Synchronized Multimodal Recording System for Laparoscopic Minimally Invasive Surgeries

Helmuth Radrich[1], Nicolas Padoy[2], Ahmad Ahmadi[2], Hubertus Feussner[3], Greg Hager[4], Darius Burschka[1], and Alois Knoll[1]

[1] Technische Universität München, Dept. of Informatics, Robotics and Embedded Systems, D-85748 Garching bei München, Germany
radrich@cs.tum.edu,

[2] Technische Universität München, Dept. of Informatics, Computer Aided Medical Procedures, D-85748 Garching bei München, Germany

[3] Klinikum Rechts der Isar, Chirurgische Klinik und Poliklinik, D-81675 München, Germany

[4] Johns Hopkins University, Department of Computer Science, Computational Interaction and Robotics Laboratory, MD 21218 Baltimore, USA

Abstract. Collecting data under regular conditions inside an operating room is a complicated and time-comsuming venture. Tight safety regulations have to be accounted for, like e.g. sterility, biocompatibility of materials, interference with regular workflow, and more. For recording high quality data, the different data-streams have to be precisely synchronized. The multimodal extendable recording system for laparoscopic minimally invasive surgeries presented here was developed to simultaneously record high quality data for multiple projects, including 3D pose estimation of the laparoscopic tools and surgical workflow analysis. At present we are recording synchronized high resolution video from 4 cameras (calibrated stereo, overview, laparoscopic) with lossless compression and audio.

1 Introduction

The goal of the work presented in this paper was to extend the basis of the available surgical data for the "Surgical Skill Analysis" and "Language of Surgery" projects at the Johns Hopkins University and for the "Surgical Workflow" project at the Technische Universität München. We recorded high quality data under regular OR-conditions during laparoscopic cholecystectomies at the 'Klinikum Rechts der Isar' hospital in Munich. The recording system was developed at the Robotics and Embedded Systems Lab of the Technische Universität München.

A brief overview of the projects just mentioned is given in the following.

1.1 Surgical Skill Analysis and Language of Surgery

The Surgical Skill Analysis and Language of Surgery project [7, 8, 11] uses kinematic motion data from the daVinci (Intuitive Surgical) robotic minimally invasive surgical system for evaluating surgical skill and for the automatic detection and segmentation of surgical motions.

The methods used are reminiscent of methods used in speech processing and are applied to suturing tasks performed on skin phantoms. A set of elementary suturing gestures, called 'gestems', is defined and these gestems are then detected in the recorded motion data. The aim of this project is to help with the evaluation and training of surgeons, to provide quantitative measures of surgical proficiency and for the automatic annotation of surgical recordings.

The recoding system presented in this paper is intended to expand the basis of usable data for this project to regular laparoscopic procedures by recording data for a marker based 3D pose estimation of the laparoscopic tools.

1.2 Surgical Workflow

The Surgical Workflow project [1, 5, 9] is aimed at an automatic recovery of surgical workflow. The methods used are able to achieve this by synchronization of multidimensional state vectors from different surgeries of the same type.

The basic idea of this project is that the laparoscopic tool usage of the surgeon is strongly correlated with the surgeries workflow. The state vectors used are composed of a binary model for the tool usage. The state vectors are sampled with a frequency of 1 Hz from manually annotated recordings of the laparoscopic image and an overview video recording of the OR.

The recording system presented in this paper was used to record data during laparoscopic cholecystectomies (gall bladder removals), the same procedure that was previously used by the Surgical Workflow project. By using a unique color-marker for each laparoscopic tool, an automatic extraction of the state vectors of a surgery could be achieved in the future.

2 Operating room recording system

Since the recording system was used during regular laparoscopic surgeries, very tight security requirements had to be fulfilled. The intended later use of the data imposed minimum performance requirements on the sensors used.

2.1 Requirements

Calibrated stereo cameras: The accuracy of the 3D-pose estimation depends on the stereo-camera baseline (depth), the accuracy of the synchronization and calibration, the camera resolution and the quality of the image (signal to noise ratio) for the marker segmentation. The position of the cameras should ensure minimal occlusions of the estimation targets. Concerning the frequency, we set 15Hz as a lower bound (the kinematic data from the daVinci system was sampled at 10Hz).

OR overview camera: We wanted to record an overview of the whole OR for later reference and analysis of surgical workflow concerning interaction in between surgeons and the OR staff. For this we needed an overview camera with a wide field of view, positioned at a good spot.

Endoscopic camera: The endoscopic camera's view used by the surgeons during the procedure had to be recorded with maximum resolution and a minimum of 15Hz.

OR-environment audio: The audio recording of the ambient OR sounds was intended for extraction of special events like coagulation or HF-cutting that are indicated by high pitch sounds to the surgeon.

Data synchronization: Data synchronization is an important issue for every multiple stream recording system. Section 3 is devoted to this topic.

Minimal interference with regular OR workflow: The recording system as a whole had to be mobile and compact, was not allowed to interfere with other OR equipment and had to be fast to mount and dismount. The last point is especially important! In case of emergency situations we had to leave the OR as fast as possible. We were not allowed to make any permanent changes to the operating room.

No contact to the sterile area: This was the most important point, since our cameras can't be sterilized and the stereo-cameras recording the surgeons workspace are close to the sterile area. Under no circumstances were the cameras or any other part of the system allowed to touch the sterile area. The stereo-cameras were required to withstand very rough handling without dropping from the attachment point.

Fig. 1. Stereo-camera-rig during a surgery

2.2 Solution

Recording hardware and mounting: We used a standard PC mounted on a mobile cart for the recordings (Fig. 2). Using a single computer for the recording enables a smaller solution and easier synchronization of the data streams. For the system to be able to cope with the large data rates we used a dual-core AMD Athlon™ 64 X2 – 6000+ CPU and two Seagate Baracuda 7.200 320GB SATA300 hard-drives in a RAID 0 configuration exclusively used for the recordings.

Fig. 2. Mobile recording system cart in the operating room

For the stereo cameras and the overview camera we used three Guppy F080C FireWire color cameras by Allied Vision Technologies (AVT). They are light and small (30mm x 30mm x 48mm, 50g) and the power is supplied by FireWire, so only one cable connection is needed per camera. The laparoscopic view is recorded via an S-Video port that is available inside the OR.

We added an Texas Instruments PCI FireWire-400 controller to the recording system. Together with the onboard FireWire-400 controller we had two busses available to connect the cameras. The S-video was recorded via an Hauppauge WinTV 878/9 frame grabber. For the audio recordings we used the onboard sound card connected to a Sennheiser freePORT™ microphone.

The main component for the stereo camera rig is the 'Velbon® Super Mag Plate' (supporting more than 5kg) on which we mounted two Guppy cameras, each on top of a 'FLM Centerball CB18' ball and socket joint (supporting max. 2kg each). The resulting base-line for the stereo cameras is 25cm. The Super Mag Plate itself is attached to a clamping support (Hama Klemm- und Tischstativ 3) that we use for attaching the resulting stereo camera rig on a monitor handle in the operating room. The weight of the total solution including cameras and lenses is around 1.1kg. The stereo rig is attached to the handle in a way that makes it impossible for it to fall, even if the clamp loosens. Fig. 1 shows the stereo camera rig mounted in the OR during a surgery.

The whole recording system, including the cameras, camera attachment solution, camera calibration patterns, cables, etc. fits onto the mobile cart, making it easier to transport the whole system to the OR and back. The cables were all laid along the walls, so that they didn't make contact to the sterile area and didn't cross the OR floor. Besides the data cables to the S-video (located close to the carts position) and to the three Guppy cameras, only one power outlet connection was needed by the system. The mobile cart was located in close vicinity to the anesthetics area from where we connected to the stereo cameras. This area can not be traversed by the OR stuff anyway, because of the anesthetics equipment.

Recording software and data rates: The recording software we used was originally developed by Balazs Vagvolgyi at the Computational Interaction and Robotics Lab (CIRL) of the Johns Hopkins University. Its original use was for recordings of the daVinci systems stereo endoscope. It uses the DirectX API to record from any compatible video devices. We extended the software to record 4 video streams and one audio stream in parallel. The log-file that is also recorded contains the time-stamps of the incoming video and audio frames.

We recorded the 3 Guppy cameras at a resolution of 1024x768 (8bit - Y800) @ 15Hz. The S-video source was recorded with 720x576 (RGB24) @ 25Hz. The corresponding raw data rates are 12 MB/s (one Guppy) and 32 MB/s (S-video). The audio streams (44100Hz, 16bit, stereo) data rate is only 0.17 MB/s. The total raw data rate of the streams recorded by the system is 68 MB/s. For comparison, here are the maximum data rates of some of the system buses the recorded data has to pass: FireWire 400 (video mode) 32 MB/s; PCI-Bus (32-bit/33 MHz) 133.3 MB/sec; Hard drives (RAID 0) 62 MB/s (random write) or 143 MB/s (sequential write). The hard drives performance was measured with the 'SiSoft Sandra XII' benchmark utility, the other values are calculated manually or taken from the specifications. All data transfer operations of the system mentioned so far are performed by DMA (direct memory access), so the CPU is not directly involved.

Connecting all three Guppy cameras to one FireWire bus would excess the FireWire-400 maximum data rate. So we connected the two stereo cameras to one FireWire bus (because of automatic synchronization, see next section) and the overview camera to a second one. The total raw data-rate to be recorded is exceeding the maximum random write performance of the hard-drives. One solution would be to always keep the data hard drives empty (no fragmentation) before starting a recording, so that they are always operating in sequential write mode. Another solution (the one we used) is to compress the data. Since the recorded data was also going to be used for computer vision algorithms we didn't want to loose any information. So we used the HuffYUV lossless video compression algorithm. It was chosen, because it performed best concerning computation costs, according to the excellent technical review by Vatolin et. al. [12]. Even so, we came very close to the maximum performance of our CPU during recording. By using lossless compression on the video data, the data-rate that had to be written to the hard drive was reduced to 48 MB/s.

Data recording procedure: For recording data inside the OR with the system presented, two people were needed. One of the reasons is that when getting in and out of the OR, the cart has to be lifted over the cabling of the regular OR equipment. Another is that the camera orientations cannot be adjusted while watching the monitor of the recording system at the same time.

Basically, the recording system can be set up inside the operating room, as soon as it is confirmed that the next surgery is the one to be recorded and the operating room has been cleaned. Interference with the regular workflow is minimal, and since patient preparation takes enough time, there's no need to hurry. The stereo-camera-rig can also be mounted at this point, but orientation adjustments and cable connections can only be made later, when the stereo-cameras reached their final position (the monitor-handle on which we attach the stereo cameras is adjusted to a comfortable position for the surgeon once the laparoscopic part of the procedure starts). Un-mounting of the system and getting out of the OR is done in the same way at the end of the procedure.

What proved to be very tedious and error-prone is the stereo camera calibration that we needed to perform for the later 3D pose-estimation. The stereo cameras orientation has to be adjusted to correctly picture the surgeons workspace. Since the surgeon is changing the position of the monitor he watches at the beginning of the laparoscopic procedure, when the area is already sterile, this has to be done at the end of the operation. Also, for a good quality calibration we need to present the calibration pattern in many different position with as little motion blur as possible. We performed the calibration at the end of the procedure, during the waking up of the patient. Since we could not unmount the stereo camera rig in order not to cause a relative movement of the cameras, we presented the calibration pattern, whenever the anesthetics team was not busy for a few seconds. Some suggestion on possible improvements of this step are given in the discussion at the end of the paper.

The cameras are oriented in a way that the patients face is never visible, thus protecting the patients privacy. In case of emergency the two people recording the data can perform the basic equipment de-connecting/removal tasks in under one minute. In such a case, the stereo cameras are left mounted and are only quickly disconnected, so as not to interfere with the anesthetics team.

3 Data synchronization

The synchronization of the different data streams is done by a log-file that records the time (system-time) of the processed frames. For the Guppy cameras there are also possibilities for a very exact synchronization of the cameras. One is by an external trigger signal, a solution that was not feasible for us, because of the additional cabling needed. For Guppy cameras connected to the same FireWire-bus the synchronization is done automatically by using the bus clock pulse. We used this solution for the stereo-cameras, since good synchronization was most important here.

3.1 Problem

The time-stamp that is reported by the DirectX API is the point in time when the data-frame has been completely processed (NOT the time when it was actually recorded or when it arrives in system memory)! This is a poorly documented fact that is highly relevant for our system, since buffering of frames and small delays in writing the data onto the hard drive can lead to relative time-stamp values that significantly differ from the truth. Using a stress test software to simulate extreme conditions we found maximum deviations in relative timing of up to 0.3 seconds (5 frames @15 Hz).

3.2 Synchronization analysis

To analyze the exact timing behavior of the system we recorded ground truth timing data and compared it to the data written to the log-file. The three Guppy cameras were recording a CRT screen displaying a stopwatch (V-synced) at a refresh rate of 85Hz. The exposure-time of the cameras was set to the exact value of one CRT screen refresh cycle (11720 us) in order to get clear, unmixed images of the displayed digits. The recorded images were undistorted and normalized (projective transformation determination via control points; resizing) and the images of the single digits of the clock were extracted. The classification of the digits was performed by a simple feed-forward neural network. This data formed the ground-truth for our synchronization analysis.

The comparison of the ground-truth to the log-file data revealed that the synchronization of the two cameras connected to the same FireWire bus is indeed perfect (up to our experiments maximum precision). The maximum frame-offset of +-1 frame is easily seen in the log-file and there is no drift. It takes up to 10 frames at the beginning of the recording until the synchronization converges. The synchronization (frame offset) of the third Guppy camera relative to the other two can easily and precisely be calculated by a linear fit of the log-file data, which corrects for the times when the log-file is lagging behind (and catches up again). The maximum deviation of the relative timings of the log file data from the ground truth data using this method was found to be 0.1535 frames. There were no dropped frames, not even under extreme test-conditions that included a running stress-test.

4 Discussion

The presented recording system was used to record data during 20 laparoscopic gall-bladder removals at our partner hospital "Rechts der Isar" in Munich. During three emergencies the hardware was handled quite rough but without causing troubles to the surgical staff or to the patient. The hardware solution is small enough not to hinder the surgical staff too much and the interference with the regular surgical workflow is inside tolerable margins. The software and hardware can handle the high data rates that have to be processed.

Still, there is much room for improvements. A fixed solution inside the operating room would be highly desirable. The overview camera could easily be mounted permanently on the ORs back wall, and the stereo cameras could be permanently attached on the top of the monitor displaying the laparoscopic image to the surgeon. A rigid casing mounted on top of a pan-tilt unit would help remedy the problems mentioned with the stereo camera calibration and orientation adjustment procedure. From a computer vision point of view, the stereo camera baseline should be wider than the 25 cm used now for improving the maximum depth resolution.

References

1. Ahmadi, A., Sielhorst, T., Stauder, R., Horn, M., Feussner, H., Navab, N.: Recovery of Surgical Workflow Without Explicit Models. MICCAI 2006, LNCS 4190, 420–428 (2006)
2. Bishop, C.: Pattern Recognition and Machine Learning. Springer (2006)
3. Brown, M., Burschka, D., Hager, G.: Advances in Computational Stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (8), 993–1008 (2003)
4. Burschka, D., Corso, J., Dewan, M., Lau, W., Li, M., Lin, H., Marayong, P., Ramey, N., Hager, G., Hoffman, B., Larkin, D., Hasser, C.: Navigating inner space: 3-D assistance for minimmaly invasive surgery. Robotics and Autonomous Systems 52, 5–26, (2005)
5. Blum, T., Padoy, N., Feußner, H., Navab, N.: Workflow Mining for Visualization and Analysis of Surgeries International Journal of Computer Assisted Radiology and Surgery 3 (5), 379–386 (2008)
6. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. (second ed.) Cambridge (2003)
7. Lin, H., Shafran, I., Murphy, T., Okamura, A., Yuh, D., Hager, G.: Automatic Detection and Segmentation of Robot-Assisted Surgical Motions. MICCAI 2005, LNCS Vol. 3749 (2005)
8. Lin, H., Shafran, I., Yuh, D., Hager, G.: Towards automatic skill evaluation: Detection and segmentation of robot-assisted surgical motions. Computer Aided Surgery 11(5), 220–230 (2006)
9. Padoy, N., Blum, T., Feussner, H., Berger, M., Navab, N.: On-line Recognition of Surgical Activity for Monitoring in the Operating Room. Proceedings of the 20th Conference on Innovative Applications of Artificial Intelligence, 1718–1724 (2008)
10. Radrich, H.: Vision-based motion monitoring through data fusion from a chirurgical multi-camera recording system. Diploma Thesis, Technische Universität München (2008)
11. Reiley, C., Lin, H., Varadarajan, B., Vagvolgyi, B., Khudanpur, S., Yuh, D., Hager, G.: Automatic recognition of surgical motions using statistical modeling for capturing variability. Medicine Meets Virtual Reality (MMVR 2008)
12. Vatolin, D., Seleznev, I., Smirnov, M.: MSU Lossless Codecs Comparison '2007. Technical Report, CS MSU Graphics and Media Lab (2007)