames/sec times an estimated length of 10 sec result in 38 Mbyte/clip. M-JPEG compression (1:10) reduces clip sizes on 3.8 Mbyte/10s video. With real 4 Mbit/s network performance, a doctor needs about 8 seconds for video access. To calculate the archive sizes we estimate 20 videos per day (two videos/patient * 10 patients)* 200 days/year * 3.8 MByte/video = 15.200 GByte/year.

Table 1.

Application	Clip length	Image resolution	Video scheme	Compression ratio	Frame /s
Endoscopy	< 10 sec	280x200	M-JPEG	1:10 to 1:15	25
Second Opinion	< 10 sec	355x288	M-JPEG	1:10 to 1:30	25

5 Indexing of Endoscopic Video

Approaches and preliminary results for the automatic video cutting of endoscopic sequences are presented in [4, 5]. The approach is based on the assumption that the doctor performing endoscopy is fixating his endoscope more or less motion free to inspect current regions of interest. After coding the video in MPEG-2, the motion information of succeeding frames is investigated to automatically detect sequences of interest directly on the compressed video (see [7]). We assume that the variations of the motion vector pattern in the P-Frames (predicted frames) give almost enough information to differentiate between areas that are ignored and intensively inspected areas. A future analysis step will be to differentiate between translatorial motion and motion due to zooming.

6 Conclusion

The system presented here represents a step toward an optimized functionality of the multi-medial patient record. The interface with the doctor has to be intuitive and adequate. Voice control seems to be one of the most promising approaches in sterile environments. The current overall cost of the underlying hard- and software components for the presented video front-end is below US$3000 (PC, M-JPEG compression board, Speech Control). In our opinion, using multi-platform viewing software and existing digital workplaces for doctors, only the addition of a suitable video archive with a medical interface is needed.

In the next step, video indexing of endoscopic video can be integrated into the workplace as soon as the cost of MPEG-2 boards is substantially lower.

Acknowledgments

MedSec is funded by the Fed. Ministry of Science, Research and Technology. The database integration is part of the INFORMIX research grant program.

References

1. G. Bellaire, D. Steines, G. Graschew, A. Thiel, J. Bernarding, T. Tolxdorff, P. M. Schlag,
   “OP 2000 and MedSeC: A concept to extend the DICOM Standard to Include Digital Stereoscopic Video Sequences”, Proc. of the BVM 98, Lehmann et al ed.1998, 234-238
2. G. Bellaire, F. Hasenbrink, G. Graschew, M. Göbel, P.M. Schlag, “The OP2000 Medical Augmented Immersive Environment (MAIE)”, Proc. of CVE´98, 33-40, Manchester,1998
3. G. Bellaire, D. Steines, G. Graschew, A. Thiel, J. Bernarding, T. Tolxdorff, P. M. Schlag, ”OP 2000 and MedSeC: Implementating a Medical Video Server Using a Data-secured DICOM Interface”, Proc. of the EuroPACS, Barcelona, pp. 205-207, Spain, 1998
4. G. Bellaire, D. Steines, G. Graschew, A. Thiel, J. Bernarding, T. Tolxdorff, P.M. Schlag, “Storage, acces and retrieval of endoscopic and laparoscopic video”, accepted for Capture, Analysis and Display of Image Sequences, SPIE EI, San Jose, Januar, 1999
5. G. Bellaire, D. Steines, G. Graschew, A. Thiel, J. Bernarding, T. Tolxdorff, P.M. Schlag, „Standardized Access, Display and Retrieval of Medical Video“, accepted for SPIE Medical Imaging, San Diego, March, 1999
6. CTN, DICOM Software Documentation,
   ftp://ftp.erl.wust1.edu/pub/dicom/software/ctn/doc
7. A. K. Elmagarmid, H. Jiang, A. A. Helal, A. Joshi, M. Ahmed, Video Database Systems: Issues Products and Applications, Kluwer, 1997
8. Informix Video Foundation DataBlade Module User's Guide, Version 1.1
9. P.M. Schlag, G. Graschew A Vision of Surgery, the concept OP2000 Langenbeck’s Arch Surg, 385:194-197, 1998
10. A. Thiel, J. Bernarding, J. Hohmann, D. Cosic, G. Bellaire, T. Tolxdorff, “Security extensions to DICOM”, Proc. of the EuroPACS, Barcelona, pp. 177-180, Spain, 1998