

Motion Tracking for Minimally Invasive Robotic Surgery

Martin Groeger, Klaus Arbter and Gerd Hirzinger
*Institute of Robotics and Mechatronics, German Aerospace Center
 Germany*

1. Introduction

Minimally invasive surgery is a modern surgical technique in which the instruments are inserted into the patient through small incisions. An endoscopic camera provides the view to the site of surgery inside the patient. While the patient benefits from strongly reduced tissue traumatisation, the surgeon has to cope with a number of disadvantages. These drawbacks arise from the fact that, in contrast to open surgery, direct contact and view to the field of surgery are lost in minimally invasive scenarios. A sophisticated robotic system can compensate for the increased demands posed to the surgeon and provide assistance for the complicated tasks.

To enable the robotic system to provide particular assistance by partly autonomous tasks e.g. by guiding the surgeon to a preoperatively planned situs or by moving the camera along the changing focus of surgery, the knowledge of intraoperative changes inside the patient becomes important.

Two main types of targets can be identified in endoscopic video images, which are instruments and organs. Depending on these types different strategies for motion tracking become advantageous.

Tracking of image motion from endoscopic video images can be based solely on structure information provided by the object itself or can involve artificial landmarks to aid the tracking process. In the first case, the use of natural landmarks refers to the fact that the genuine structure of the target is used to find reference positions which can be tracked. This can involve intensity or feature based tracking strategies. In the second case of artificial landmarks, markers with a special geometry or colour can be used. This enables particular tracking strategies, making use of the distinctive property of these markers.

This chapter describes different motion tracking strategies used to accomplish the task of motion detection in minimally invasive surgical environments. Two example scenarios are provided for which two different motion tracking strategies have been successfully implemented. Both are partly autonomous task scenarios, providing automated camera guidance for laparoscopic surgery and motion compensation of the beating heart.

2. Motion tracking and visual servoing

Visual motion tracking is dealt with here, i.e. tracking of motion from video images. This enables the use of the video endoscope for tracking, as used in minimally invasive surgery

Source: Medical Robotics, Book edited by Vanja Bozovic, ISBN 978-3-902613-18-9, pp.526, I-Tech Education and Publishing, Vienna, Austria

(MIS). Other tracking strategies with special markers and sensors, e.g. optical tracking (e.g. by ARTrack) or magnetic tracking (e.g. by NDI), which they are hard to be applied in minimally invasive surgery, are not covered.

2.1 Motion tracking

Visual tracking deals with objects of varying positions in a sequence of images. The challenge is to determine the image configuration of the target region of an object as it moves through the field of view of a camera (Hager & Belhumeur, 1998). The task of visual tracking is to solve the temporal correspondence problem, which is to match target regions on successive frames of an image stream.

Tracking involves particular difficulties due to variability in the following parameters:

1. Target pose and deformation: the object can change its position and orientation, and its image can also be deformed, eg. when viewed from different perspectives.
2. Illumination: pixel intensities may change significantly as the scene or parts of it are exposed to different lighting conditions.
3. Partial or full occlusion: the object may vanish from the scene or be partially occluded by other objects.

Tracking strategies Two different tracking strategies can be distinguished: tracking based on image features and tracking of complete regions or patterns in an image. Feature-based tracking requires the extraction of features, which yields robustness against changes of global illumination. But image features may be sparse, which requires additional constraints for the tracking process (Hager & Belhumeur, 1998). While region-based tracking saves the cost of feature extraction, it is burdened with a relatively high computational expense to find the best matching pattern in subsequent images. Direct operation on image intensities requires illumination compensation but has the advantage of using all intensity information available.

Tracking targets The target of tracking, to be detected and followed in a sequence of images, can be a particular image pattern of the object of interest with a distinctive structure. This distinctive structure implies a sufficient contrast in intensity and uniqueness to avoid losing the target in favor of a similar object in the image. Since these criteria may be difficult to fulfill in some environments, it can be advantageous to aid tracking by the use of so-called *artificial landmarks*. These artificial landmarks are designed with a unique and distinctive structure or colour and are put on the object to be tracked. While this kind of tracking is often referred as being based on artificial landmarks, the other case, in which no additional markers are placed on the object of interest to aid the tracking process can be denoted as tracking based on *natural landmarks*. In this way, natural landmarks refer to prominent parts of the target object in the image. The use of natural landmarks is especially attractive when objects such as organ surfaces are tracked, where artificial landmarks would be difficult to fix. Artificial landmarks often involve a tracking approach, in which image features are extracted which relate to these landmarks. For the case of natural landmarks, the choice between a region- and a feature-based tracking strategy depends on the property of the scene and the target object.

2.2 Tracking of surgical instruments (rigid objects)

In principle, tracking surgical instruments seems much easier than of deformable objects such as organ surfaces, since the tracking targets are rigid. The rigidity property combined

with the fact that the geometry of the objects is known enables the use of a predefined target model. Also, the application of artificial landmarks is much easier, as e.g. colour markers as in (Wei et al., 1997) or (Tonet et al., 2007). However, in the case of surgical instruments with direct contact to human tissue, particular medical requirements such as the biocompatibility and the sterilisability of the artificial markers have to be met (Wintermantel & Ha, 2001).

Most approaches for instrument tracking can be categorised into the two main classes of colour-based strategies and approaches without colour which mainly rely on a geometric model of the instrument.

The use of *colour markers* is particularly attractive, if the environment occupies only a limited range of colour, as is the case for the situs in laparoscopic surgery, which makes the design of a unique colour marker possible (Wei et al., 1997). Similarly, in a more recent publication (Tonet et al., 2007), a colour strip at the distal part of the instrument shaft is used to facilitate segmentation for the localisation of endoscopic instruments. As shown in (Wei et al., 1997) the use of an appropriate colour marker can yield a robust solution for the tracking of surgical instruments.

The approach in (Doignon et al., 2006) does without the aid of artificial markers but uses region-based colour segmentation to distinguish the achromatic surgical instrument from the image background (Doignon et al., 2004) to initiate the search for region seeds. Based on this a special pose algorithm for cylindrically shaped instruments is used to localise the instrument, which can be regarded as the second class of model-based approaches.

Doing without the aid of colour information leads to approaches which base their tracking strategy on the geometry of the instrument. These approaches often involve the extraction of edge images of the scene including the instrument, as shown in Fig. 2. As this example shows, this brings along a lot of difficulties to distinguish the instrument from its surroundings. Therefore, these approaches tend to be time consuming and prone to errors, which means that robustness is hard to achieve. A common strategy to detect the instrument without the aid of colour is to use the Hough transform, e.g. in (Voros et al., 2006).

2.3 Tracking of organs (deformable objects)

The particular difficulty with tracking the motion of deformable objects arises from that fact that, in contrast to rigid objects such as surgical instruments, the shape of the object itself changes. Moreover, in the case of organs, an appropriate and precise motion model is hard to estimate and is nonlinear in general (McInerney & Terzopoulos, 1996). Tracking deformable objects often involves the estimation of deformation in a particular image area, e.g. to extract face motions (Black & Yacoob, 1995) or to track surfaces in volume data sets of the beating heart (Bardinet et al., 1996; McInerney & Terzopoulos, 1995).

However, if the temporal resolution of the image stream is sufficiently high, such that changes between two subsequent images are small, approximating the deformation by a rigid motion model (consisting of e.g. translation and rotation) is often sufficient, as investigated in (Shi & Tomasi, 1994). This enables local structures of deformable objects to be tracked efficiently.

Fixing artificial markers to deformable objects is difficult, in particular in the case of organ surfaces. Therefore, tracking approaches based on natural landmarks are advantageous, which often involve a region-based strategy.

A region-based approach designed to enable robust motion tracking of the beating heart surface using natural landmarks (Groger et al., 2002) is described in more detail below in a scenario to compensate motion of the beating heart by a robotic system (5).

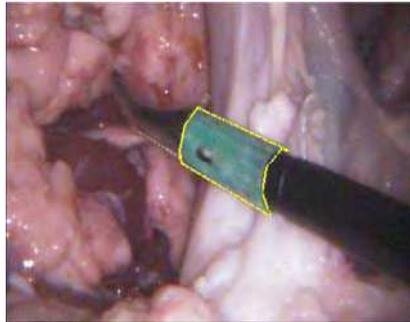


Figure 1. Laparoscopic instrument with colour marker (DLR)

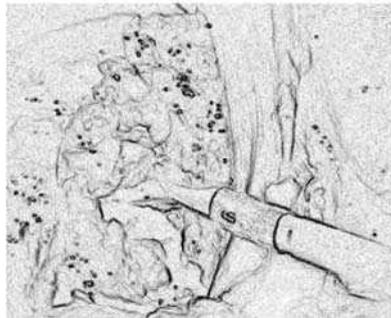


Figure 2. Edge image of laparoscopic instrument with colour marker

2.4 Visual servoing

"Visual servoing" in robotics denotes the control of an end effector in a control loop closed by imaging sensors. This requires the estimation and tracking of position and orientation of objects in the three-dimensional space, based on camera images (Corke, 1993; Hutchinson et al., 1996). Visual servoing involves the use of methods from realtime image processing, from visual tracking, and from robot control theory.

Many existing systems for visual servoing are based on artificial landmarks, which are mounted to the object of interest. However, this often increases the effort to set up the system or is hard to achieve, as e.g. with tracking of deformable objects such as organs. A region-based approach, which does not need any particular kind landmarks is described in (Hager & Belhumeur, 1998). This tracking system is successfully applied in a system for robust hand-eye coordination based on images of a stereo camera (Hager, 1997).

The use of stereo imaging from a stereo endoscope enables to estimate the three-dimensional position of the target, which is necessary for visual servoing tasks in 3D environments.

Two visual servoing scenarios for minimally invasive surgery are presented below. The automated la-paroscope guidance system enables a robot to automatically adjust the camera position to the current field of surgery (section 4). It is based on tracking a rigid object (surgical instrument) with aid of an artificial landmark (colour marker) mounted on it. The second scenario of motion compensation of the beating heart applies a region-based strategy with natural landmarks to track the motion of a deformable surface (the heart). Robust

tracking of the heart, combined with a sophisticated robotic system enables to compensate the motion of the beating heart during surgery (section 5).

3. Minimally invasive robotic surgery

3.1 Minimally invasive surgery

Minimally invasive surgery only requires small incisions into the patient body. These incisions are used to introduce endoscopic instruments into the patient body, and also to insert an endoscopic camera, which provides a view of the site of surgery inside the patient. In contrast to open surgery, minimally invasive surgery minimises trauma for the patient, decreases the loss of blood, speeds up patient convalescence, and reduces the time of the patient in the hospital.

While minimally invasive surgery brings along clear benefits for the patient, the surgeon is faced with strongly increased demands, especially since direct contact to the field of surgery is lost. A sophisticated robotic system can compensate for the increased demands posed to the surgeon and provide assistance for the complicated tasks.

3.2 Robotic support for invasive robotic surgery

Surgical robots have been developed for a variety of specific applications, as summarised in (Taylor & Stoianovici, 2003). Most early first uses of robots in surgery occurred in neurosurgery (Y. S. Kwok & et al., 1988), but the field soon expanded to other disciplines such as orthopaedics (Taylor et al., 1989,1994; Kazanzides et al., 1995) and laparoscopy (Sackier & Wang, 1996).

The use of robots allows to increase the accuracy of surgical interventions, as shown by early robotic systems in neurosurgery (Y. S. Kwok & et al., 1988) and orthopaedics (Mittelstadt et al., 1996; Bargar et al., 1998). In minimally invasive surgery the drawbacks caused by loss of direct access to the field of surgery can be compensated by the aid of robotic systems, combined with techniques from the field of telepresence. Cartesian control, e.g., overcomes the so-called "chopstick effect" when performing surgery through small incisions (Ortmaier & Hirzinger, 2000). Combined with increased dexterity of specially designed instruments (Rubier et al., 2005) this enables the surgeon to lead the instruments similar to in open surgery and to regain the dexterity as in open surgery. Force feedback (Preusche et al., 2001) together with specially designed sensorised instruments (Kubler et al., 2005) enables the surgeon to feel forces occurring at the tip of the instrument during surgery. Moreover, the use of stereo endoscopes enable a three-dimensional view to the field of surgery and as in open surgery. Different techniques of 3D display devices, such as head-mounted displays (HMDs) or a stereo console as in the daVinci system (Guthart & Salisbury, 2000).

The combination of preoperative planning with the surgical intervention enables intraoperative support for the surgeon by medical robots, such as the guidance of instruments (Ortmaier et al., 2001).

3.3 Visual servoing for robots in medicine

Visual servoing closes the control loop between imaging sensors and robot control. It enables to perform partly autonomous tasks, depending on the current situation in the field of surgery.

Examples for autonomous robot functions are tasks to supervise the working room or the automated guidance of the camera in laparoscopy (Wei et al., 1997). Different scenarios, in particular from soft tissue surgery are the guidance of the robot end effector to particular positions, in relation to given tissue structure, e.g. to hold a lighting source or tissue parts, or the autonomous movement of the robot end effector to particular positions, as e.g. in liver biopsy.

Moreover, one can think of compensating the motions of organs, such that the relative configuration and distance between instrument and organ surface remains constant. Thus, the organ is stabilised virtually. In this case, however, it is also necessary to integrate the video image provided to the surgery into the motion compensation procedure and to maintain overall consistence of motion compensation.



Figure 3. ZEUS robotic system by Computer Motion Inc



Figure 4. DaVinci robotic system by Intuitive Surgical Inc

3.4 Robotic systems for minimally invasive surgery

Many robotic systems that have been applied to surgery are based on industrial robots, as e.g. the *Robodoc* system for hip surgery (Kazanzides et al., 1995; Taylor et al., 1994; Bargar et al., 1998), and are therefore large, heavy and hardly flexible. Other robotic systems, specially designed to be applied for surgery, such as the *ZEUS* system ((Sackier & Wang,

1996), Fig. 3) by Computer Motion Inc. (Goleta, CA, USA; now: Intuitive Surgical Inc.) and the *daVinci* system ((Guthart & Salisbury, 2000), Fig. 4) by Intuitive Surgical Inc. (Cupertino, CA, USA) are much more flexible and light-weight. These systems are sufficient for laparoscopic assistance tasks such as the automated guidance of a laparoscope to provide the surgeon with a view of operating field. This is shown in the first example scenario below (section 4).

These robotic systems, however, lack the high degree of precision needed for orthopaedic surgery and additionally the high dynamics required for following the motion of the beating heart (section 5). The newly developed KineMedic surgical robot was specially designed to account for these increased demands, providing both light-weight and flexibility and the required high dynamics and precision.

The design of the KineMedic robot is based on the method of soft robotics pursued at the Institute of Robotics and Mechatronics, DLR, which leads to robotic systems such as the DLR light-weight robot (Hirzinger et al., 2001), which are light-weight, flexible and modular, and still maintain a high degree of dynamics and accuracy. Based on these techniques, the newly designed KineMedic robot has been developed as a joint partnership of DLR and BrainLab AG (Heimstetten, Germany), focussing on the demands of surgery (Ortmaier et al., 2006).

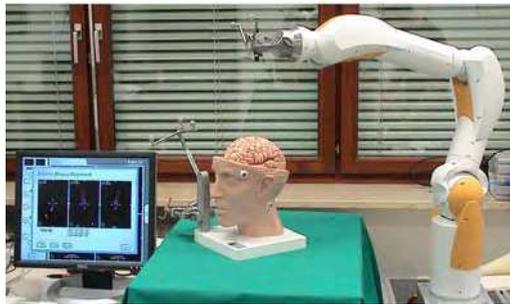


Figure 5. DLR-KineMedic medical robot arm

The KineMedic surgical robot (see prototype in Fig. 5) consists of sophisticated light-weight robotic arms, which reach a payload of 3 kg at a dead weight of only 10 kg. The redundant design of the robotic arm with seven joints enables, using null-space motion, to reconfigure the position of the robot, while the position and orientation of the instrument remains in the same position. With force-torque sensors, implemented in addition to the redundant design of the robot, the reconfiguration of the position can be performed in an intuitive way by touching and pushing the robot into the desired direction. Furthermore, the redundancy can be used to implement an arm control system which avoids collisions, which enables a more flexible setup in the operating room. Since the robot is built in light-weight design, it can be mounted or removed easily by a surgeon or nurse during a surgical intervention. This reduces mounting times in the operating room. For minimally invasive surgery usually two of these robot arms are used to manipulate surgical instruments, while a third arm moves the endoscope. An example of such a scenario is presented in Fig. 6. The KineMedic robot arm is controlled at a rate of 3 Hz and has a high relative positioning accuracy. This way, the robot provides the dynamics required for following the motion of the beating heart.

The new KineMedic robot shows significant improvements on the medical robots available so far and enables highly demanding scenarios such as compensation the motion of the beating heart.

4. Automated laparoscope guidance

In laparoscopic surgery, the surgeon no longer has direct visual control of the operation area, and a camera assistant who maneuvers the laparoscope is necessary. Problems of cooperation between the two individuals naturally arise, and a robotic assistant which automatically controls the laparoscope can offer a highly reliable alternative to this situation. In this section a autonomous laparoscopic guidance system for laparoscopic surgery is described, developed at the DLR's robotics lab and thoroughly tested at MRIC (Department of Surgery at the Klinikum rechts der Isar (MRIC) of the Technical University of Munich) (Wei et al., 1997; Omote et al., 1999).

A robot holds the laparoscope and directs it to the operative field by means of image processing techniques. The method is based on colour coded instruments. The system originally operated at a maximum rate of 17 Hz for stereo-laparoscopes and 34 Hz for mono-laparoscopes (Wei et al., 1997). It now easily runs on a standard PC in realtime for stereo-laparoscopic images delivered at a framerate of 25 Hz. For mono-laparoscopes, tracking only in lateral directions (left/right and up/down) is enabled, but for stereo-laparoscopes tracking in the longitudinal direction (in/out), too.

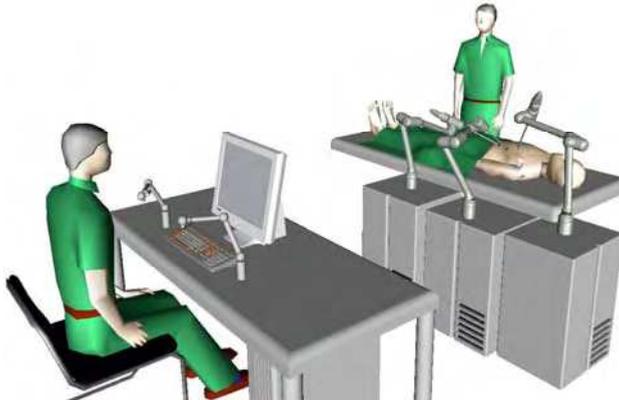


Figure 6. DLR scenario for minimally invasive robotic surgery

During the initial period of clinical evaluation 20 laparoscopic cholecystectomies have been performed and compared with those using human camera control. The longer set-up time was finally compensated by a shorter operation time. The frequency of camera correction caused by the surgeon as well as the frequency of lens cleaning was much less than with human control. The smoothness of motion was much better with the robot than with human assistants. Subjective assessments by the surgeon revealed that the robot performed better than the human assistant in a significant majority of cases.

4.1 Introduction

Laparoscopic surgery is minimally invasive, which offers the advantages of reduced pain, shorter hospital stay, and quicker convalescence for the patients. Unlike open surgery, laparoscopic surgery needs only several small incisions in the abdominal wall to introduce instruments such as scalpels, scissors, and a laparoscopic camera, such that the surgeon can operate by just looking at the camera images displayed on a monitor screen. While in open

surgery vision and action are centered on the surgeon, he loses direct visual control in laparoscopic surgery. Another person, the camera assistant, has to point the laparoscope to the desired field of vision. The surgeon has to give instructions as to where the scope should be focused, and the camera assistant has to follow them. This naturally gives rise to problems of cooperation between the surgeon and the camera assistant. A certain amount of the assistant's experience and a mutual surgeon-assistant understanding are necessary, but usually difficult to obtain. The surgeon frequently has to give the commands to move the laparoscope onto the desired area of view. This gives him an additional task, detracting his attention from his main area of concentration. The laparoscopic image may become unstable in a long operation due to fatigue of the camera assistant.

To deal with these problems, several robotic assistance systems have been developed (Hurteau et al., 1994), (Taylor et al., 1995), (Sackier & Wang, 1996) to provide more precise positioning and stable images. For a more comprehensive review of robotic systems in other surgeries see (Taylor et al., 1994), (Troccaz, 1994), (Moran, 1993), and (Taylor & Stoianovici, 2003). Investigations indicate that the use of robots in surgery reduces personnel costs while almost maintaining the same operation time (Turner, 1995). A surgical robot may be controlled either by an assistant using a remote controller or by the surgeon himself using a foot pedal (Computer Motion Inc., 1994). Voice control seemed to be another attractive alternative, as it was available with the AESOP3000 medical robot arm (by Computer Motion Inc., Goleta CA, USA).

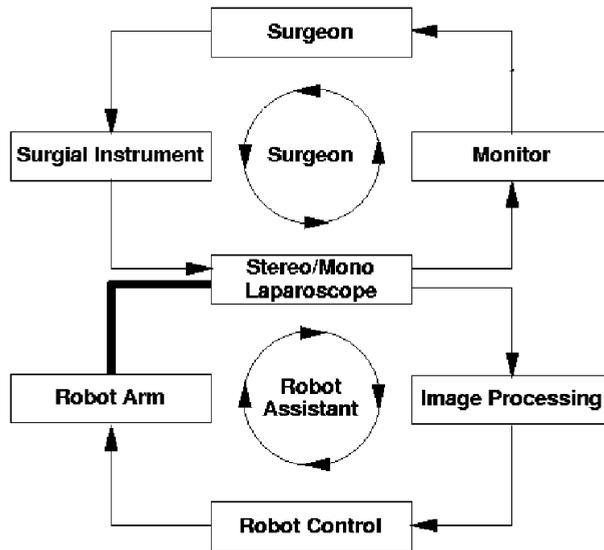


Figure 7. The structure of image-based robot-assisted minimally invasive surgery

To avoid the need for another assistant and to free the surgeon from the control task, an autonomous system that automatically services the laparoscope is highly desirable. The basic structure of an image based system is shown in Fig. 7. The surgeon handles the surgical instruments dependent on his observations on a monitor where the laparoscopic image is displayed, as usual in minimally invasive surgery. But instead of a human, the laparoscope is held by a robot arm which is controlled via an image processing system in order to track

the surgical instrument smoothly. The use of the laparoscope as a sensor for the tracking system sounds attractive, because no extra sensor is needed, but it's hard to obtain reliable control signals under realistic clinical conditions and safety requirements. The dominant problems for image processing are ambiguous image structures, occlusions by blood, organs or other instruments, smoke caused by electro-dissection, and the need of (quasi-)real-time image processing. With respect to the slow robot motions during a surgery, an image processing rate of about ten frames per second (10 Hz) is considered to be enough, but is a lower limit to maintain the impression of smooth motion.

Several researchers have tried to use image processing techniques to track the instrument such that it is always centered in the visual image. Lee, *et al.*, (Lee *et al.*, 1994) used the colour signatures of the image to segment the instrument. Since the instrument and background often possess the same colour components, much post-processing, such as shape analysis, has to be done to remove false segmentations and to extract the position of the instrument in the image. No real-time implementation was reported in, (Lee *et al.*, 1994), and it is not known whether the complexities of their shape analysis may allow implementations applicable to surgical operations. Casals, *et al.* (Casals *et al.*, 1995), used patterned marks on the instrument to facilitate image segmentation by searching for the presumed structure in the contour image. The method was reported to operate at a rate of 5 Hz for a mono-laparoscope using customised image processing hardware. Since both the methods in (Lee *et al.*, 1994) and (Casals *et al.*, 1995) rely on the existence of a preassumed shape or structure, they may fail if the camera is too near to the instrument, or if the instrument is partially occluded by organs or contaminated by blood. In both cases, the preassumed shape may not be present. Taylor, *et al.* (Taylor *et al.*, 1995), used multi-resolution image correlation to track an anatomical structure specified by the surgeon with an instrument-mounted joystick that controlled a cursor on the video display. A problem with this method might be that the anatomical structure deforms and may completely change its appearance due to manipulation of the organs.

We propose a visual laparoscope-tracking method which is simple and robust (Arbter & Wei, 1996), (Arbter & Wei, 1998). The laparoscope may be a mono-laparoscope or stereo-laparoscope. A mono-laparoscope enables the robot to track the instrument in the lateral directions left/right and up/down, while a stereo-laparoscope provides depth information and can be used to control the distance between the tip of the laparoscope and the instrument. Due to the multiplicity of problems with shape analysis, we do not check for the presence of any particular shape or structure. Instead, we use colour information alone for instrument segmentation. The non-uniqueness of the instrument colour inspires us to use an artificial colour-marker to distinguish the instrument (Fig. 12a). To mark the instrument, the colour distribution of typical laparoscopic images is analysed and a colour is chosen which does not appear in the operational field (the abdomen here). With colour image segmentation, the marker can be correctly located in the image and used to control the robot motion. Thus, even if only a very small part of the marker is visible, reliable data can still be obtained for robot control.

To build up an experimental system, only commercially available hardware was used, with the instruments from Bausch Inc., Munich, Germany, the stereo-laparoscope system from Laser Optic Systems Inc., Mainz, Germany, the AESOP 1000 robot from Computer Motion, Goleta, USA, the MaxVideo 200 image processing system from Datacube Inc., USA, and a M68040/25MHz host-CPU from Motorola Inc., USA. The coloured markers have been

placed on the instruments by the manufacturer. The electronic components are integrated into an electronic radiation protecting cabinet being mobile and used as transportation car for the robot arm, too.



Figure 8. DLR automated camera guidance scenario with AESOP robot

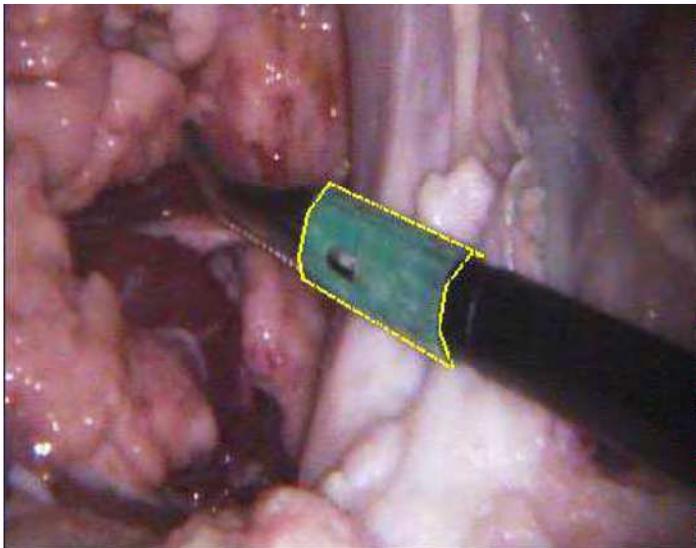


Figure 9. Laparoscopic instrument with colour marker (DLR)

For the initial period of clinical tests the system was evaluated in 20 laparoscopic cholecystectomies and compared with those using human camera control.

In the following, the image processing module, the robot controller module, and the experimental results are presented.

4.2 Image Processing

Figure 10 shows a block diagram of the image processing module. The inputs are the analog video *RGB-signals*, either from the CCD-camera pair of the stereo-laparoscope, or from only one CCD-camera of a mono-laparoscope. The inputs are **time multiplexed** at the video frame rate of 25 Hz (CCIR) in order to spend only one image processing hardware to process the stereo-images. Then the analog signals are converted into digital *RGB-signals* of 8 bits each. The *RGB* data stream is **converted** to the *HSV* format (Hue, Saturation, Value) for reasons explained below. The **classifier** separates two classes of pixels, those having the colour of the marker and those not. The result is a binary image containing the object separated from the background. The classifier is the kernel of the image processing module and will be explained in detail below. The **localiser** computes the bounding box and the centre of gravity of the object pixels as well as their number (size of region). The bounding box is then used to define the region of interest (ROI) for segmentation of the next frame. The use of an ROI speeds up the segmentation procedure and improves the robustness against misclassification. Figure 11 shows a stereo-laparoscopic image (of an experimental environment) superimposed by the centres of gravity and bounding boxes of the segmented marker.

4.2.1 Colour representation

A colour can be represented by its red, green, and blue components (*RGB*). In digital 8bit-images, the *RGB* values are between 0 and 255. Thus colours can be represented by the points within the *RGB* cube of size 256 x 256 x 256. The *RGB* colour space can be transformed to another colour space, the *HSV* colour space (Hue, Saturation, Value) where only two components *H* and *S* are directly related to the intrinsic colour and the remaining component *V* to the intensity. Different *RGB-to-HSV* transformations are known in video technology and computer graphics. We have used the following one (Foley et al., 1990):

$$r = R/R_{\max}, \quad g = G/G_{\max}, \quad b = B/B_{\max} \quad (1)$$

$$\min = \text{Minimum}\{r, g, b\} \quad (2)$$

$$\max = \text{Maximum}\{r, g, b\} \quad (3)$$

$$\Delta = \max - \min \quad (4)$$

$$H = \begin{cases} 60^\circ (g - b)/\Delta & : \max = r \\ 60^\circ (b - r)/\Delta + 120^\circ & : \max = g \\ 60^\circ (r - g)/\Delta + 240^\circ & : \max = b \end{cases} \quad (5)$$

$$S = (\max - \min)/\Delta \quad (6)$$

$$V = \max \quad (7)$$

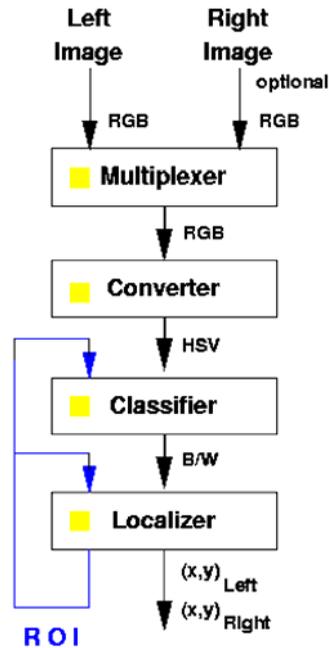


Figure 10. Structure of the image processing module

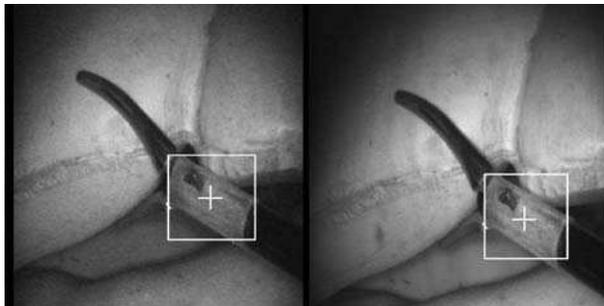


Figure 11. A stereo laparoscopic image superimposed by the centres of gravity and bounding boxes of the segmented marker

In laparoscopic surgery, we would like the image segmentation results to be insensitive to the strength of illumination. The H and S are insensitive to the strength of illumination, if only one light source, having a certain colour temperature, is used, as is the case in laparoscopy. One advantage of the $H S$ colour space is its 2-dimensionality in contrast to the 3-dimensionality of the RGB colour space, so that the colour signature of a colour image can be directly analyzed in the $H S$ plane. Figure 12d shows a colour space of the $H S$ representation, filled with the corresponding colour, where the brightness is set to 255. In this coordinate system, the H value is defined as the angle from the axis of red colour, and the S value (normalised to the range of zero to one) is the length from the origin at the centre.

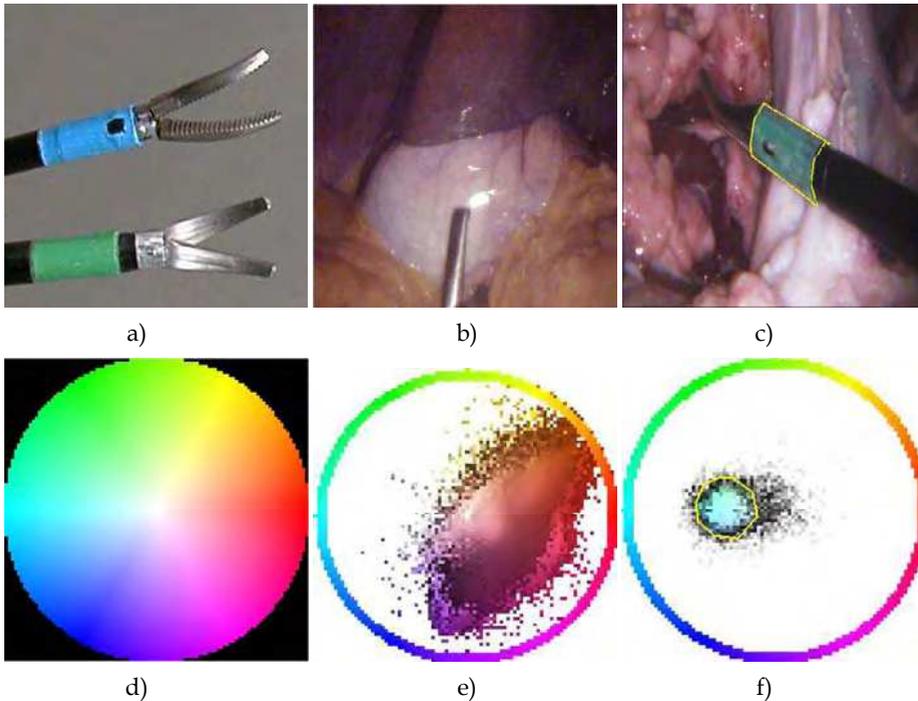


Figure 12. (a) Distal ends of colour coded minimally invasive surgical instruments (b) A typical laparoscopic image (c) Marker image with polygonal boundary (d) HS colour space (e) Colour Histogram of the abdominal scene (f) Colour histogram of the marker superimposed by the polygonal classifier boundary (cluster-polygon)

4.2.2 Marker colour selection

To choose the colour to be brought onto the instrument, we analyzed the colour components of real laparoscopic images recorded on a video tape. Typical abdominal images containing variations of colours are manually selected. Figure 12b shows one of the 17 images used in our colour analysis. An array of counters in a quantised HS domain is set to zero at the start. Then, for each pixel in the images, we compute its HS values and increment the counter by 1 at the corresponding HS position. The result is a 3-D histogram, which indicates the frequency of occurrence of all the colours in the analyzed images. To give an intuitive perception of the histogram, we display it in a colour image format, with the brightness (V) set proportional to the frequency of occurrence and the HS values equal to the HS coordinates in the HS plane. Figure 12e shows such a histogram, where the ring near the image boundary is used to help perceive the overall colour distribution. The crescent bright region within the ring represents the colours that do not appear in the images and can thus be used as the colour to be marked on the instrument. For the marked colour to be optimally distinguishable from those present in the image, the colours near the cyan are preferred, as can be seen in Fig. 12e. After the admissible colours have been determined, we have to consider the material which carries the desired colour, its commercial availability, and its

biocompatibility. On account of these factors, we have used a near-cyan plastic ring, as shown in Fig. 12c.

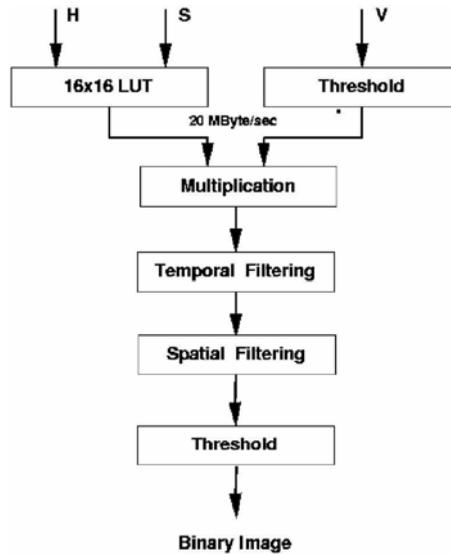


Figure 13. Classifier structure

4.2.3 Colour training selection

Due to the colour distortion through laparoscopes, we have to locate the actual position of the chosen colour in the HS plane. We select a set of typical images showing the marker in different situations, e.g., near, far, slanted, or orthogonal to the view direction, and calculate the colour histograms from the marker regions only. We first manually outline the marked instrument in the image with a polygonal boundary as shown in Fig. 12c. Then, the pixels within the polygon are used to compute the colour distribution in the HS plane. Figure 12f shows the colour cluster of the marker.

To represent the corresponding individual marker colour space, we again use a polygonal approximation of the cluster boundary Fig. 12f. By backprojection of the enclosed colours to the original training set and by modifying the boundary, we iteratively minimise the number of misclassified background pixels by simultaneously maximizing the number of correctly classified marker pixels. We repeat this procedure for all the images out of the training set, resulting in a set of colour regions. The union of the individual regions represent the marker colour space, and we call its border cluster-polygon. The above process is called colour training, and is of the type of supervised learning.

4.2.4 Colour classifier

The kernel of the colour classifier (Fig. 13) is a 16-bit look-up table (LUT). This LUT is the implementation of the region being bounded by the cluster-polygon. Its input is a data stream of 16-bit HS values, which are formed by concatenating the 8-bit H and 8-bit S values. Its output is binary and indicates whether the input value falls within the cluster-polygon or not. Low intensity pixels do not provide reliable HS values and are themselves of no interest. Thus,

pixels being classified as marker pixels, but having an intensity below a certain threshold are reset to zero (background) by multiplication of the LUT output with the thresholded intensity V . This step of postprocessing would not be necessary, if RGB values would be used as input. But the use of a 16-bit LUT requires to reduce the resolution to 5 bits for each RGB component. This is an other reason, why we preferred the HSV colour space. Fig. 14 shows the initial segmentation of Fig. 12c using the cluster-polygon of Fig. 12f. It can be seen from Fig. 14 that most of the marker pixels are correctly classified, yet some of them and a few background pixels are misclassified. Misclassification of marker pixels is much less critical than of background pixels and can be accepted up to a considerable amount since no shape analysis is used. Although we could avoid false segmentations of background pixels by choosing a smaller cluster-polygon, but this would also eliminate too many pixels belonging to the marker. The classification errors tend to be scattered, as well in the space as in the time domain. Furthermore, the space-frequency bandwidth of the marker region is much lower than the bandwidth of the scattered errors. Therefore the initial segmentation can efficiently be improved by spatio-temporal lowpass postprocessing. We add successive binary frames (time-domain lowpass) and convolve the result with a 7×7 box operator (space-domain lowpass, local 7×7 average). By thresholding the low-pass filtered image, not only misclassified background pixels are removed, but also misclassified marker pixels are recovered, thus the marker region becomes more compact, as shown in Fig. 15. A special colour classifier design tool (CCDT) has been developed, which allows for an easy design of a colour classifier (Arbter & Kish, 2004).

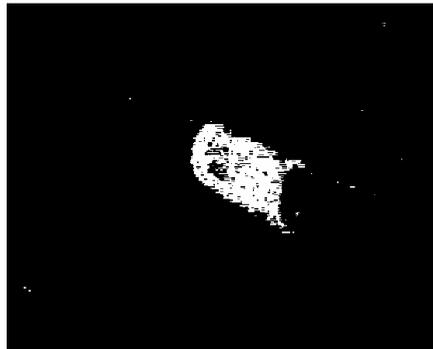


Figure 14. Colour segmented marker



Figure 15. Postfiltered segmentation result

4.3 Robot Controller

The task of the controller is to bring the actual image of the instrument to a desired location at the monitor screen by smoothly moving the robot according to the incoming signals from the image processing system. The desired location is either prestored, or it can be redefined on-line by moving the instrument to the desired monitor position, while the tracking mode is switched off. In the second case the image processing module extracts the actual location values and stores them as reference coordinate values for the future.

As input to the robot controller, we have used the centers of gravity as well as the corners of the bounding boxes. We made the experience that corners are much more reliable in most cases than centres of gravity, especially in the case where the marker is partially occluded.

Since the AESOP 1000 robot system (Computer Motion Inc., 1994) provides direct motion control in the image plane, no user-involvement in the robotic kinematics is necessary. The commands **MoveLeft ()** and **MoveRight ()** specify robot motions such that the laparoscopic image moves to the left and right of the human eyes looking at the monitor image; that is, they specify the x -direction motion in the image coordinate system. Similarly, **MoveUp ()** and **MoveDown ()** control the motion in the y -direction in the image plane. Motions orthogonal to the image plane (longitudinal z -direction motions) are specified by the **ZoomIn ()** and **ZoomOut ()** commands.

Suppose (x_L^0, y_L^0) and (x_R^0, y_R^0) are the reference coordinate values in the left and right images, respectively. Suppose (x_L, y_L) and (x_R, y_R) are the current coordinate values of the colour marker location in the left and right camera images, respectively. Then, we determine the 3D-speed command of the robot motion as follows:

$$\begin{aligned} v_x &= \alpha[(x_L - x_L^0) + (x_R - x_R^0)] \\ v_y &= \alpha[(y_L - y_L^0) + (y_R - y_R^0)] \\ v_z &= \beta \frac{\sqrt{(x_L - x_R)^2 + (y_L - y_R)^2} - \sqrt{(x_L^0 - x_R^0)^2 + (y_L^0 - y_R^0)^2}}{\sqrt{(x_L^0 - x_R^0)^2 + (y_L^0 - y_R^0)^2}}. \end{aligned}$$

The equations reduce in the case of mono-laparoscope to:

$$v_x = \alpha(x - x^0), \quad v_y = \alpha(y - y^0).$$

This intermediate commands are then converted to the specific **Move . . . (speed)** commands by separating magnitudes and signs for speed and direction.

With this control law the closed loop system has approximately a first order low-pass transfer function. The bandwidth (dynamics) depends on the values a for lateral motions and 0 for longitudinal motions, respectively, and may easily be adapted to the surgeon's needs. The system follows asymptotically slow instrument motions, as they occur if the surgeon changes the operational field, but damps fast motions, as they occur if the surgeon treats the tissue. This behavior provides the surgeon with smoothly moving images in the first case and with quasi-stable images in the second case, as is desired.

4.4 Robustness

Safety is of the highest priority in surgery. The correct segmentation of instruments is crucial for correct visual guidance. A problem particular to laparoscope images is that the received light by the narrow lens system is usually very weak, so that the CCD signal (including noise) has to be highly amplified. For this reason, the signal-to-noise ratio is considerably lower than that of a standard CCD-camera. In our system, the high rate of correct colour segmentation in the presence of noise is attributed to the use of spatio-temporal low-pass filtering.

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- HTML (Free /Available to everyone)
- PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)
- Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below

