2 Endoscope Robots and Automated Camera Guidance

2.1 A Survey of Motorized Endoscope Holders

This section describes in chronological order the development of motorized endoscope[1] holders. In the literature these are often referred to as active endoscope holders. However, this term is avoided here as it suggests active robots as defined in the previous chapter (p. 1.2) as opposed to telemanipulated robots. All endoscope holders treated in this section belong into the category of telemanipulated robots.

The focus in this section is on the mechanical structure and human-machine interfaces of the endoscope robots. A more in depth look at the approaches for the automation of camera guidance, is the topic of the following section (2.2). Given the short history of modern minimally-invasive surgery (cf. 1.1.1), the first attempts to replace the camera assistant with a robot have been quite early.

Before motorized endoscope holders, simple mechanical holders have been employed that must be unlocked, readjusted by hand and then relocked for each change in endoscope position. Examples of these mechanical holders are the Robotrac, Aesculap, the First Assistant, Loanard Medical Inc., the Omni-Tract, Minnesota Scientific and the Iron Intern, Automated Medical Products Corp. The system by Erbse et al. [42][43] is already a mechatronical system that uses piezoelectric actuators for locking and unlocking. A step beyond these devices originating in mechanical retracting systems for open surgery were

[1]Only rigid endoscopes that remain partially outside the patients body are included, e.g. laparoscopes. In particular this excludes flexible endoscopes, such as gastroscopes, and capsule endoscopes.

A. Bihlmaier, *Learning Dynamic Spatial Relations*,
DOI 10.1007/978-3-658-14914-7_2

special kinematics with a built in remote center-of-motion.[2] One of these representatives is the TISKA Endoarm [44], which was developed at the Karlsruhe Research Center and commercialized by Karl Storz GmbH. TISKA is not purely mechanic, but uses electromagnetic friction to keep a position that is enabled and disabled by a foot pedal.

Although motorized, the Roboscope [45] is excluded from the survey because the only control mechanism it provides is manual hand guidance. In the same manner, modified endoscopes with panoramic optics augmented by motorized optical zoom and purely visual translation of the image, e.g. as presented by Kimura et al. [46] and the ImagTrac system [47], Olympus, are considered outside the intended scope. The Kaist Laparoscopic Assistant Robot (KaLAR) [48][49] that builds on a custom endoscope featuring an additional bending mechanism is therefore also excluded. The Robotic Flexible Laparoscope System (RFLS) [50] is controlled by head motion using a gyroscope, but since the control modality is very similar to the system in section 2.1.21 this flexible endoscope is not further elaborated. Tamadazte et al. [51] insert two additional miniature high-definition cameras with the endoscope into the trocar, thereby increasing the surgeon's field of view. The ViKY endoscope robot (cf. 2.1.9) is used to manually position the augmented endoscope, yet, further details about this system fall outside the scope set here.

Two crucial advantages of endoscope holders compared to a human assistant are a more stable image and the potential to take up less space around the operating table [17]. Even the fine motor tremor of the human camera assistant is reported to deteriorate the image stability [25].[3] Furthermore, the idea of solo surgery, i.e. performing laparoscopic procedures without human assistance, is a big motivational factor for many surgeons [52].

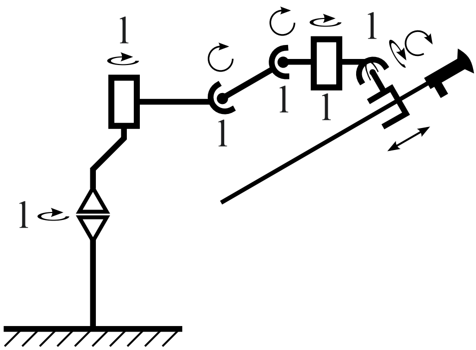
[2]Given the robot is positioned correctly, the mechanical structure of robot kinematics with a remote center-of-motion guarantee that the trocar constraint (Fig. 1.10) are maintained.

[3]See section 6.1.4 for a model and an empirical evaluation how vibrations influence robot-assisted endoscope guidance.

2.1.1 Endex

The Endex endoscopic positioning system [53], Andronic Devices Ltd., can be seen in the middle between mechanic endoscope holders and motorized ones (Fig. 2.1a). It only features a single actuated degree of freedom (Fig. 2.1b). Otherwise the Endex is a passive mechanical structure that is pneumatically locked and unlocked. The single motor uses the endoscope optics as a linear joint by means of a silicon roller. A foot pedal allows to move the orientationally fixed endoscope in and out of the trocar. The authors report their experience for using the Endex in over 60 gynecologic surgical procedures. Best user experience is reported for tasks that require an extended period of time in the same operative field. Spatial constraints are improved by attaching the holder to the other side of the OR table.

(a) Image of the Endex endoscope holder (Source: [53]).



(b) Endex kinematics: All passive joints *l* employ pneumatic locking, which is controlled with a release button. The endoscope is used as rail in the actuated linear joint.

Figure 2.1: The Endex endoscope holder.

2.1.2 AESOP (Automated Endoscope System for Optimal Positioning)

The Automated Endoscope System for Optimal Positioning (AESOP) was first described by Sackier and Wang [54] in 1994 (Fig. 2.2). A custom manipulator with 4 actuated and 2 passive joints is presented with a SCARA-like structure (Fig. 2.3). The AESOP joints are backdrivable, this feature is used to enable manual repositioning of the endoscope as with a mechanical holder. However, the surgeon has to pay attention to the robot's kinematic structure and must thus make sure to obey the trocar constraint by himself.[4] Other interfaces for manual control are pressure sensitive foot pedals and a joystick-like hand controller. The current endoscope position can also be stored and recalled later on.

Figure 2.2: The AESOP endoscope holder (Source: [54]).

The AESOP, Computer Motion, was FDA-approved in 1994 [55] and subsequently evaluated in a number of clinical studies. Jacobs et al. show that using the foot pedal to control the robot leads to

[4]Employing the backdrivability with unpowered motors thus feels quite different than a robot that measures the external forces exerted by the surgeon and actively complies to them in task space instead of joint space. The latter is referred to as "hands-on" mode or "kinesthetic" mode. See the description of the LARS robot in the current section.

increased task completion times compared to manually guiding the endoscope [24]. In 1999 Arezzo et al. compared the duration and subjective rating of 70 laparoscopic experiments performed as solo surgery [52]. The original AESOP (AESOP 1000) controlled by foot pedals [56] and the later introduced voice-controlled AESOP 2000 [57] were compared with the passive TISKO Endoarm and human assistance. Human assistance had slightly, but significant, lower completion time compared to all other setups. Furthermore, it was found that although completion time with unmotorized endoscope holders was lower, the best subjective rating was given to the voice-controlled AESOP. Foot pedals and (finger-ring) joysticks were the least favorable control options, both in terms of duration and user experience. The authors note the relatively large space requirements of the AESOP and lack of support for 30 degree optics. The former is largely due to the large cart the AESOP is attached to and not the robot arm.

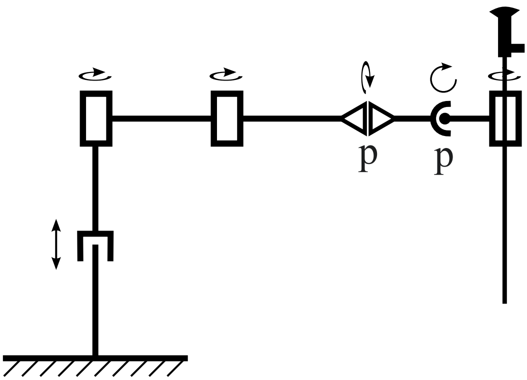


Figure 2.3: Kinematics of the AESOP endoscope holder. Joints marked with a letter *p* are unactuated passive joints.

For urological interventions Partin et al. [58] report results from 17 procedures performed with AESOP assistance. No increase in operating time compared to human assistance was found. However, in three cases intraoperative bleeding occurred and required human camera assistance. For gynecological surgery Mettler et al. [59] describe their experience with the voice-controlled AESOP. Although the AESOP provides a more stable image and results in reduced operative times, at the same time “the whole procedure requires more concentration from the surgeon”. Kavoussi et al. reports a significantly steadier endoscopic image and no difference in operative times with robotic assistance [60]. Ballantyne reports similar results with respect to image stability and operating time in solo surgery of laparoscopic colectomies [55].

The newer AESOP 3000 was evaluated by Nebot et al. [61] in a phantom box and compared to the EndoAssist with head-tracking as input modality (cf. p. 35 in this section). Even with the improved speech recognition of the newer AESOP, the authors report frequent voice recognition errors that adversely affect the surgeon’s performance. Punt et al. [62] evaluate control of endoscope zoom and light intensity through voice commands, a touch panel and a human assistant. Although there are also several recognition failures, voice control is deemed the best control option. Kipfmüller notes that use of the AESOP results in less frequent cleaning of the optics [63]. Another aspect reported by Shew et al. [25] is the prevention of motion sickness in small operative spaces due to the steadier endoscope image. On the other hand, the authors note a relevant learning time for the surgeon. For thoracic sympathectomy (VATS) for hyperhidrosis procedures[5] Martins Rua et al. [64] compared the performance of the AESOP to human assistance. Each group of the randomized study comprised 19 patients. Performance endpoints were number of wrong camera movements, number of optics cleaning, duration of procedure and several endpoints related to patient outcome. No difference was found in the patient endpoints and camera movements. The optics had to be cleaned less often, on average 0.22 to 0.42, but operation time was longer with an average of 12.89 to 9.89 minutes. In a study of 11 AESOP-assisted and 15 human-assisted colectomies, Merola et

[5] Clamping or dividing the sympathetic nerves inside the chest to cure abnormally increased sweating.

al. [65] find that neither length of operation nor patient outcome are significantly different. At the same time the robot reduced the number of required OR staff. The preference of some surgeons for robot assistance is believed to stem from "a subjective sense of better control of their surgical field contributes to this preference. The fluid movement of the robotic camera holder reduces motion sickness, eliminates the wandering of the laparoscope caused by an inattentive human assistant, keeps the picture still and steady until the surgeon commands it to move, and prevents inadvertent rotation of the laparoscope." Proske et al. [66] find no difference in terms of complications, duration of procedure or hospitalization between human-assisted and robot-assisted cholecystectomies and colectomies. The study group included 47 patients operated a year earlier with a human assistant and 50 patients operated with AESOP assistance.

Despite these positive results, there are also known general limitations of the AESOP [55]:

- Control of the camera robot with voice or joysticks is slow compared to human assistance for rapid changes of the visual field.
- Voice control can distract other OR staff.
- Manual control encourages settling for a single visual field instead of going back and forth between better suited ones.

Kraft et al. [67] compared the AESOP to human assistance for laparoscopic cholecystectomy and hernioplasty in a randomized study with two groups of 120 patients. For cholecystectomy, they found that the preoperative setup time was 5 minutes longer in case of robotic assistance. Also the operating time was about 5 minutes longer than with a human assistant for a median overall time of 30 minutes. Unsurprisingly, the voice-controlled AESOP required many more explicit commands than the human assistant. On a scale from 1 to 5 the subjective evaluation of the AESOP was 0.6 lower and the ability to achieve optimal focus was also worse by 0.6 points. However, the authors note that the human assistants in the control group were remarkably well trained with an average of 100 past assistances.

The AESOP is also part of the ZEUS telemanipulation robot described later in this section (p. 69).

2.1.3 Begin and Hurteau et al.

Begin and Hurteau et al. [68][69] presented a robotic camera holder in 1995.[6]. The system used a small industrial robot (CRS A460) with six degrees of freedom (DoF). An additional passive universal joint was used to fix the endoscope to the robot flange (Fig.2.5). Positioning of the endoscope was performed manually by the surgeon by means of a simple joystick. The joystick input was mapped to spherical coordinates centered around the camera trocar (Fig. 2.4). Thereby, the surgeon directly controlled the altitude (left-right motion), zenith (up-down motion) and radius (zoom motion). After an animal study, the system was evaluated in eight patients. It was shown that the surgeon was able to complete all procedures without a human camera assistant.

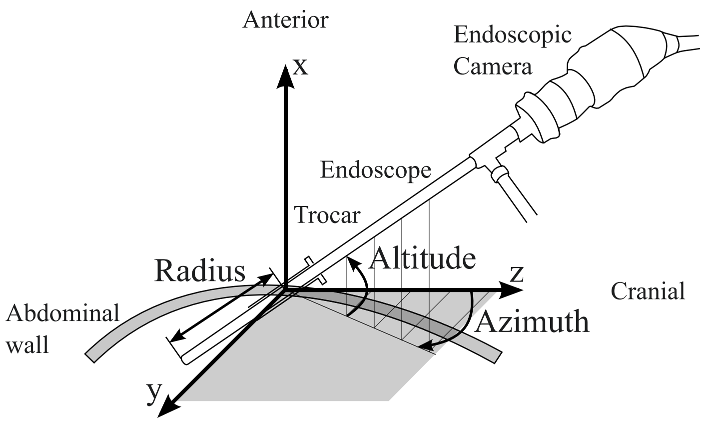


Figure 2.4: Illustration of a spherical coordinate system centered around the endoscope trocar.

[6]The first human applications and the paper submission were even two years earlier in 1993 [70]

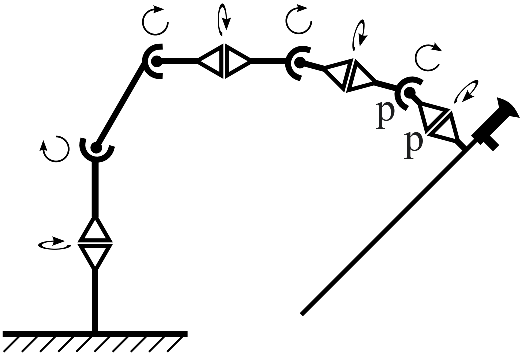


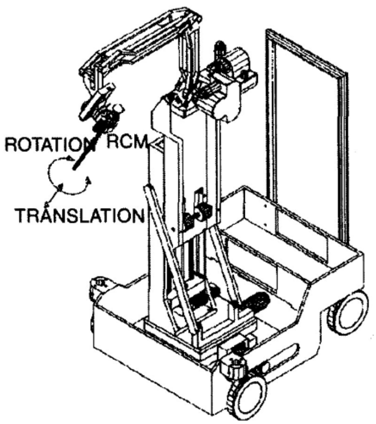
Figure 2.5: Kinematics of the CRS A460 together with a passive universal joint as used in the system by Begin et al.

2.1.4 LARS (Laparoscopic Assistant Robot System) / PLRCM (Parallel-Linkage Remote-Center-of-Motion)

The LARS [71] or PLRCM [72] robot[7] was published in 1995 by Taylor and Funda et al. (Fig. 2.6). It features a custom robot manipulator with a remote center-of-motion defined through four bar linkage kinematics serially coupled with a 3-axis linear cartesian stage (Fig.2.7). As a result, the trocar constraint is already taken care for in hardware. This improves patient safety by largely prohibiting injuries of the abdominal wall through unintended robot movements. A force-torque sensor mounted at the manipulator flange provides additional safety by monitoring the forces exerted by the robot. However, friction in the trocar and the necessary deformation of tissue does not allow highly sensitive force thresholds.

LARS implements two different human-machine interfaces for manual control. The first is an instrument mounted joystick that is basically directly mapped to the robot's rotational degrees of freedom (DoF). Alternatively, the robot can be controlled by directly exerting additional forces on the endoscope. These external forces are measured by

[7] The name LARS will be used for both publications.



(a) Schematic drawing of LARS (Source: [71]).

(b) Upper part of the LARS endoscope holder (Source: [73]).

Figure 2.6: The LARS / PLRCM endoscope positioner.

the force-torque sensor and the robot complies with them by moving along the force vector, i.e. “hands-on”. This cannot be used at the same time as the described safety feature based on the force-torque sensor. Furthermore, the authors consider adding speech commands as a option for direct control. Beyond the manual control modes, LARS also comprises a less direct control modality: The surgeon can use the joystick to move a superimposed cursor on the endoscopic display. He thereby selects an anatomical feature in the image that the endoscope is supposed to put in the center of the field of view. In order to map the two-dimensional monitor selection of the surgeon to a 3D point inside the patient triangulation is used. If a stereo endoscope is in use[8] this can be accomplished from the two camera perspectives. In case of a monoscopic laparoscope the robot slightly displaces it in order to acquire the two required images. Once the target point is known, the robot moves the endoscope in order to center on the target. LARS was in-vivo evaluated on pigs in 1994.

[8]Stereo laparoscopes were not available at the time.

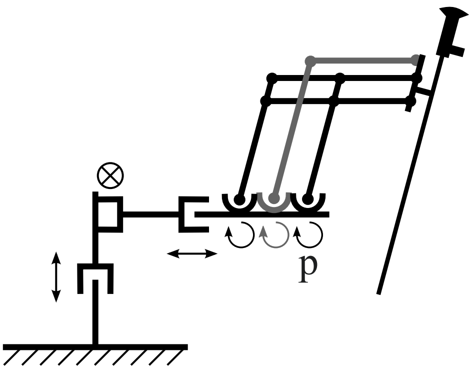


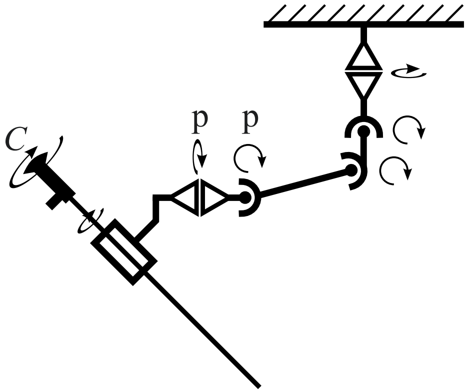
Figure 2.7: Kinematics of the Laparoscopic Assistant Robot System (LARS). The structure of the two parallel linkages is shown simplified. Passive joints are marked by *p*.

In a later publication [73], optimal control of the robot under trocar constraint (cf. Fig. 1.10) is formulated as an optimization problem. This enables a generic description of task-dependent control modes including motion constraints. The control approach is compatible with kinematically deficient systems, i.e. having less than 6 degrees of freedom, as well as kinematically redundant systems, i.e. more than 6 DoFs. An example of the former is the LARS with 5 DoF and an example of the later is the HISAR with 7 DoF (see 2.1.5).

2.1.5 HISAR (Hopkins-IBM Surgical Assistant Robot)

Funda et al. [74] report on application of the Hopkins-IBM Surgical Assistant Robot (HISAR) as endoscope holder (Fig. 2.8a). HISAR is ceiling-mounted and features 7 degrees of freedom (Fig. 2.8b). Two of these degrees are passive. One of these DoFs is the actuated rotation of the camera around the endoscope optics. Therefore four actuated DoFs are used for positioning the optics' eye piece and thereby the spherical orientation, insertion depth and rotation along the endoscope shaft.

(a) HISAR (Source: [72]).



(b) HISAR kinematics. Passive freely moving joints are marked with p. The rotation of the camera C around the endoscope optics is also motorized.

Figure 2.8: The Hopkins-IBM Surgical Assistant Robot (HISAR).

A comparison of the LARS/PLRCM with the HISAR with respect to safety, ergonomics and control are presented by the same authors [72]. They find that there are trade-offs to be made between these criteria. Main differences relate to precision of motion, especially under external disturbances, robot workspace and the space required by the robot in the OR. In case of the robots under comparison, the LARS is more precise, but with a much smaller working volume and it blocks a lot more relevant space in the OR. For safety reasons, the ability to measure force-torque with LARS due to the fully constrained endoscope is beneficial. On the other hand the backdrivability of the HISAR can also be seen as a safety feature in case of electronic failures.

Funda et al. conclude their evaluation with valuable insights into robotic assistance for surgery: “It is difficult to design a general purpose surgical robot. The workspace, ergonomic and precision requirements associated with different procedures vary greatly. Once a promising class of applications for robots in surgery is identified, a specific mechanism and design approach may be required to adequately address the application requirements within cost constraints. However, the manipulator itself is only one part of an overall system which includes control electronics, computers and software, human-machine interfaces, surgical end-effectors, and much more."

2.1.6 Laparobot, EndoSista, EndoAssist

In course of the development of the EndoAssist, Armstrong Healthcare Ltd., the intermediate prototype systems were known as Laparobot and for a longer period of time as EndoSista [75]. The EndoAssist was then introduced as a commercial product in 1998 [76]. For the sake of brevity and given their minor differences, only the kinematics of the final EndoAssist are described. The EndoAssist is not attached to the OR table, but is integrated into a wheeled cart that is positioned around the OR table. With its high vertical axis and long central boom, EndoAssist is intended to be placed behind the surgeon and to reach over him. The actuated kinematics are made up of the vertical axis and two rotational degrees of freedom at the end of the overhead boom (Fig. 2.10). A remote center-of-motion is implemented by an additional offset linkage that connects to the endoscope [75]. Proper positioning of the EndoAssist with respect to the camera trocar is facilitated by two lasers in the overhead boom.

A very relevant feature of the EndoAssist in this survey relates to its unconventional control interface. The EndoAssist uses head-tracking together with a simple foot switch [78]. The surgeon wears a headband that contains electromagnetic transmitters which provide the relative movement to an receiver mounted at the endoscopic monitor. In order to allow free head movement and to prevent unintended movement of the endoscope, a foot switch is used to engage control over the robot movement. If the surgeon wants the endoscope to move left, he presses the foot switch and turns his head slightly left. The robot will move into this direction until the surgeon turns his head into another direction or disengages the robot by releasing the foot switch. This is quite different from the approach of directly linking head

Figure 2.9: The EndoAssist endoscope holder (Source: [77]).

motion to motion of the endoscope. For example, Voorhost et al. [79] performed experiments on directly linking endoscopic perspective and head rotation of the surgeon. Although this research has been resumed recently [80], the relationship between instruments and position of the surgeon poses many obstacles in practice.

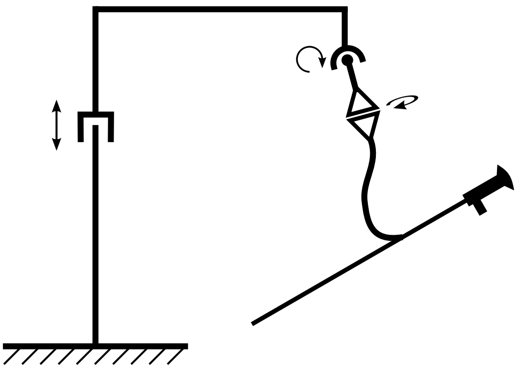


Figure 2.10: Kinematics of the EndoAssist endoscope holder.

The first clinical trials of the prototype system were conducted in 1993 [81]. A major benefit of the system, as reported by Finlay [82], the founder of Armstrong Healthcare Ltd., compared to human assistant is the steady image - especially at the end of long procedures. Aiono et al. [78][83] report results of a randomized clinical study with 96 laparoscopic cholecystectomies conducted by six surgeons. About half of the interventions were performed with robot assistance and the other half with human assistance. In their results mean operating time with robotic assistance was 8 minutes shorter for an average duration of 70 minutes. Furthermore, they found a smaller variation in the operating time. Only three interventions were required to fully master the learning curve for head control. Regarding trainees, Aiono et al. see a benefit in freeing them from camera work and thus allowing them to focus more on relevant matters. Yet, they acknowledge the concern of junior team members not taking part in robot-assisted procedures anymore in the future. In the domain of urological surgery, Kommu et al. [77][84] report the results of a clinical study with 51 procedures, half of which are performed robot-assisted. The evaluated endpoints were: body part discomfort score (BPDS), subjective usability, number of required lens cleanings, setup time, overall operative time, subjective surgical performance and number of required rearrangements of the EndoAssist. Although all procedures could be performed with the EndoAssist, in several occasions the arm had to be relocated. In general more frequent lens cleaning was required in the robot-assisted case. BPDS scores were equal for both assistances. Depending on the procedure type, usability was equal or lower to human assistance. No difference was found in setup time and surgical performance.

Yavuz et al. [85] compare the performance of the AESOP and the EndoSista (EndoAssist) for standard camera tasks in training boxes. Their results show a much better performance in terms of task duration for preprogrammed movements (position memory) with the AESOP. Furthermore, voice control of the AESOP was found to be superior to head control in the EndoSista. Yet, the results must be taken with caution since all experiments were performed by a single operator.

Qualitatively, the authors note that the AESOP feels less bulky and quality of voice control is diminished in noisy environments. Wagner et al. [86] also compare AESOP to EndoAssist. However, their results are from two groups of 20 patients each undergoing laparoscopic radical prostatectomy (LRP)[9] with assistance by AESOP or EndoAssist. EndoAssist was controlled by the surgeon through the head motion interface. AESOP was controlled by an experienced assistant. Average setup time of AESOP was lower, 2 minutes, compared to EndoAssist, 5.3 minutes. Only in 1 out of 11 steps in the intervention was a significant difference between assistant-controlled AESOP and surgeon-controlled EndoAssist, which favored the EndoAssist. Even though an assistant is required for LRP, the procedure subjectively benefited from robot assistance because the assistant could fully focus on his other tasks. Nebot et al. [61] performed a similar study in which they evaluate the EndoAssist and AESOP 3000 performance for complex camera tasks in box trainers. They find the head-tracking controlled EndoAssist to be significantly faster than the voice controlled AESOP. A large part of the result is attributed to the continuous control of the EndoAssist compared to the discrete step-wise control of the AESOP. However, better usability is ascribed to the AESOP voice control, given the complexity of precise head movements together with the requirement to press a foot switch. Den Boer et al. [87] compare AESOP to a passive endoscope holder (PASSIST) and to human assistance. From 78 laparoscopic cholecystectomies 30 were performed as solo surgery. Operative times were not significantly different with 42 to 49 minutes average for assisted surgery compared to solo surgery. The number of endoscope repositioning was with an average of 49 times much lower with AESOP than the 114 times with human assistance.

2.1.7 FIPS Endoarm

The FIPS Endoarm [88] is a remote controlled endoscope positioning arm developed by the Karlsruhe Research Center (FZK) and the

[9] Surgical removal of the whole prostate gland.

University of Tuebingen. FIPS was later brought to market by Karl Storz GmbH.

Figure 2.11: The FIPS Endoarm (Source: [88]).

Over 300 phantom procedures were performed without human assistance. The authors report shorter operative times compared to human assistance. However, even shorter times are reported for assistance by a passive endoscope holder (TISKA Endoarm, cf. p. 24). They conclude that "none of the solutions tested showed an intuitiveness and reliability competitive with hand positioning". Furthermore, the findings suggest best overall ergonomics when the endoscope robot is placed on the opposite side of the OR table. The FIPS Endoarm was later used as part of the ARTEMIS teleoperation system [89] for MIS. A version of FIPS was also experimentally used for gynecological laser laparoscopy [90].

Buess et al. [91] report about a randomized study involving phantom experiments of laparoscopic cholecystectomy. Three groups of 15 interventions each were performed: human assisted, assistance by FIPS Endoarm controlled with the finger-joystick and voice-controlled FIPS Endoarm assistance. Although the joystick interface was judged more intuitive, intervention times were slightly lower with voice control. One major disadvantage in the voice command interface was found to be the limited motion of each command. For example, a long move

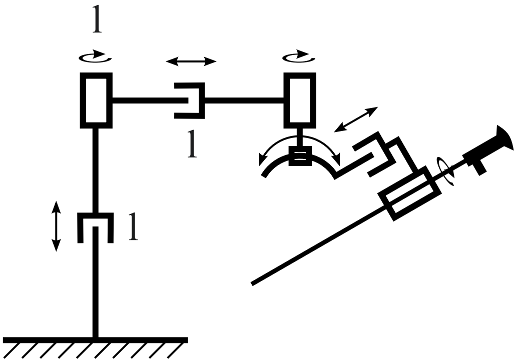


Figure 2.12: Kinematics of the FIPS Endoarm. Unactuated mechanically locked joints are marked with l. The remote center-of-motion is realized through a C-arch mechanism.

into the same direction requires a sequence of multiple identical voice commands. While the intervention time with the FIPS Endoarm was lower than with human assistance, the overall time including setup and break-down was significantly longer. Yet, the systems allows to reduce the number of required assistants [92]. Arezzo et al. [93] attests FIPS the best results and most intuitive user experience in a series of 400 phantom cholecystectomies in comparison with AESOP and EndoAssist.

A later iteration of the FIPS Endoarm, named FELIX, featured visual tracking of instruments [94]. The image was divided into three areas, central, inner and outer region, if the instrument leaves the inner region the endoscope moves towards it until it is within this region again.

2.1.8 Munoz et al., ERM (Endoscopic Robotic Manipulator)

The setup presented by Munoz et al. [95] in 2000 consists of a Stäubli RX60 with two passive rotational DoFs in the endoscope adapter. This kinematic structure is identical to the one in Begin et al. (see

Fig. 2.5). A mapping of Cartesian to spherical coordinates is employed for control (cf. Fig. 2.4). Voice-control with discrete commands and joystick with four degrees of freedom, modelled after an endoscope, is used to control the endoscope in spherical coordinates with the camera trocar as center of motion. Several phantom and pig experiments were conducted with the system to show its general feasibility [96].

Figure 2.13: The ERM endoscope holder (Source: [97]).

In 2001 Munoz et al. [98] present new custom kinematics, later called ERM, for an endoscope holder (Fig. 2.13). The new SCARA-like robot (Fig. 2.14) features one linear and two actuated rotational joints together with two passive rotational joints at the endoscope holder. An extensive discussion of the required workspaces and possible kinematics solutions can be found in a later publication [97]. Voice control and remote control via a 3D mouse, SpaceBall, provide manual control over the endoscope movement [99].

In 2005 a follow up publication describes a refined version of the ERM (Fig. 2.14) together with in vivo and clinical results [100]. A

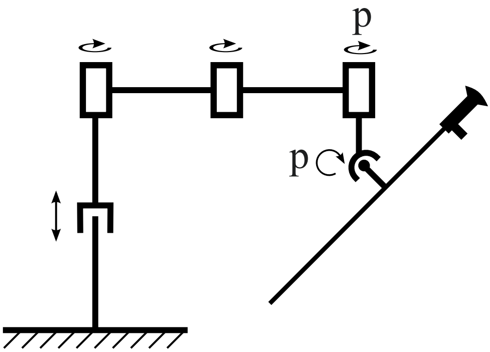


Figure 2.14: Kinematics of the custom endoscope arm (ERM) described by Munoz et al. in 2001.

failure mode and effects analysis (FMEA) is undertaken to identify possible failure szenarios and their potential hazard for the patient. Munoz et al. [101] report that in the experiments with 16 patients the robot performance was good. Robot-assisted operating times were shorter and the optics had to cleaned less often.

2.1.9 LER (Light Endoscope Robot), ViKY (Vision Kontrol endoscopY)

The Light Endoscope Robot (LER) was introduced by Berkelman et al. in 2002 [102]. From the onset the goal was to develop a much smaller and lighter endoscope manipulator compared to all previous systems. The very first design was a cable driven robot actuated by pneumatic artifical muscles actuators. An immediately following update of this design replaced the pneumatic actuators by electric motors for their higher accuracy and easier control [103]. Already the next prototype of 2003 [104] was then very similar to what became a commercial product under the name ViKY later on. Therefore the first design will not be discussed here. The kinematic diagram of LER is shown in Fig. 2.15. Two passive universal joints together with a passive rotational joint that can be mechanically locked in unison position the actuated part of the LER concentrically around the camera trocar (Fig. 2.16a). The actuated part consists of a special serial kinematic: A round rack and pinion actuator, a C-arch mechanism and a linear rack and pinion actuator that clips to the endoscope. This structure implements spherical kinematics directly in hardware. Each 'natural' movement (left ↔ right, up ↔ down and in ↔ out) is directly mapped to a single actuated joint. Given a horizontal attachment, the image horizon remains stable. Since no rotation around the endoscope shaft is possible, LER is only suitable for 0° optics. Both passive and actuated part – including gears and the brushless motors – can be autoclave sterilized and thus no part of the robot requires draping. Each joint is backdrivable if the motors are switched off. This can be used for initial positioning. Basic manual control interfaces are a miniature keypad that is clipped to the instrument handle and voice commands. Long et al. [105][106] reports on an initial pig and cadaver experiment series undertaken from 2003 to 2005 with the LER. The compactness of the LER and easy setup is reported as a practical advantage. After initial breakdowns of the hardware, the later prototypes proved sufficiently reliable.

The LER was commercialized after clinical studies (2007) with FDA approval in 2008 under the name ViKY (Vision Kontrol endoscopY) [107]. A multidirectional foot pedal and voice commands via bluetooth headsets are provided as control interfaces (Fig. 2.16b). In addition, the current endoscope pose can be saved and later on recalled by voice commands. According to Gumps et al. [108] setup time of the ViKY is significantly shorter compared to AESOP, on average 41 to 253 seconds. The authors report a similar time reduction in (simulated) emergency removal, 3 seconds for ViKY and 8 seconds for AESOP. Voice control is evaluated to have a low success rate of 71%[10] and thus the foot pedal was judged superior. Although the ViKY is much more compact with respect to space at the OR table, the relatively large footprint of the rack and pinion disk can be obstructive for the instrument trocars. The initial clinical study

[10] The authors report 67% for AESOP 3000.

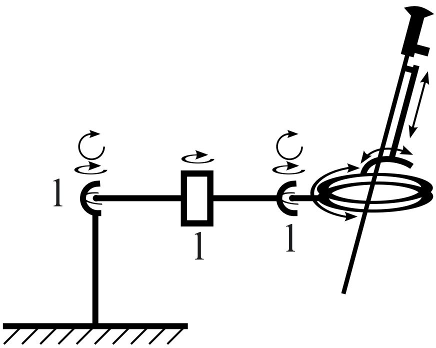
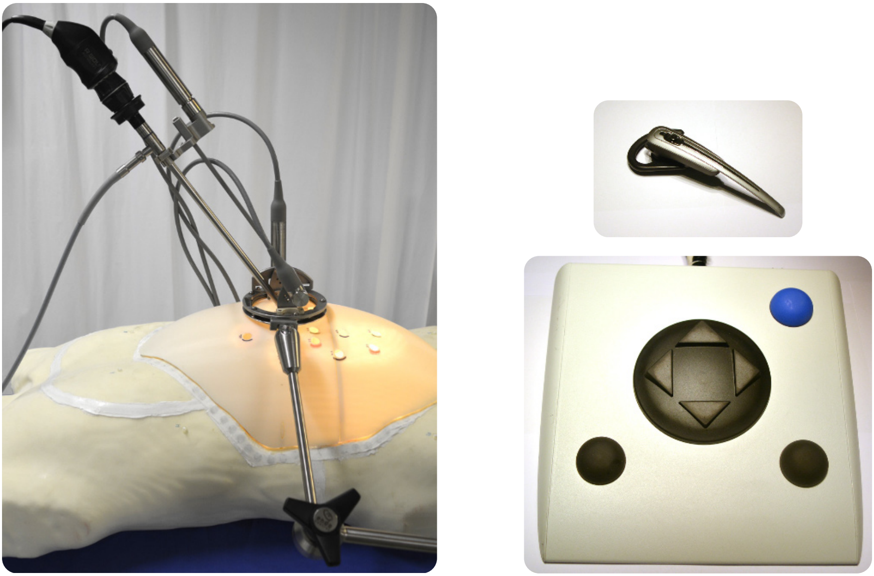


Figure 2.15: Kinematics of the LER and ViKY endoscope holders. An important difference in kinematics between LER and ViKY is the realization of the linear axis: LER uses a cable-driven design counteracted by a compression spring. ViKY employs a rack and pinion actuator. The first three joints l are mechanically locked.

comprised 53 patients with prolapse surgeries, prostatectomies[11] and cholecystectomies being the dominant procedures. Unfortunately, no comparative data to human assisted interventions is provided for the reported robot-assisted operative times. The setup time for ViKY was found to be around 5 minutes. As main advantage the freeing of one assistant's hand is noted, lack of amplitude of motion as main limitation.

Long et al. [109] present a study of 20 patients undergoing urologic surgery with ViKY assistance. One patient withdrew his consent for robotic-assistance. In two cases the robot could not be used: Once due to malfunction and in the other case due to difficult adhesiolysis[12]. Out of 17 interventions started with ViKY assistance, 5 were not completed with the robot. In one case the voice control failed and in four cases the surgical conditions required human assistance. Complete autonomy of the surgeon over the camera is stated as main surgical

[11]Surgical removal of the prostate gland.
[12]Surgical devision of irregular adhesions.



(a) The ViKY endoscope holder attached to the OR table and centered around the trocar of a medical phantom.

(b) ViKY control interfaces: Discrete voice commands and a foot pedal for continues control

Figure 2.16: The ViKY endoscope holder.

advantage. Limitations in the motion range of ViKY were found to represent a practical issue. Furthermore, the robot becomes less usable, if a lot or very wide motions are required. The authors qualify the results by stating that "a pilot study that can only assess the feasibility and the safety of the procedures using the robot. Further studies should investigate the clinical impact."

In 2009 Berkelman et al. [110][111] presented a telemanipulation setup for MIS based on the kinematics of the ViKY system. It consists of three ViKY robots, two small ones for the instruments[13] and a middle sized one for the endoscope, together with a master console

[13] Custom instruments with an actuated wrist are used in the setup.

featuring two haptic devices, Phantom Omni. Voice commands are used to control the endoscope guiding ViKY.

Research by Voros et al. [107] into automated positioning based on visually tracked instruments will be discussed in section 2.2.

2.1.10 Naviot

Kobayashi et al. published the design of a novel five-bar linkage mechanism in 1998 [112][113]. One particular goal was to mechanically limit motion to a safe region and position the manipulator and especially the motors away from the patient. The Naviot [114], Hitachi, introduced in 2003, is a remote-controlled endoscope holder based on this five-bar design (Fig. 2.17).

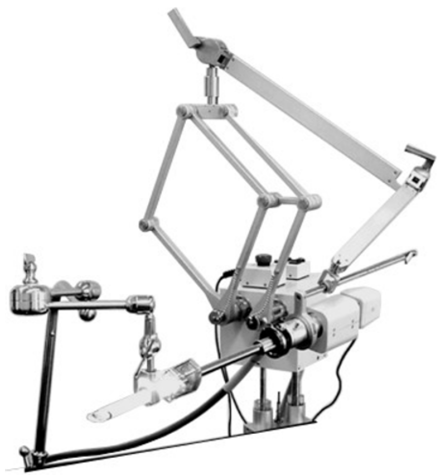


Figure 2.17: The Naviot endoscope holder (Source: [115]).

The five-bar linkage mechanism (Fig. 2.18) mechanically limits the range of motion to 45° horizontally and 25° vertically. Insertion depth is fixed, instead a motorized optical zoom is employed. A two button controller that is attached to the laparoscopic instrument is used to control endoscope orientation and optical zoom.

The system was initially evaluated through seven cholecystectomies in 2002, which showed the feasibility of the mechanism and its control. A later study by Tanoue et al. [115] compared a group of ten cholecystectomies with Naviot assistance to human assistance. The average total operative time of 89.3 minutes with the Naviot assistance was found to be significantly longer compared to human-assisted interventions with an average of 74.8 minutes. This additional time was largely due to the Naviot setup time. Using a shorter endoscope in combination with an optical zoom is reported to be beneficial because of less chance for contact with organs.

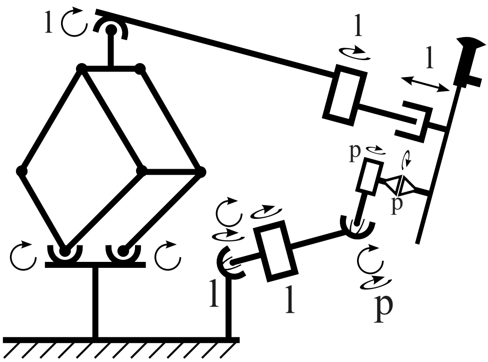


Figure 2.18: Kinematics of the Naviot. The endoscope can only be rotated in two dimensions. Zoom is achieved by optical means. The lower passive kinematic chain fixes the endoscope's center of motion. Intraoperative locked joints l and free moving passive joints p are marked.

Yoshino et al. [116] evaluated the Naviot robot in two patients for thoracoscopic surgery[14]. Although the experiments were successful, concerns are raised about usability of the Naviot for more complex procedures given its small motion range. Yamada et al. [117] also successfully tested the Naviot in two thoracoscopic interventions.

[14]VATS: Video-assisted thoracoscopic surgery

2.1.11 SOLOASSIST

The SOLOASSIST [118], AKTORmed GmbH, was introduced as endoscope holder for visceral, urological and gynecological surgery. The SOLOASSIST directly clamps to the OR table. It consists of three actuated rotational and two passive rotational joints (Fig. 2.20). All actuated joints are fluidic actuators with integrated track measurements. By means of an overpressure function, each joint is backdrivable with a defined force. The camera end of the endoscope is positioned in three translational dimensions around the camera trocar. Because the endoscope tip is situated inside the trocar and the rotational freedom provided by the passive universal joint, a spherical motion around the trocar point results (cf. Fig. 1.10). In order to allow natural movement directions, the surgeon controls the robot not by translation in a Cartesian coordinate system, but in a spherical one (cf. Fig. 2.4). To this end, the SOLOASSIST is registered to the trocar in the setup phase by means of a metal sphere that must be manually positioned above the trocar point. Once registered, a sterilizable instrument-mounted joystick is used for control.

Figure 2.19: The SOLOASSIST endoscope holder (Source: [119]).

Gillen et al. [120] compared the performance of SOLOASSIST to a human camera assistant for cholecystectomies in two groups of 63 respectively 60 patients. Three experienced surgeons performed a total of 123 procedures. A majority of 47 out of 63 robot assisted operations were performed by a surgeon having a lot of experience working with the SOLOASSIST. Human assistance was provided by well experienced residents. The authors found that total operation time was significantly lower for the human-assisted group with an average of 90 to 104 minutes. However, the OR staff minutes, identical for solo surgery using SOLOASSIST, but twice the amount in case of human assistance, strongly favor robotic assistance with 104 to 180 minutes. The authors report that 4.8% of the procedures started with robotic assistance required a switch to human assistance. In these three cases intraoperative complications occurred. The three surgeons assigned the SOLOASSIST an average of 2 on a scale of 1 to 5 for the attributes handling, required force, quality of camera view and general satisfaction with the system. About two uncontrolled or unexpected camera movements per operation were registered. For laparoscopic surgery Holländer et al. [121] report on 1033 procedures performed with SOLOASSIST. In summary eight of the nine surgeons interviewed about their experience preferred robotic over human assistance because of better image stability and control over their view. In 71 cases the robot had to be removed and replaced by a human assistant.

For gynecological surgery Beckmeier et al. [122] evaluated the SOLOASSIST in 104 patients. The SOLOASSIST was used with conventional 2D endoscopes for the first 63 patients and then replaced by the Einstein Vision system, Aesculap / B. Braun Melsungen AG. The latter uses a slightly modified SOLOASSIST as camera holder together with a 3D endoscope. For the purpose of this endoscope holder overview, both systems can be treated jointly. Setup time averaged 7 minutes, the learning curve is reported to be about 20 cases. Total operative time with robot assistance increased by 4 minutes from 110 to 114 minutes. Handling was evaluated with a score 2 (good) on a scale from 1 to 5. An average of 1 unwanted

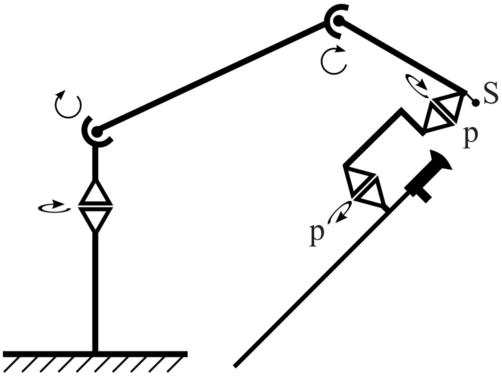


Figure 2.20: Kinematics of the SOLOASSIST. The sphere S at the end of the second link is used for registration with the pivot point. The two joints p close to the endoscope are free moving passive ones.

camera movement per operation was recorded. Furthermore, the authors stress the economic advantage of not requiring a camera assistant.

In the area of head and neck surgery Kristin et al. [123] evaluated and adapted SOLOASSIST as a motorized camera holder. Starting with feasibility evaluations of the SOLOASSIST for interventions in nose, nasopharynx[15] and larynx[16] conducted in anatomical specimens, it was found that the unmodified robot was not well suited. Points that severely hindered the application were: Too high movement speed, not adjustable motion speed and lack of force sensing and force feedback. Especially the latter threatens patient safety because of potential tissue perforations. In summary, the much tighter spaces around the endoscope in head and neck surgery compared to visceral surgery make robots such as the SOLOASSIST unfit for them. A first modification replaced the endoscope adapter by a quick release fastener, such as a magnetic connector, for faster removal of the endoscope from the

[15] Part of the upper respiratory system between mouth and nasal cavity.
[16] The voice box.

surgical area. In a later publication [124] measurements of forces that occur during sinus surgery were undertaken. Computer-aided design (CAD) was used to evaluate different adapted kinematics for head and neck surgery. Although prototypes were built, no detailed results are provided. Furthermore, the interaoperative movements of a hand-held endoscope in head and neck interventions were recorded and analysed [119]. The maximum motions, volumes and enveloping bodies were calculated on a total of 27 sinus, 30 mastoid cavity and 14 larynx endoscopies. Based on these findings in the ENT[17] area, modifications of the SOLOASSIST were proposed. Foremost, the authors state that the "endoscope holder needs five degrees of freedom for use in head and neck surgery instead of only three degrees of freedom (as in abdominal surgery)."

In his MD thesis[18] from 2014, Maifeld [125] evaluates the SOLOASSIST with respect to: Emergency conversion to open surgery; Performance in phantom tasks; In vivo performance on pigs with special attention to the reachability of all abdominal quadrants; and comparison of SOLOASSIST to AESOP and manual camera guidance in phantoms with respect to trajectory stability. Maifeld also poses the question why after more than 30 years of motorized endoscope holders, no system has gained wide clinical acceptance. In the experiments SOLOASSIST had a longer setup time, 297 seconds, compared to the AESOP, 129 seconds. Yet, both times are much shorter than the average preparation time for laparoscopic cholecystectomy of 22.2 minutes [126]. Further results are the mean task time of 7.7 minutes for SOLOASSIST compared to 4.5 minutes for AESOP. It was also more difficult to exactly follow a predefined trajectory with the SOLOASSIST. Subjective evaluation of intuitive use and cognitive load also favored AESOP. In contrast, stability of the horizon was judged to be superior for the SOLOASSIST. Concerns are raised about tension on camera cable and light cable that can build up undetected during the course of an intervention. Qualitatively, the delay between press

[17] Ear, Nose and Throat.

[18] The thesis is written in German.

of the joystick buttons and movement of the robot was found to be too long. Control of the robot was found to be difficult close to the trocar point. The latter is likely due to the singularities that occur close to the trocar point when mapping from Cartesian to spherical coordinates (cf. 5.3.2 and 6.1.5). The author reports from a series of twelve cholecystectomies on pigs that operative time was about equal to human assistance. The surgeons preferred the SOLOASSIST in these experiments because of a reduction in corrective camera motions. Even though in some parts of the intervention, the surgeon and SOLOASSIST impeded each other in their workspace. Part of the thesis is also an acceptance survey based on the answers of 25 surgeons. 76% have no experience with robotic assistance systems, yet, 88% regard mechatronic support systems as desirable.

2.1.12 LapMan

The LapMan endoscope manipulator (Fig. 2.21), Medsys, introduced in 2004 [127], is primarily used in gynecological surgery. The LapMan robot is mounted on a rolling unit. Kinematics of the manipulator are shown in Fig. 2.22. During setup LapMan is registered to the camera trocar using a laser pointer mounted on a part of the robot that is only influenced by the relative positon of the robot to the OR table and the first linear joint. The first linear joint is not used intraoperatively. The three actuated joints are made up of two parallel kinematics and one linear joint at the shaft that attaches to the endoscope. The unconventional design of the LapMan is supposed to improve endoscope motion by locating the camera trocar in the geometric center of the manipulator.

LapMan is controlled by a wireless joystick, named LapStick, which is mounted to a laparoscopic instrument [129].[19] The joystick contains two redundant sets of controls for in/out and left/right motion, the up/down motion is controlled through a distinct 1 DoF flap.

[19] Earlier versions of LapMan were controlled by palm interface that was worn under the surgeons glove (cf. [130]).

Figure 2.21: The LapMan endoscope manipulator (Source: [128]).

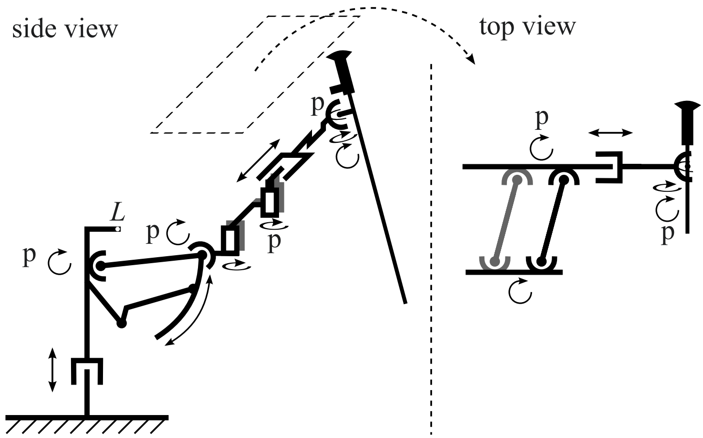


Figure 2.22: Kinematics of the LapMan. The Laser L is used to register robot and trocar. Passive free moving joints are marked with p.

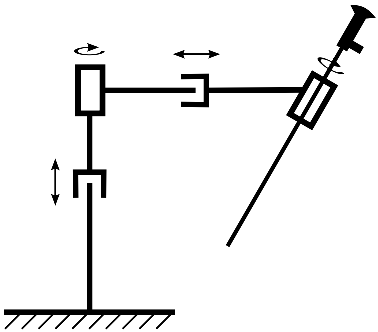
As the title of the paper “How to Maintain the Quality of Laparoscopic Surgery in the Era of Lack of Hands?” by Hourlay [130] already indicates, his focus is on the question how well a human camera assistant can be substituted by a robot. Before going into detail on this question, Hourlay looks at the larger picture in the OR with respect to control of equipment by the surgeon. On one hand, there is an advantage if the surgeon controls more aspects himself, e.g. faster reactions and less coordination overhead. On the other hand, “it would be completely uneconomical to have the surgeon perform all those tasks” given that the surgeon and OR time is a scarce resource. In case of LapMan as an assistance system, Hourlay concludes: First, the investment pays off in less than two years, if the robot is used in 25% of laparoscopic procedures and replaces the human assistant. Second, the stable image contributes to patient safety and increases surgeon’s productivity. Third, after an initial learning curve LapMan improves the surgeon’s work because of decreased stress level and eye fatigue due to a stable view and autonomy of work.

Tchartchian et al. [128] provides the results of a study in gynecological surgery with 50 patients operated by a single surgeon with robotic assistance and compared to the same number of human-assisted interventions. No statistically significant difference in operative times were found. Further objective results are a lower number of corrections of the endoscope and less time for the manual corrections. In the surgeon’s subjective assessment the image stability was also better in the robot-assisted case as well as the satisfaction score. The authors emphasize the improved image stability and the surgeon’s autonomy of vision. Finally, concerns are raised about the management of complications in case solo surgery is widely adopted.

2.1.13 SWARM

The SWARM endoscope holder developed in 2005, recently published by Deshparade [131], is a free standing device (Fig. 2.23a) that is controlled by voice or foot pedal. SWARM has four active degrees of freedom in the kinematic configuration shown in Fig. 2.23b. In addition to the usual left/right, up/down, zoom in/out commands diagonal commands have been implemented (for a discussion of this aspect see 2.1.19). The speed of the resulting motion can also be changed by voice. Voice recognition rates of about 95% are reported. Setup times are reported to be as low as 1-2 minutes, although it needs to be covered as it is not sterilizable. In total 784 laparoscopic interventions were performed under SWARM assistance. The author compares SWARM to ViKY and AESOP. In his comparison, SWARM is superior to both other systems in nearly all aspects, such as setup time, task time, voice-control success rate and obstructions.

(a) The SWARM endoscope holder (Source: [131]).



(b) Kinematics of the SWARM.

Figure 2.23: The SWARM endoscope holder.

2.1.14 COVER (Compact Oblique-Viewing Endoscope Robot)

The Compact Oblique-Viewing Endoscope Robot (COVER), see Fig. 2.24, aims to exploit oblique-viewing (e.g. 30°) endoscopes to reduce the number of degrees of freedom in endoscope holders [132]. COVER is mechanically positioned over the camera trocar and has three actuated degrees of freedom (Fig. 2.25). Vertical and insertion movements are conventionally implemented by moving the endoscope. However, instead of moving the endoscope horizontally, the authors propose to rotate the oblique-viewing endoscope in order to view different left/right regions.

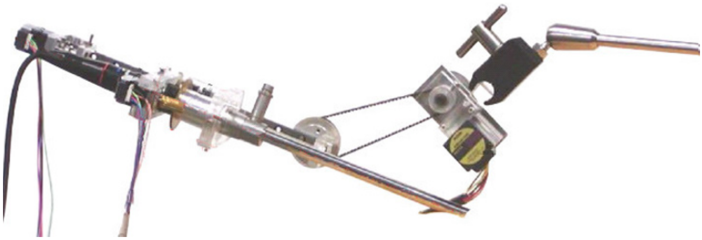


Figure 2.24: The Compact Oblique-Viewing Endoscope Robot (COVER) (Source: [132]).

For control visual head-tracking, the FAce MOUSe (FAMOUS) interface (see p. 89), is utilized. The authors report that it was possible to perform an in vivo cholecystectomy on a pig.

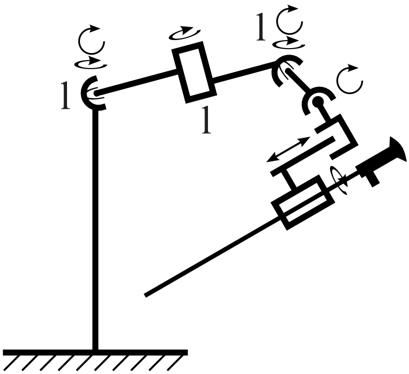


Figure 2.25: Kinematics of the COVER. The first three joints l are mechanically locked during setup time.

2.1.15 SMART (Synthetic Muscle Actuator based Robotic Technology) / P-arm

Taniguchi et al. [133] propose a six DoF endoscope holder based on the kinematics of the Stewart-Gough platform (Fig. 2.27). One version of the system, named SMART (Synthetic Muscle Actuator based

Robotic Technology), uses shape-memory alloy (SMA) as actuators. A later version, named P-arm [134][135][136], uses hydraulic actuators instead (Fig. 2.26). Water is used as fluid in the hydraulic cylinders in order to be biocompatible. The goal for this manipulator is to be very light weight, compact and sterilizable. The authors argue that these attributes, together with accuracy, are best achieved by a parallel structure, such as the chosen Stewart-Gough platform. The endoscope is held in place by a magnetic coupling.

Figure 2.26: The SMART / P-arm endoscope holder (Source: [134]).

In vivo laparoscopic cholecystectomies could be successfully performed on pigs. The endoscope holder was remote controlled by a human camera assistant through a 6 DoF joystick interface, which is also based on a Stewart-Gough platform. The manipulator was found to be about as obstructive as a human camera assistant.

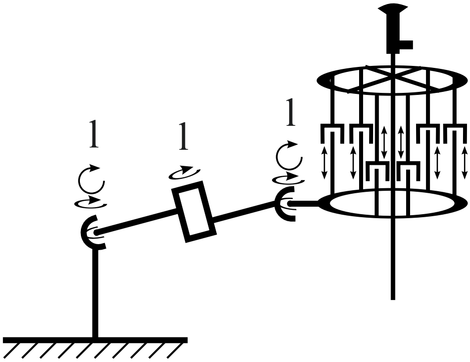


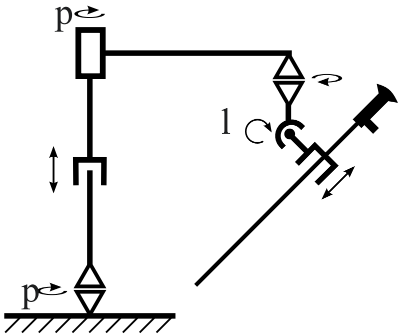
Figure 2.27: Kinematics of the SMART / P-arm (Stewart-Gough platform). The joints l position the actuated part of the endoscope holder over the camera trocar during setup, intraoperatively they are mechanically locked.

2.1.16 Tonatiuh II

The Tonatiuh II endoscope holder (Fig. 2.28a), introduced by Martinez et al. [137] in 2007, Tonatiuh II provides three active degrees of freedom, one freely moving passive joint and two mechanically locked joints for alignment to the camera trocar (Fig. 2.28b). These position the endoscope handle on a sphere around the trocar. The center of the horizontal rotation is aligned with the trocar point during setup of the system. Vertical orientation is achieved by a passive degree of freedom together with translation relative to the trocar point.

The robot is controlled by a gaming joypad. Initial tests in box trainers and animals found that only small forces were exerted at the trocar, even if the endoscope is very close to the abdominal wall. Later on, several clinical tests with different laparoscopic procedures were performed with the Tonatiuh II. Unfortunately, it is unclear whether the surgeon controlled the robot or this was done by an assistant.

(a) The Tonatiuh II endoscope holder (Source: [137]).



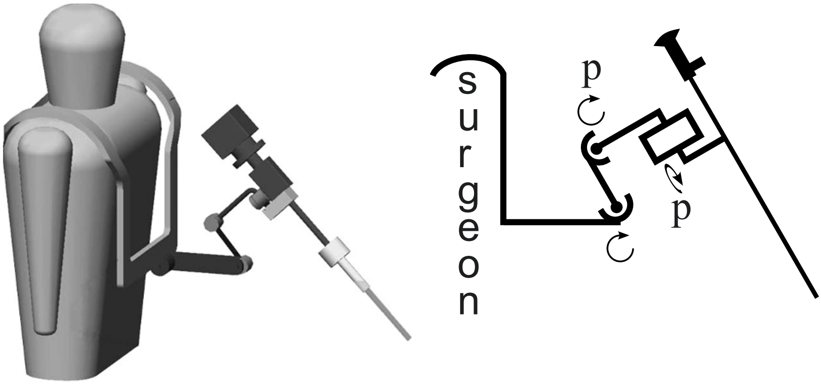
(b) Kinematics of the Tonatiuh II. Passive joints p and joints mechanically locked during setup l are marked accordingly.

Figure 2.28: The Tonatiuh II endoscope holder.

2.1.17 PMASS (Postural Mechatronic Assistance Solo Surgery)

The Postural Mechatronic Assistance Solo Surgery (PMASS) [138] is a wearable endoscope holder (Fig. 2.29a) presented in 2009. PMASS is worn by the surgeon like a harness or a simple exoskeleton and provides one actuated degree of freedom and two passive ones (Fig. 2.29b). Weight without endoscope is about 0.5 kg. The actuated joint provides up/down motion of the endoscope and is controlled by a foot switch. Left/right motion and insertion depth is changed by the surgeon rotating his torso respectively moving towards or away from the patient. The two passive joints in conjunction with the trocar point, map these movements to the corresponding endoscope motion.

Initial evaluations were undertaken in veterinary surgery. Reported limitations are restriction to zero-degree optics. Furthermore, one surgeon reported that the equipment and requirement to hold a certain postures results in discomfort in longer procedures. Mishra et al. [139] describe a clinical study undertaken with PMASS. A total of 13 low-



(a) CAD rendering (Source: [138]). **(b)** PMASS kinematics.

Figure 2.29: The wearable Postural Mechatronic Assistance Solo Surgery (PMASS).

risk patients underwent laparoscopic interventions (appendictomies, ovarian cystectomies and sterilizations) performed by three surgeons. Overall the autonomy and ability to control the optics received good and above scores with 4 and higher on a 0 to 5 scale. However, comfort was only judged with 3 to 3.5 points. Shoulder and neck fatigue was criticised as well as a feeling of being tethered to the patient.

2.1.18 FreeHand

The FreeHand camera holder (Fig. 2.30), Prosurgics Ltd., is commercially available since 2009 and the intended successor of the EndoAssist (see p. 35). FreeHand uses head-tracking together with an activation foot switch for control in the same manner as EndoAssist. The surgeon receives visual feedback through a small indicator display showing the recognized motion direction. Motion speed can be adjusted on the robot. The robot is attached to the OR table and positioned by means of a mechanically-locked passive links.[20] In order to register

[20]This mechanism is very similar to the one employed by the ViKY endoscope holder (cf. p. 42).

FreeHand's remote center of motion, laser pointer guidance is provided. The actuated part of the kinematics consist of one rotational joint, one C-arch mechanism and a linear joint (Fig. 2.31). In contrast to the ViKY, the motor cables are not free-hanging, but part of the passive linkage.

Figure 2.30: The FreeHand endoscope holder (Source: [140]).

Stolzenburg et al. [141] report on a randomized study that compares FreeHand assistance to human assistance in endoscopic extraperitoneal radical prostatectomy (EERPE)[21]. 50 EERPEs were performed by three surgeons, half by robotic assistance and half by human assistance. Human assistance was provided by an inexperienced assistant. No significant difference in operative times were found. The camera robot was faster in horizontal and zooming motions, but slower in vertical ones. Less errors in camera motion were observed and the optics had to be cleaned less often in case of robot assistance. The authors qualify their results with respect to the human assistant by stating that "co-operation between the surgeon and the camera assistant is significantly enhanced when the latter becomes familiar with the

[21] Full surgical removal of the prostate gland with access outside the abdominal cavity.

anatomy and the procedure. Thus, the results of the present study may have been different if assistants experienced in camera-holding were compared with the FreeHand".

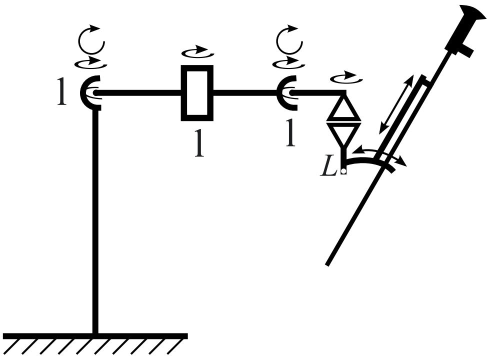


Figure 2.31: Kinematics of the FreeHand. The first three joints l are mechanically locked during setup time. The Laser L is used to register the robot's remote center of motion to the camera trocar.

2.1.19 EVOLAP

Herman et al. [142] developed the EVOLAP endoscope holder (Fig. 2.32) in 2009 with the objectives of high rigidity, low setup time, small footprint at the OR table and flexible placement. The EVOLAP robot is mounted at the OR table. It consists of a parallel mechanism for horizontal and vertical motion located at the edge of the OR table and a linear axis for insertion depth located at the endoscope (Fig. 2.33). The translation of the endoscope handle piece results in a change of orientation due to the combination of the trocar point acting as a pivot point and the two passive joints at the endoscope. A miniature analog joystick that is attached to a laparoscopic instrument is used to control the EVOLAP holder. Fixed and proportional speed control was implemented. Appropriate coordinate system for the joystick control are discussed by the authors and compared those of ViKY, AESOP and LapMan. An initial in vivo experiment of laparoscopic salpingectomy[22] was successfully completed with EVOLAP assistance.

Figure 2.32: The EVOLAP endoscope holder (Source: [143]).

A later study by Herman et al. [143] compares the motion performance of EVOLAP to ViKY and AESOP in box trainer tasks. Also the influence of the mapping from joystick to robot is evaluated. AESOP was found to have the most efficient and intuitive control due to the coordinate frame presented to the user. This was particularly apparent in endoscope positions close to the vertical. Independent of the robot kinematics, a joystick interface that does not only allow three directions of motions, but also diagonal ones was found to improve task performance.

[22]The surgical removal of a Fallopian tube.

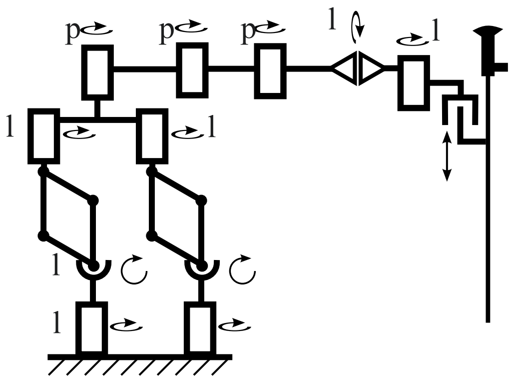


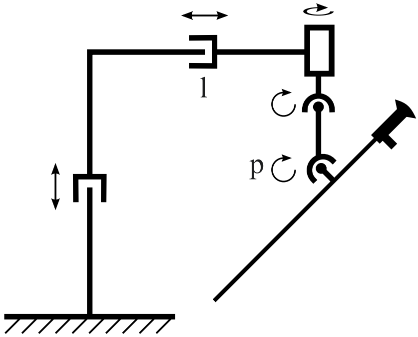
Figure 2.33: Kinematics of the EVOLAP endoscope holder. Only three joints are actuated. Most joints are freely moving passive joints p or mechanically locked during the setup l.

2.1.20 RoboLens

RoboLens is a motorized endoscope holder published by Mirbagheri et al. [144] in 2010. The endoscope positioner is mounted on a cart that is positioned around the OR table. The RoboLens kinematics consist of one linear, two active and one passive rotational joint (Fig. 2.34b). All joints are active, the trocar point is maintained in the robot controller. The authors provide an analysis of the workspaces and manipulability of the RoboLens kinematics. RoboLens can be controlled by a foot pedal and voice commands.

After evaluation of the trajectory following behavior in a lab setting, clinical trials were conducted. RoboLens assisted several surgeons in 30 laparoscopic interventions, most of which were cholecystectomies. The overhead boom allowed to position the bulk part of the endoscope holder away from the surgeon and occupying only the workspace over the camera trocar from above. Although no further details are reported, the authors note that the camera image was stable and no camera cleaning was required. Furthermore, the overall procedure time was lower with robotic camera assistant. Due to issues with the reliability of voice recognition, some surgeons preferred the foot

(a) The RoboLens endoscope holder (Source: [144]).



(b) RoboLens kinematics with one passive joint p. The joint marked as l is not in active use intraoperatively.

Figure 2.34: The RoboLens endoscope holder.

pedal interface. At the same time the foot pedal was found to be problematic because of the multiple foot switches that are already used in the OR, e.g. for electrocautery.

2.1.21 Tadano et al.

Tadano et al. [145] introduced a pneumatic laparoscope holder in 2015 (Fig. 2.35). The design is based on a pneumatic surgical manipulator (IBIS IV) developed by the same team [146]. The holder features a remote center of motion, which is implemented by a parallel link mechanism. All four degrees of freedom possible through the camera trocar (cf. Fig. 1.10) are pneumatically actuated. The device is placed over the trocar point and fixed there by mechanical links. The complete structure consists of two rotational joints, one linear and one rotational joint (Fig. 2.36). The second rotational joint is made up of a slider-crank mechanism and actuated by a pneumatic cylinder. Pneumatics are favored for their inherent compliance due to their compressibility and thus safety as well as the possibility to build a compact and more lightweight manipulator. In combination with pressure sensors and encoders, a sufficiently high position accuracy is achieved.

Figure 2.35: The pneumatic endoscope holder by Tadano et al. (Source: [145]).

The endoscope holder is operated through head motions. These are not tracked externally, as with the EndoAssist and FreeHand, but only based gyroscope measurements of two inertial motion sensors (IMU) worn by the surgeon. Robot motion through head movement is only enabled while a foot switch is pressed. Relating to human machine interfaces, the use of a head mounted display (HMD) instead of a conventional monitor is discussed.

Experiments in a box trainer were conducted. The task completion time for simple grasping tasks was similar between human and robot assistance. Finally, the system was also tested in vivo with three gastric resections on pigs. The interventions could be completed with the robotic assistance without lens cleaning.

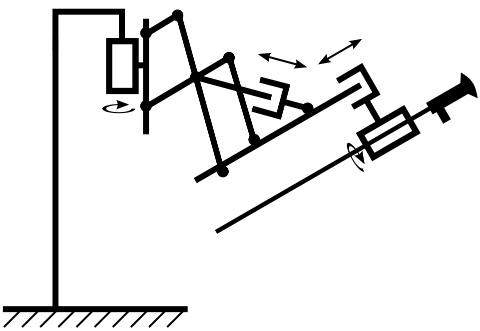


Figure 2.36: Kinematics of the pneumatic endoscope holder. Actuators are two cylinders and two vane motors.

2.1.22 Further systems

The following endoscope holders are summarized in this section because only brief descriptions of the systems are available.

A survey of surgical robotics by Pott et al. [147] from 2005 mentions the Compact Laparoscopic Endoscope Manipulator (CLEM) based on pneumatic muscles and controlled by voice or foot pedal. In addition, the PAROMIS is mentioned, which is based on hexapod kinematics, features force torque sensing and is controlled by speech or touch screen.

In 2009 Kraus et al. [148] describe a system named DeltaScope consisting of Delta kinematics that provide three degrees of translational motion. The trocar serves as fixed point, which alters the translational motion into spherical motions around the trocar point. DeltaScope is operated by joystick. The system was tested in tympanoplasty[23].

The AutoLap image-guided laparoscopic system, Medical Surgery Technologies Ltd., is a compact endoscope holder mounted to the OR table. It provides manual control and instrument tracking based on

[23] Surgical reconstruction of the eardrum or small bones.

visual servoing (cf. 2.2.3). Unfortunately, no system description is available in the literature.[24]

Another commercial system is the HIWIN endoscope holder (MTG-H100), HIWIN Technologies Corp., can be used on a trolley or mounted to the OR table. It features a remote center-of-motion and is controlled by a foot pedal. Again, no publications on the system could be found.[25]

2.1.23 Telemanipulation Systems

Although all telemanipulation systems provide an actuated endoscope holder, these systems will not be discussed in detail here for two reasons: First, the scope of this thesis is an assistance system working at the OR table alongside the surgeon. Therefore, surgical telemanipulators in terms of endoscope holding capability do not fall into the intended application domain. Secondly, in a telemanipulation setting the surgeon is already located away from the patient and uses special input devices to control the instruments. Switching between instruments and camera control does not pose as big of an additional obstacle as in conventional MIS. Nevertheless, the knowledge-based endoscope guidance system presented in the following chapters could be directly applied to remote controlled surgical robots.

Rassweiler et al. discuss [149] the cost trade-offs associated with robotic assistance systems for MIS, such as endoscope holders, compared to full telemanipulation setups. Three recent perspectives on remote controlled surgical systems can be found in [150], [151] and [152]. A brief overview of the most well-known telemanipulation systems with a strict focus on their endoscope holding capability and control interface for the endoscope follows.

[24] Information about the system is available in a brochure by the manufacturer: `http://mst-sys.com/wp-content/uploads/2014/11/AutoLAP_Brochure.pdf`.

[25] Information was taken from HIWIN's medical product catalog: `http://www.hiwin.tw/download/tech_doc/robot/Medical_Equipment(E).pdf`.

da Vinci

The da Vinci [153][33], Intuitive Surgical Inc., is by far the most clinically used surgical robot.[26] The 3D endoscope is mounted to an arm identical in kinematics to the instrument arms, i.e. with a remote center-of-motion and a linear joint for insertion depth. Camera control is performed at the surgeon console by switching control of the input devices from instruments to camera by a foot switch. While the camera is moved the instruments are stationary and vice versa. Nagi et al. developed a low-cost camera arm for the da Vinci Research Kit (dvrk) [155], named CALap.[27]

ZEUS

The ZEUS robotic surgical system [156][157], Computer Motion Inc., was a clinically used system between 1998 and 2003. It consists of three independent manipulators at the patient side. Two of these are remote controlled through hand guided input devices. The third robot is the AESOP (2.1.2) which is controlled by the surgeon via voice using a headset. From a system perspective, the AESOP works as a standalone component. ZEUS was used in the famous first transcontinental robot-assisted laparoscopic cholecystectomy, also known as the "Lindbergh operation" [158].

MiroSurge

The MiroSurge [35] developed by the German Aerospace Center (DLR) is a research system consisting of three independent lightweight robots at the patient side. These robots, named MIRO, share a common ancestry with the KUKA lightweight arms (cf. 4.2.4), however they are much smaller with a reduced payload. Two of the robots are used to position surgical instruments, named MICA, which provide two additional DoFs inside the patient. The third robot positions a conventional endoscope attached to it. The endoscope position and

[26] As of September 2015, nearly 3.500 units are installed (Source: [154]).

[27] The da Vinci Research Kit only comprises two manipulator arms by default.

orientation is completely determined by the robot position including rotation along the shaft. Control of the robots is performed through two haptic input devices, Omega.7, Force Dimension. In normal control mode each input device is mapped to one of the instrument robots. To move the endoscope, one input device is used in alternation with instrument control.

RAVEN

The RAVEN surgical robot system [159], Applied Dexterity, is an open research platform consisting of two patient side manipulators with remote center-of-motion kinematics. A particular research focus is long distance telesurgery, e.g. for astronauts on long-term space missions. However, currently only two arms for instrument positioning are part of the RAVEN. Experiments either happen with direct visual observation or use a human camera assistant.

CoBRASurge

The Compact Bevel-geared Robot for Advanced Surgery (CoBRASurge) [160][161], is a research system consisting of multiple small robot manipulators. As the name already indicates, CoBRASurge employs a special kinematic based on bevel gears to achieve spherical motion around the trocars in a compact manner. Each manipulator features four degrees of freedom that correspond to those available by the trocar constraint (cf. Fig. 1.10). Instruments and Endoscope are controlled in the same manner by a joystick.

Telelap Alf-X

A recent telesurgical platform is the Telelap Alf-x [162] approved for clinical use in 2012. It has been evaluated in about 150 gynecological interventions. The Telelap Alf-X consists of several large manipulators mounted on wheeled carts that can be individual positioned around the OR table. Long overhead booms allow to move the bulk of the required floorspace for each unit away from the patient, enabling good intraoperative access to the patient. The surgeon controls all robots, at least two for instruments and one for the endoscope, from a surgical console. The instruments are controlled by hand-operated input devices. For control of the 3D endoscope, two interfaces are available: First, an eye-tracking system, which centers the endoscope at the point that the surgeon is looking at on the screen while he presses a button on each input device. Second, zoom and insertion depth change through forward and backward head motion.

2.1.24 Summary

In 1998 Dunlap et al. [163] answer the question whether robotic arms are a cost-effective surgical tool very positively: "The study results indicate the robotic arm not only outperforms human camera holders, but also reduces laparoscopic surgical operating time, resulting in improved efficiency and cost savings to the institution."

Studies that compare human assistance to one of the systems above as well as those that compare two of these have already been described in the previous sections. Arezzo et al. [164] compare the motorized endoscope holders EndoAssist, AESOP and FIPS and the passive mechanical ones, TISKA, Martin arm and Endofreeze in a phantom study.[28] In their results, human assistance always performed best in terms of operative time and all passive systems were superior to the motorized ones.

In a survey on camera and instrument holders from 2004, Jaspers et al. [165] summarize about 70 publications related to passive – PASSIST, Tiska, Martin arm, Unitrack, Ball trocar, Leonard arm – and motorized – AESOP, EndoSista, LapMan, FIPS, Image track camera – endoscope holders. Overall, the literature they surveyed "recognized that using a camera holder in laparoscopic surgery provides an optimal and stable image of the operation field compared with human assistance. The control of the endoscope by the surgeon him/herself is also generally considered to be superior to human assistance. There was no difference in these respects between the

[28] The same publication also introduces Endofreeze.

passive and active (robotic) camera holders." At the same time, all interfaces for manual control have clear limitations: There are already multiple foot pedals present in the OR, both hands of the surgeon are already occupied rendering joysticks unergonomic, voice control only allows discrete motions and head control requires artificial head motions. The authors conclude that passive holders appear to work as well or even better than motorized ones and are more cost-efficient.

Feussner et al. state in the context of an overview of technical and digital advancements of surgery from 2014 [166] that motorized endoscope holders clearly failed to gain wide acceptance in the past. Yet, newer systems, such as SOLOASSIST and ViKY, are being reevaluated for their clinical benefits.

In summary, a multitude of endoscope robots have been built in the past 30 years. Many of these were evaluated in animal experiments. Several were evaluated in clinical studies. Nearly ten have been commercialized. Especially for the latter ones, a large number of studies have been published. Table 2.1 summarizes the endoscope holders that were covered in detail above. For the thesis at hand, four aspects are import to note:

- No motorized endoscope holder has reached widespread clinical use.
- Studies clearly show the potential for ergonomic improvements in MIS through usage of endoscope robots.
- A lot of research focused on the evaluation of various kinematics.
- Many interfaces for endoscope control have been researched: foot pedals, joysticks, voice, head, gaze, image pointing and position memory. However, all follow the paradigm of manual control by the surgeon (cf. Fig. 1.8).

Table 2.1: Overview of the motorized endoscope holders described in this section. For joints the following abbreviations are used: Revolute R; Twisting T; Linear L; Universal U; Parallel mechanisms P, e.g. four-bar linkages; and C-arch mechanisms C. Passive free moving joints are marked with a subscript 'p', e.g. R_p. Intraoperative locked joints are marked by a subscript 'l'. $U_l R_l U_l$ often is a standard mechanical instrument holder arm. Note: Telemanipulators and endoscope holders that were not described in detail (2.1.22) are not listed in this table.

Year	System	Robot	Kinematics	Control	References
1993	Endex	Custom	$(5R_p)B_pL$	Foot pedal	[53]
1993	EndoSista	See EndoAssist			[81], [82], [75], [85]
1994	AESOP	Custom	$LRRT_pR_pR$	Backdriving, Foot pedal, Joystick, Position memory, Voice control	[54], [58], [60], [24], [56], [57], [59], [52], [55], [65], [87], [61], [25], [67], [64]
1995	Begin and Hurteau et al.	CRS A460	$(6R)U_p$	Joystick	[68], [69]
1995	LARS / PLRCM	Custom	$LLLPP$	Joystick, Hands-on, Image pointing	[71], [72], [73]
1995	HISAR	Custom	$TRRRTR$	Joystick	[74]

1996	FIPS	Custom	$L_pR_pL_pRCLR$	Finger-ring Joystick, Voice control	[88], [91], [92], [93]
1998	EndoAssist	Custom	LRT	Head-Tracking	[78], [83], [76], [61], [77], [84], [86]
2000	Munoz et al.	Stäubli RX60	$(6R)U_p$	4 DoF Joystick, Voice control	[95], [96]
2001	ERM / Munoz et al.	Custom	$L(2R)(2R_p)$	3D mouse, Voice control	[98], [97], [99], [100], [101]
2002	LER / ViKY	Custom	$U_lR_pU_lRRL$	Backdriving, Foot pedal, Voice control, Position memory	[103], [102], [104], [105], [106], [108], [111], [110], [107], [110], [109]
2003	Naviot	Custom	$P(2R_p)L_p$	Hand controller	[112], [113], [114], [116], [115], [117]
2004	LapMan	Custom	$L(2P)LU_p$	Joystick	[127], [130], [129], [128]
2005	SWARM	Custom	$LRLR$	Foot pedal, Voice control	[131]

2006	COVER	Custom	$U_l R_p U_l RLR$	Head-tracking	[132]
2006	SMART / P-arm	Custom	Stewart-Gough	Joystick	[133], [134], [135], [136]
2007	Tonatiuh II	Custom	$R_l L R_l T R_p L$	Joypad	[137]
2008	PMASS	Custom	$R(2R_p)$	Foot pedal and body posture	[138][139]
2009	SOLO-ASSIST	Custom	$T(2R)(2R_p)$	Backdriving, Joystick	[118], [123], [124], [120], [121], [122], [119]
2009	FreeHand	Custom	$U_l R_p U_l RCL$	Head-Tracking	[141]
2009	EVOLAP	Custom	$RPR_p(3R_l)$ $T_p R_p L$	Joystick	[142], [143]
2011	RoboLens	Custom	$LL_l RR_p$	Foot pedal, Voice control	[144]
2014	Tadano et al.	Custom	$RPLR$	Head-Tracking	[145]

2.2 Approaches for Automated Camera Guidance

The previous section comprehensively surveyed motorized endoscope holders. The main focus was on robotic aspects, such as kinematics, the human machine interfaces and suitability for endoscope positioning. Several of these systems will reappear in this section, however, not as motorized endoscope holders under manual control, but as active robots (cf. p. 12). The question is how can positioning of the endoscope – by means of any motorized endoscope holder – be performed autonomously without continuous manual control by the surgeon. As discussed above (cf. 1.1.2), manual control further increases the mental workload of the surgeon. Pandya et al. [167] summarize the problem: "[It] is difficult for the user to obtain optimal camera viewpoints in a dynamic environment or to react effectively to irregular events in the scene due to task overload, latency issues, and complex camera positioning issues. The surgeon has to continually manage the camera position to achieve effective viewpoints through manual control or verbal communication with a second operator."

First related research areas to the automation of endoscope positioning will be briefly discussed, common features and differences will be highlighted. Afterwards a classification of approaches to automated camera guidance is presented. Based on this classification, the approaches known in literature are surveyed.

2.2.1 Related Problems

Active Perception [168] explores strategies for active acquisition of perceptional information. Instead of relying on a sensory system with fixed properties, expertise about sensors, reproduction properties and the information gathering task are employed to create a system that actively adapts to external circumstances. Which information is gathered does not simply depend on contingencies about the external makeup of the environment, but is gathered purposefully with respect to the data acquisition task.

For visual perception, Active Vision [169] looks at an observer that is able to change geometric parameters of the scenery. In relation to robotics, this usually means that the visual sensor can change position within the environment or manipulate the environment by rearranging objects. Many visual ambiguities about object properties can be easily resolved, if sensory information from multiple direction can be captured. A simple perspective example is shown in Fig. 2.37: Resolving the spatial ambiguities without getting different perspectives must rely on very minute details or can be impossible altogether. If the scene can be explored, i.e. viewed from different points, the actual spatial relations can be recognized.

The problem of Next Best View [170] is a sequence of view points to achieve a given vision task. For example, it is not possible to obtain a complete spatial model of most scenes from a single view point due to occlusions. “Next” is part of the problem name because the time component of the problem becomes important when executed on physical systems that have to move through space. Not only the more apparent minimization of total path length is a reason, but also that some view sequences could not be possible or dangerous.

Active Perception, Active Vision and Next Best View are insofar quite different from automated camera guidance that problems are often posed as (a series of) offline problems[29]. Camera guidance is an assistance function performed in real-time alongside the actions of the surgeon. Thus, the target objective of the perception task change together with the surgical task. To some extent the same holds true for the environment, which is constantly changing over the course of the intervention. This is especially true with respect to the surgical instruments, which are the most salient feature of task and environment.

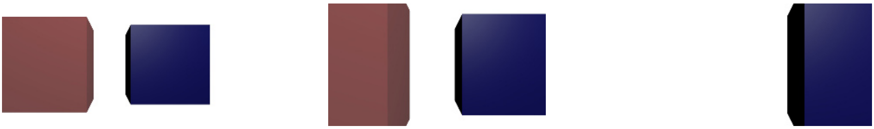
Another related problem is control of virtual cameras in computer graphics, e.g. computer games. However, there are many differences between positioning a virtual camera and a real (endoscope) camera. First of all, in the virtual environment the camera placement algorithm has access to the ground truth about the position of all scene objects. Secondly, the virtual camera is not constrained by physical limitations, for example in terms of (re)positioning velocity. Thirdly, it is assumed that the optimal camera position is already

[29]The term ’offline problem’ is derived from the distinction between offline and online algorithms. Offline algorithms have access to the full problem statement from the beginning, e.g. the input data. Online algorithms on the other hand receive the input piecewise while doing computation on the already received data.

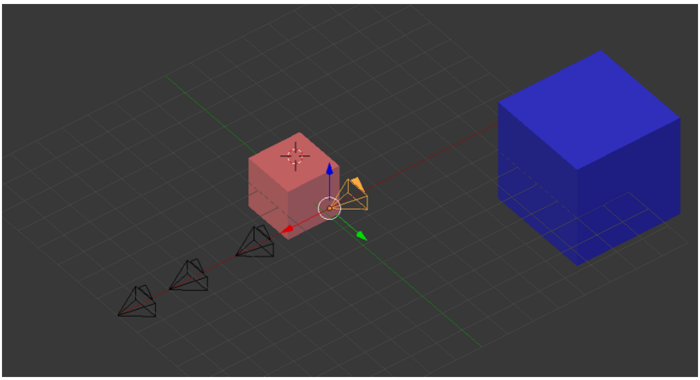
specified by a set of constraints or a known cost function. Therefore, none of the approaches can be directly applied to real-world endoscope positioning. The overview paper by Christie et al. [171] provides many more details on this subject.

As shown in the survey by Kober et al. [172] the field of reinforcement learning is widely applied to learning tasks in the domain of robotics. In reinforcement learning the system explores different strategies and receives rewards depending on its task success. However, the paradigm seems unsuitable to train an endoscope robot for the following reasons: First, the reward function must be engineered before the system can start to learn. Yet, this either presupposes already having a model of good endoscope positions, which is actually part of the learning task, or requires the use of very coarse reward functions. In the latter case, too many interventions would be required until the system behavior converges. In general, even if safety issues would be solved, there are significant practical obstacles with respect to the number of interventions required for training. Most case studies on reinforcement learning in robotics require on the order of 100 episodes to learn short dynamic tasks. This poses a big obstacle to apply the methods in the context of surgical robotics research, where recording and annotation of 10 interventions is already a major effort. In contrast, as shown in chapter 7, already five recorded interventions can be enough to train the system presented in this thesis for a 25 minute intervention.

(a) From this perspective the cube on the left appears to have the same size as the cube on the right. Both seem to be next to each other and have the same distance from the virtual camera.



(b) From left to right the observer approaches the two blocks straight ahead. It soon becomes obvious that the blocks are neither next to each other nor have the same size.



(c) Orthographic view of the actual scene: The left block is half the size and far in front of the right block. The four positions of the virtual camera for the images above are also indicated.

Figure 2.37: Illustration how active vision can easily resolve scene ambiguities that cannot be easily resolved from a single perspective.

2.2.2 Classification of Approaches

For the survey of current research approaches to the automation of endoscope guidance in the following section, a systematic classification is put forth in this section. In a recent survey paper on camera viewpoint automation Pandya et al. [167] propose such a classification. Their classification will be described next. However, due to the limitations of this classification, a novel extended classification is then proposed and argued for.

The top-level partition proposed by Pandya et al. is between reactive, proactive and combined control strategies (cf. Fig. 2.38):

- *Reactive*: The camera moves in direct response to changes in tracked data, such as instrument position or the surgeon's gaze.
- *Proactive*: Given the current state, preexisting domain knowledge is employed to predict the best camera position.
- *Combined*: Control has aspects of reactive and proactive control.

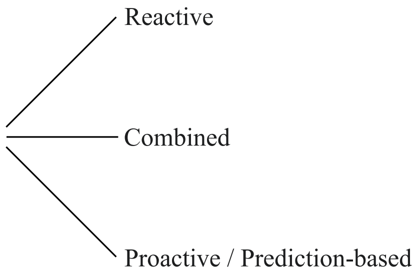


Figure 2.38: Classification of viewpoint automation proposed by Pandya et al.

The classification for automation of camera positioning proposed here is two-dimensional and structured as follows (cf. Fig. 2.39):
First dimension (A):

- *Direct*: A simple algorithmic mapping between current features and camera motion. Features can be simple, e.g. position of instrument tip, or complex, e.g. surgical gestures[30].
  - *General-purpose*: The algorithmic mapping can be directly applied to other camera positioning problems.
  - *Domain-specific*: The algorithmic mapping was specifically designed for the task domain and does not easily translate to other tasks.
- *Model-based*: Camera motion depends on a model that is part of the algorithms input data, i.e. given the same scene information, but with a different model will significantly alter the camera motion.[31]
  - *Reactive*: The resulting motion depends only on the current point in time.
  - *Planned*: The resulting motion depends on a time interval. This can include a memory of past states or planning ahead by means of a simulation model.

Second dimension (B):

- *Constant*: The system always uses the same mapping between input and camera motion.
- *Context-aware*: The system changes the mapping depending on a high-level property such as intervention type or surgical task.

Thus an approach must be assigned one class from each dimension A and B, e.g. an approach can be "constant, direct general-purpose" or "context-aware, model-based planned".

[30] A surgical gesture, or surgeme, is located between the task level, e.g. suturing, and motion primitives (cf. [173]).

[31] To explicitly deny the loop hole of simply loading a 'direct' program as data (into an interpreter) and referring to this as 'model-based', the following restriction must hold: The loaded model must not be Turing complete.

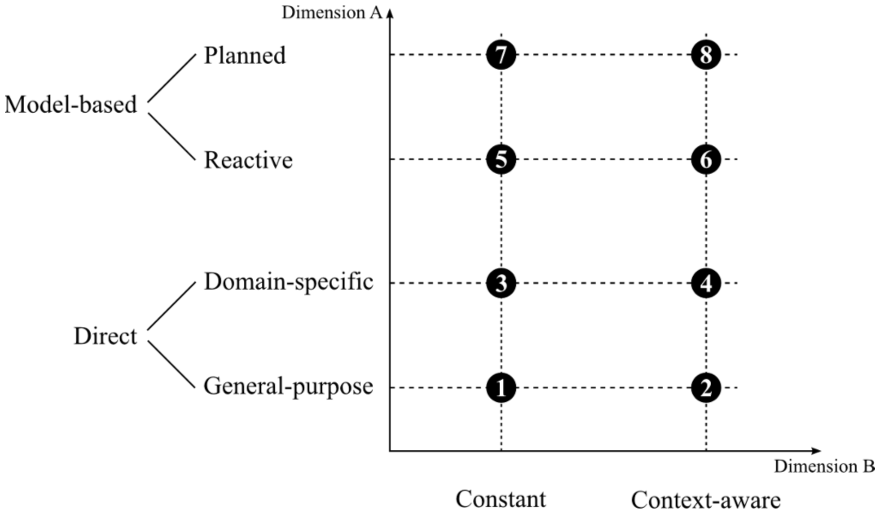


Figure 2.39: Proposed two-dimensional classification of camera automation.

2.2.3 Survey of Approaches

Employing the classification proposed in the previous section, existing approaches to camera automation will be summarized and classified in this section.[32] In order to keep the survey focused and at a reasonable length, papers that deal with the automation of other laparoscopic tasks, e.g. suturing, are excluded, even though some apply similar methods. Publications are sorted into the categories and described chronologically within each category.

Direct General-purpose (1,2)

Visual Servoing Visual Servoing is a particularly prominent direct general-purpose approach. Therefore, this approach will first be described in general terms without reference to specific publications. Afterwards, all publications that employ visual servoing can

[32]The numbers of the following subsections correspond to those in Fig. 2.39.

be described more concisely with reference to this general description. Publications that only treat instrument detection and tracking, but do not apply it for automated endoscope positioning are briefly discussed in 5.2.1.

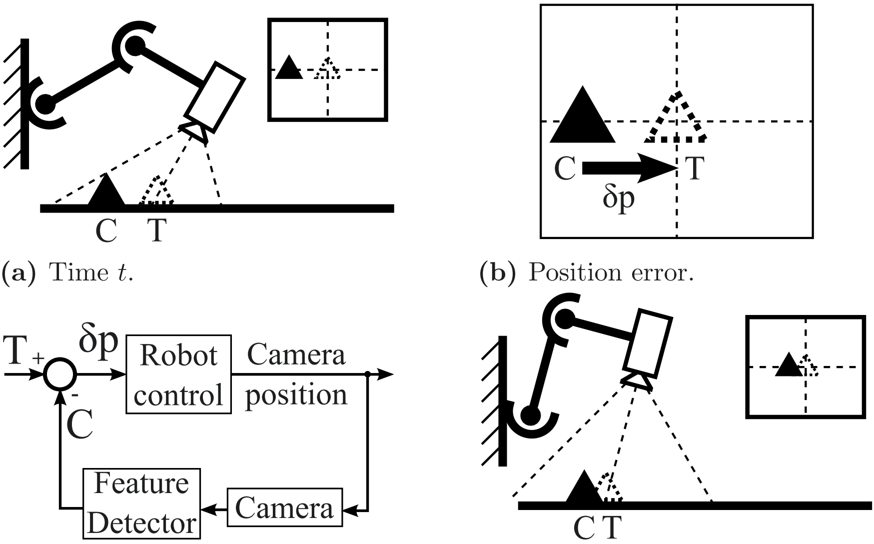
Visual Servoing refers to the process of tracking a feature in an image stream, comparing the feature's current position to a desired position and moving an actuator in order to reduce the distance between current and desired feature position. Fig. 2.40 illustrates this principle. In order to reduce the alignment error in each timestep[33], the kinematical relation between change of robot position and the camera's field of view must be known. If robot dynamics can be ignored in the outer control loops – which is assumed here given the low accelerations and velocities required for endoscope motion, often a velocity-based control approach is employed. Given a transformation between robot and camera coordinates, the time derivatives of the image coordinates can be written as a Jacobian matrix J. The (pseudo-)inverse of the Jacobian J^{-1} can then be used to compute the robot motion for the next time step δq from the alignment error in the image δp.

For more details, Azizian et al. [174][175] provide an extensive survey on visual servoing in medical robotics. An older but complementary article is available by Groeger et al. [176].

Approaches The pioneering work in 1994 by Lee and Uecker et al. [177][178][179] implemented markerless visual servoing on the AESOP (cf. 2.1.2). The term 'markerless' refers here to the use of standard laparoscopic instruments that were not modified in any manner, e.g. through the addition of colored rings at their distal end.[34] Further information on the instrument detection and tracking can be found in section 5.2.1. Given the position of the instrument

[33]Cameras typically provide 30 images per second, assuming features can be detected at this frame rate of 30 Hz, the gap to the much higher robot control rates (100 - 1000 Hz) must be bridged by interpolation. The arising problems of cascade control will not be addressed here.

[34]A corresponding distinction is between natural and artificial landmarks.



(a) Time t. **(b)** Position error. **(c)** Control loop sketch. **(d)** Time $t+1$.

Figure 2.40: A simple example of Visual Servoing. The robot position is changed in each timestep in order to make the current feature position C coincidence with the target feature position T.

tip, the authors propose a velocity controller based on the inverse Jacobian that keeps an instrument tip in the center of the image. Only left/right and up/down motion was implemented, insertion depth is manually controlled by the surgeon. The approach is classified as general-purpose because the model is directly applicable to visual servoing tasks outside surgery. The system is not context-aware.

A US patent by Kudo et al. [180] from 1998 describes an endoscope positioning system that tracks instruments by means of color markers or by electromagnetic three-dimensional position sensors. Servoing of orientation and zoom is performed. Furthermore, the system comprises different manual control interfaces such as a joystick, head-tracking and position memory.

Zhang et al. [181] present monochrome markers for instrument tracking in 2002, realized as strips along the endoscope shaft. They elaborate on the peculiarities of camera calibration for endoscopes because the marker shape together with an intrinsically well calibrated endoscope can provide depth information from monocular images. Experiments showed that the markers were robust to specular reflections and could be tracked well. Unfortunately, no experimental results for the visual servoing performance are provided.

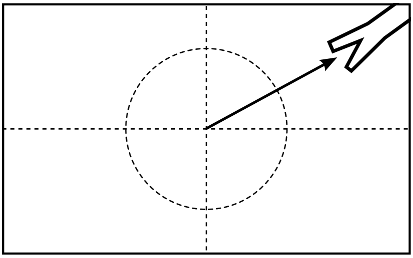
The master thesis of Sivadasan [182] from 2006 compares the mental workload and task completion time between eye gaze and joystick control in a simple lab setup. The subjective workload according to the NASA Task Load Index (TLX) questionnaire for a simple camera aiming task was lower for gaze control. Task performance was very similar to joystick control.

Bourger et al. [183] apply the model-free efficient second-order minimization (ESM[35]) tracking algorithm based on histogram matching. The method is insofar MIS specific as the trocar constraint (Fig. 1.10) is incorporated into the servoing control. However, all other aspects can be directly applied to other visual servoing tasks. In particular, the histogram-based tracking is directly applicable for tracking outside of the surgical context. Results with AESOP (cf. 2.1.2) in a box trainer showed relatively slow convergence attributed to a significant amount of noise in the histogram matching, non-linearities in camera positioning and the high latency of the AESOP controller.

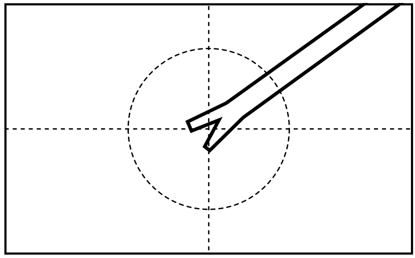
Polski et al. [185] describe servoing of the SOLOASSIST (cf. 2.1.11) by means of electromagnetic trackers in 2009. Markers are attached to the endoscope camera and the instrument. The servoing task evaluated in a test box was to keep a fixed distance between the endoscope tip and the instrument tip. Positioning accuracy was between 10 mm and 30 mm depending on size of the test target. Motion hysteresis

[35] Mali [184] formulates visual servoing as minimization problem and employs results of the latter to the former. In analogy to the use of a first-order Tayler series approximation in the Gauss-Newton method (GNM) for minimization, a second-order Taylor series is used for ESM.

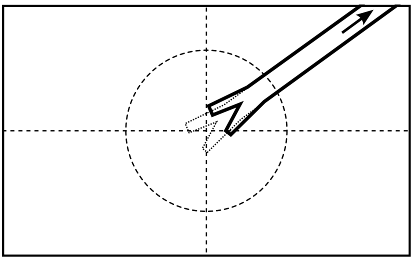
(Fig. 2.41) was used to keep the endoscope position stable for small instrument motions.



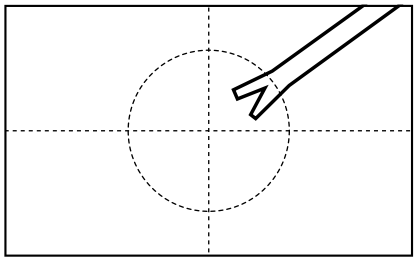
(a) Instrument distance above hysteresis threshold (dashed circle): Camera moves.



(b) Camera has centered on instrument tip – assuming this being the optimal position.



(c) Instrument moves, but new distance below hysteresis threshold.



(d) Camera remains still until instrument leaves hysteresis region.

Figure 2.41: Illustration of motion hysteresis to stabilize camera position.

In 2010 Osa et al. [186] implement visual servoing of the endoscope on a telemanipulation system, named Endo[PA]R[36]. Special attention is given to a modified interaction matrix that takes the coupling between linear and angular velocities into account, which is due to the trocar constraint (Fig. 1.10). Apart from this, standard visual

[36]The Endoscopic Partlally-Autonomous Robot (Endo[PA]R) system [187] consists of four ceiling mounted Mitsubishi MELFA 6SL robots equipped with force-torque sensors.

servoing with color markers on the instruments is used for automatic camera control.

Noonan et al. [188] describe gaze contingent control for an endoscope with additional degrees of freedom at the endoscope tip in 2010. Since the flexible instruments are outside the scope of this survey, only the control aspects that generalize to the control of a conventional rigid endoscope will be explained. A stand-alone eye tracker, Tobii x50, records the 2D fixation points on a monitor at 50 Hz. The surgeon must keep his head in a volume of $30 \times 16 \times 20$ cm in front of the screen. To control the endoscope motion, the screen is split into five regions: a central area and the four corner quadrants. After selecting a GUI button by fixation, the endoscope rotates towards the four image corners until the central area is fixated again. Motion speed is selected by the distance from the screen center. To rotate along the endoscope shaft the surgeon has to first fixate on a GUI button and afterwards move the endoscope by fixating on the left or right part of the screen. Evaluation was performed by a pointing task in a box trainer, but not compared to other control modalities.

King et al. [189] describe automatic instrument tracking in a lab test setup consisting of a box trainer and a pan-tilt-zoom webcam in 2013. Color markers are attached to the instrument tips. Based on the instrument tip position, the following rules have been implemented: If two instruments are visible, the camera is centered between their tips. Motion hysteresis is used to reduce the amount of camera motion as long as the tips are sufficiently close to the image center. If both instruments are close to the image center, the camera zooms in and out if they are too close to opposing image edges. No motion is performed if no or only one instrument is visible.

Lie and Zhang et al. [190][191] present gaze control of the CoBRASurge endoscope holder (cf. 2.1.23) in 2013. A stand-alone eye tracking system, S2 Eye Tracker, is used to acquire the gaze data at 60 Hz in a working volume of $25 \times 11 \times 30$ cm. After low-pass filtering of the raw gaze data, fixations lasting for 2 seconds are detected. If the surgeon confirms the intent[37] to move the endoscope by a foot clutch the fixation point is used to reposition the endoscope accordingly. The authors report that a simple pointing task in a box trainer could be performed by multiple users.

Building on their earlier work (cf. p. 2.2.3), in 2015, Li et al. [193] present gaze control that distinguishes natural eye movement from control intended eye movement based on fuzzy logic. The automatic classification between gaze for control and gaze for observation was evaluated by visual exploration tasks in a box trainer. Compared with the earlier approach based on pure dwell time, the fuzzy interface had a better average response time (1.4 to 2.1 seconds), but with a higher variation. In a subjective assessment conducted through a usability questionnaire, the fuzzy interface scored better (100 to 88.5 points on a scale from 0 to 130), yet, the dwell time interface was found slightly superior for repetitive tasks.

Direct Domain-specific (3,4)

In 1997 Wei et al. [194][195] published a visual servoing approach based on color markers attached to the instruments in combination with a stereo laparoscope.[38] The AESOP 1000 (cf. 2.1.2) robot was used for endoscope positioning. All three degrees of freedom are determined by the visual servoing control. The insertion depth, i.e. the distance to the instrument tips, was determined by means of stereo disparity of the detected color markers. Motion gain for small image displacements was set to a low value in order to have a stable static image position. Furthermore, the detected instrument positions were low-pass filtered in the temporal and spatial domain. So far the algorithm is general-purpose, however, additional domain-specific rules that address characteristics of MIS are employed, too: The robot will not immediately change the endoscope position, when a

[37] The problem with gaze control is that control is always 'on' because the surgeon does not only look at the screen in order to interact with it, but also to just see its content (cf. [192]).

[38] A patent for the procedure was filed in 1996 [196].

new instrument enters the field of view. Normal instrument motion is distinguished from motions such as retraction of the instrument by manual thresholds. In the latter case, the robot will not follow the instrument. However, the control rules are not context-aware and constant throughout the intervention. An evaluation in pigs showed that the system followed the instruments well, even if they were partially occluded, stained or vision was deteriorated by smoke caused by electrocautery[39]. The authors argue that adding color-marks to laparoscopic instruments would have negligible cost and not impact their surgical properties. In 1999 Omote et al. [197] report the application of the system, named Self-Guided Robotic Camera Control for Laparoscopic (SGRCCS), in 20 laparoscopic cholecystectomies and compared these to 58 human-assisted ones. Setup time of the system was 21 minutes on average, 83% of the procedures could be completed with SGRCCS assistance. Average operative times with the robot were 54 minutes and 60 minutes without robot assistance. Furthermore, a reduced number of corrective camera movements and lens cleanings are reported for robot assistance. The subjective assessment in comparison to human assistance favored the robot in 12 cases, found it to be equally good in three cases and worse in two.

The line of research started in 2001 by Nishikawa et al. [198] looks into head control of endoscopes by means of visual face tracking. After the initial publication the system was named FAce MOUSe, or FAMOUS. The first method described is based on iris detection from a grayscale image provided by a camera mounted to the monitor that displays the endoscopic video. A second method recognizes a black marker tape that is attached to the surgical cap. The following interaction method assumes that the surgeon remains in a constant distance to the screen and faces it in parallel. Pan and zoom motions are initiated by different face gestures. If the surgeon tilts his head first left then right the pan mode is initiated, if he performs these motions the other way round zoom mode is toggled. In each mode

[39]Cauterization, the selective burning of tissue, by means of electrically generated heat.

left/right and up/down rotation of the surgeon's head determines the relative motion speed of the endoscope. A finalizing head gesture stops endoscope control. An in vivo experiment, laparoscopic cholecystectomy on a pig, could be performed without a human assistant and no lens cleaning was required. During the 44 minute intervention, 97 face motions had to be performed by the surgeon. In 2002 the system was compared to a voice control interface [199] using the Virtual Laparoscopic Interface (VLI), Immersion Corporation, which simulates simple laparoscopic manipulations and is controlled by two mechanically tracked laparoscopic instruments and a foot pedal. Three tasks used in the evaluation were derived from actual surgical actions performed during cholecystectomy, but much simplified: retrieval of gallbladder, dissection and ligation of cystic duct and vessels, dissection of gallblader from liver. Suitability of the endoscope view was scored by rate of compliance to the rule that the target must be within a fixed distance to the endoscope and sufficiently close to the center of the image. No significant difference in task completion time were found. The amount of time spend for camera manipulation and endoscope path distance were smaller with face control. On the other side, instrument paths were longer and task errors were more frequent with face control. In an extended article from 2003 [200], the analogy between face control and mouse control, giving FAce MOUSe its name, is expounded. Also further details on the custom endoscope holder are provided, which employs kinematics that are similar to the LARS robot (cf. 2.1.4). In addition, more in vitro experiments are described. The system was also compared to instrument tracking in a two box trainer interventions with pig organs [201]. Qualitatively, instrument tracking was found to be problematic for very fine manipulations with high zoom.

Voros et al. [202][203] exploit the special geometric layout of MIS in the abdomen (cf. Fig. 1.2) in 2006. It is known that both camera and instruments enter the abdominal cavity through the trocar points located on the surface defined by the inflated abdominal wall. These points can be measured by means of a calibrated endoscope robot. The diameter of the laparoscopic instrument's shaft is also known.

Calibration includes intrinsic camera calibration and extrinsic, hand-eye, camera to robot calibration. The trocar points are viewed from two different positions by the manually controlled endoscope robot. Because the transformation between the recording positions is known, conventional stereo vision algorithms can be used to compute the 3D location from the 2D images. This information is projected on to the image plane and simplifies finding the instrument axis and instrument tip in the image. For experimental purposes the ViKY endoscope holder is used. Instrument detection results from a cadaver experiment show a detection error of less than 11 pixels for a 200×100 pixel image resolution in 87% of the images. However, systematic wrong detections are reported, too. No visual servoing results are provided.

The master thesis by Munduri [204] from 2010 implements two automatic camera modes on the ZEUS telemanipulator (cf. 2.1.23). Based on the known transformation between the three robotic arms deployed in the ZEUS system together with a registration of the trocar points, the instrument position is known from the forward kinematics. Given the endoscope and instrument position, a zoom mode and a following mode are implemented. The zoom mode calculates the angle between camera center and the instrument in order to keep it visible ($< 35°$) and not too close ($> 15°$). In following mode, the algorithm tries to keep the endoscope centered on the instrument. Motion speed is low in order to avoid too much motion. Switching between static, zoom and following mode is performed manually.

Song et al. [205] base their approach on tracking of color markers at the instrument tip. The image distance between instruments is used to determine the zoom ratio. Yet, the authors note that since this measure is based on the image distance only, the distance for instruments with different depths can be easily misjudged. As a workaround a second color marker was attached to the instrument that must be within the field of view, thus keeping a minimum distance between endoscope and instruments. Experiments were performed with a Webcam on a pan-tilt unit. Motion hysteresis is used for tracking stabilization.

Yu et al. [206] describe automatic positioning of the endoscope in the Laparoscopic Minimally Invasive Surgical Robotic System (LMISRS) telemanipulation system in 2013. Their algorithm is based on the known position of the endoscope and telemanipulated instruments. Given an initial view angle of the instruments, the authors elaborate on the kinematic transform required to keep the endoscope in the same view angle while the instruments move. No experimental results for the servoing scheme are provided.

In 2013 Fujii et al. [207] present gaze control of an endoscope by means of explicit gaze gestures. The authors favor gaze gestures because of the inherent intentional ambiguity provided by eye motion. Other works have to use an additional input channel, such as a foot pedal, for disambiguation in gaze control. Recognition of gestures is achieved by modeling each one as a Hidden Markov Model (HMM). The forward-back algorithm is used to assign a probability to each HMM for a given observation. The most likely gesture is then selected. Two gestures for zoom and pan were trained and evaluated. The system is evaluated in a box trainer with a KUKA LWR IV (cf. 4.2.4) that positions the endoscope. Cartesian impedance control is employed to provide force-torque constrained positioning. The trocar constraint is obeyed by specifying sufficient intermediate poses that are followed in a point-to-point (PTP) manner by the controller. Once an eye gesture is recognized, a GUI is overlaid on the laparoscopic image. The actual zoom or pan is then determined through the eye fixture on the screen. Activation of zoom and pan mode by gaze gestures was compared to activation by foot pedals. Gaze control was used in both cases. Eleven surgical residents took part in the box trainer experiments. The first task consisted of navigating the endoscope on a numbered grid. The second task was to remove a set of ‘lesions’ in an upper gastrointestinal phantom, which required bimanual manipulation. As control group to the gaze gesture and the foot pedal activated gaze control, conventional human assistance was evaluated. Assessment comprised task time and camera path length as well as the NASA TLX questionnaire on mental task load. Robot assistance in both modes was found to result in significantly shorter camera paths compared to human assistance. Task completion time did not differ between all three modes. Although the median score for mental load in case of foot pedal activation was lower, the difference was not significant. The score for robot and human assistance did not differ significantly either.

In their paper from 2014, Zhao et al. [208] combine the instrument markers by Zhang et al. (cf. p. 85) with the geometrical method by Voros et al. (cf. p. 90) for visual servoing with the ViKY endoscope holder (cf. 2.1.9). Given the acquired trocar points in 3D relative to the instrument trocar and an intrinsically calibrated endoscope, the 2D marker position in the image can be triangulated to 3D positions. The reconstructed positions are smoothed by a Kalman filter for tracking. Visual servoing of all three actuated degrees of freedom provided by the ViKY is performed. No experimental results beyond the tracking accuracy are provided.

Model-based Reactive (5,6)

In 1996 Casals et al. [209] describe situation-dependent camera control by means of visual servoing using tracked instruments that are marked with two straight lines along the tool axis. The tool diameter in the endoscopic image is used to approximate the distance between endoscope and tool tip. Three surgical situations are distinguished to have the same camera control rules: two moving surgical instrument in the scene; one moving and one relatively steady auxiliary instrument in the scene; only a single instrument visible in the scene. In case of two active instruments, the endoscope is positioned in the middle of the line connecting the two instrument tips. In case of one dominant instrument, the endoscope is also positioned on the line connection both instrument tips, however, not in the middle but closer to the dominant one. Finally, in case of a single instrument in the scene, the camera centers on the predicted next position of the instrument tip. Zoom is controlled to keep the instrument(s) in the field of view, but since the optimal zoom depends on the surgeon's preferences, control is done relative to an initial parameter set by the surgeon. Filtering is applied in order to smooth the endoscope motion. The implemented filter mixes motion hysteresis (cf. Fig 2.41) with a gain proportional to the position error. In case of large deviations that risk the tool disappearing from the camera view, direct tracking is used. Thus the approach can be characterized as context-aware and model-based reactive (see number 6 in Fig. 2.39). The system was evaluated in the operating room on a commercial robot [210], unfortunately, no detailed results are reported.

Nishikawa et al. [211] describe a system based on preoperative workspace planning combined with intraoperative instrument servoing in 2006. Preoperatively, the surgeon manually defines workspaces in the abdominal cavity which determine the desired zoom ratio. Intraoperatively, when an instrument enters the predefined workspaces the endoscope is centered on the instrument tip with the respective zoom ratio. Between multiple workspaces the zoom ratio is linearly interpolated. Instruments are externally tracked by a Polaris system, NDI. Qualitative results, from a laparoscopic cholecystectomy in a box trainer with pig organs, showed that the adaptive change of the zoom ratio was well received. However, the implementation, due to a multitude of defined workspaces within the working area of the instrument, changed the zoom ratio too frequently. Since the manual definition of workspaces takes surgical phases into account, this approach can be considered context-aware.

In 2008 Nishikawa et al. [212] present a system that acquires the optimal zoom ratio by observing previous interventions. By means of externally tracked instruments, Polaris, NDI, the distance and angle between the laparoscope and instruments were recorded in a box trainer. The authors look at the correlation between angle and distance and use this relation as control law in visual servoing. The angle-distance control law was compared to a fixed-distance control law in a box trainer for cholecystectomy on pig organs.

Rivas-Blanco et al. [213] present automated camera positioning based on a cognitive architecture in 2014. The lab setup consists of a Barret WAM arm with 7 DoFs, both tracked by an external marker-based tracking system, Polaris, NDI, together with a box as abdomen phantom. Instead of a laparoscope, the camera is held on the inside of the abdominal wall by a magnet at the robot end effector. The robot translates the camera along the abdominal wall instead of changing its orientation as for a conventional endoscope. However, the details of this hardware setup are disregarded in the following explanation of the control system. A cognitive architecture comprised of short-term and long-term memory is proposed by the authors. The long-term memory consists of a semantic memory that stores information about interventions and surgical facts, a procedural memory that stores learned behaviors and an episodic memory that stores experiences of the user. In the short-term memory the surgical state is estimated (cf. 5.2.2) and the focus of attention (FOA) is estimated that in turn triggers a motion behavior of the camera. The human-machine interface (HMI) enables the surgeon to adapt the manually the camera view. For each surgical gesture known to the system a trained Hidden Markov Model (HMM) is used for recognition. The position of an object in the camera image together with the zoom factor, referred to as FOA, is specified for each state of a surgical task. This mapping is individualized for each surgeon over time: Every time the surgeon gives a voice command to manually change the camera position, the modified relation is associated with the current surgical task. The system was evaluated by five non-experts performing a suturing task in the lab setting described above. After three trials with manual control of the camera position by voice, the same task was again performed three times with automated camera positioning. While a large reduction in the number of voice commands is reported, the overall task time was not significantly reduced.[40] The camera motion employs a model of the intervention, is context-aware of the surgical task and uses reactive visual servoing.

In 2014 Agustinos et al. [214] continue the approach by Voros et al. (cf. p. 90), making it context-aware and implementing it on the ViKY endoscope holder (cf. 2.1.9). Based the recognition

[40]Since the trials were not performed in random order, the learning which is to be expected in novices was not accounted for.

of the current surgical step in a workflow model, different tracking modes are activated:[41] With a single single instrument in the scene its tip is tracked with or without motion hysteresis (cf. Fig. 2.41). In case of several instruments present, either one of these is tracked, which assumes they can be reliably distinguished without markers, or the intersection point of their shaft axis is used as tracking target. Experiments were performed in a box trainer that showed that visual tracking converges, but without motion filtering the camera image can become unstable due to tracking noise. The latter is especially a problem when the intersection between two instrument is tracked due to the projective error enlargement.

Model-based Planned (7,8)

Wang et al. [215][216] present a system in 1996 that incorporates high-level models, planned scope motions and reactive motions based on image processing. The system comprises several conceptual layers (from top to bottom): supervising & planning, guidance & control, integration & coordination and sensing & modeling. The authors characterize the system as "choreographed" because sequences of actions can be executed semi-automatically. Initial trocar insertion is given as an example. The surgeons initiates the sequence by voice command, image analysis and object recognition are then used to ground each step. First, the endoscope is positioned towards the abdominal wall, visual analysis focuses on bulging parts in the image and recognizes when the trocar tip penetrates the abdominal wall. Different voice commands initiate manual motion of the endoscope or visual servoing to an instrument tip (cf. Lee and Uecker et al. on p. 83). The high-level plans allow to specify different reactive behaviors of the lower levels, thus the system is context-aware.

Starting in 2004, Ko and Kwon et al. [217][218][219] describe interaction with the endoscope robot[42] based on a surgery task model.

[41] The authors refer to known approaches for phase identification in the literature, autonomous mode switching is not actually implemented.

[42] The KaLAR robot featuring a custom bendable endoscope is used.

The robot should not only react to direct commands, but autonomously respond to the surgeon's actions and the current task. Although the main purpose of a task model is better human-robot interaction (HRI), it also facilitates environment perception. The surgery task model consists of states (surgical stages), transitions triggered by environment information and an action strategy associated with each state. Transitions can be probabilistic, accounting for variance in procedures. An example model of laparoscopic cholecystectomy is provided by the authors. For camera assistance, each tuple of state and instrument is manually assigned with the preferred camera mode, which can be either static, voice commands or visual servoing of an instrument. Instruments are augmented by different color markers that are used for tracking of the instrument and identification of its type. A further mode that the authors plan to implement is automatic centering on an anatomical region. The system was in vivo evaluated by three cholecystectomies in small pigs. Average operating time was 26.7 minutes, which is put in relation to the surgeon's average times in human cholecystectomies of 23.3 minutes. The automatic selection of the preferred camera mode was found to reduce the number of voice commands from 71 to 50.

The autonomous endoscope guidance described by Weede et al. [220][221] in 2011 builds on a knowledge base of recorded interventions from one intervention type. All trajectories undergo a clustering algorithm that distributes a predefined number of cluster centers over the data points that minimize the distance between points to their clusters. Once the clusters have been positioned and data points assigned to them, the covariance of each cluster is determined. In the next step, each trajectory is traversed and the transition between different clusters are recorded. The resulting transition probabilities between clusters in turn are used to define a second-order Markov chain that encodes a probabilistic trajectory through the combined interventions. In order to capture the important relation between the two instruments, the points corresponding to each instrument tip are represented as a combined six dimensional vector. Intraoperatively the current position of the instruments is assigned to a cluster by means of a maximum likelihood classifier. The Markov chain is then used to predict the most likely next cluster to which the instruments will move. Based on a set of manually defined rules for an optimal endoscopic field of view, the camera is moved, if the prediction finds one instrument outside this view. To increase motion smoothness, the resulting camera positions are spatially low-pass filtered. The authors evaluated the prediction performance in a box trainer where different landmarks hand to be touched in three defined sequences with laparoscopic instruments. On average the prediction was accurate with between 89.1% to 92.3% depending on the number of clusters specified. Compared to a purely reactive camera positioning, 29% less camera motions were required. The approach is model-based planned and constant as the relation between camera position and instruments is manually specified and independent of the current surgical task.

In an abstract from 2013 Bauer et al. [222] describe the "soloassist-cognitive" project. The project's aim is to combine eye tracking control and visually tracked instruments for partially autonomous camera positioning based on a model of the surgical workflow. Unfortunately, no further details on the project are currently available. Two of the Fundamentals of Laparoscopic Surgery (FLS) manual skill tests[43], namely peg transfer and precision cutting, were modified to require camera repositioning. The pegs were placed further apart and the circle was enlarged that had to be cut. No quantitative results are provided, but the authors report the automatic camera positioning to be well suited for the tasks. Furthermore, the authors recognize that the "presented camera system currently has a very basic movement scheme that essentially keeps both instruments in the camera's view at all times. This corresponds to the simplest instruction that could be given to a novice camera holder. Such a movement scheme may be adequate in some situations, but it is not ideal or sufficient for many tasks."

[43] These are comprised of peg transfer, pattern cut, ligating loop, extracorporeal suture and intracorporeal suture [11].

Summary

Table 2.2 summarizes all approaches detailed above. The majority of approaches either focuses on improving manual control modalities or looks into visual servoing strategies. Several approaches include context-awareness based on a model of the surgical workflow that modifies the underlying control strategies, e.g. a different zoom level is assigned to each surgical step. There is a clear correlation between model-based approaches and context-awareness. In terms of Fig. 2.39, the bottom-left (1,3) and the top-right (6,8) are dominant. Authors that recognize a need for endoscope control that is not derived from hand coded direct algorithms, but employs a model of the surgical task, also recognize the requirement for adaptation to different surgical tasks. In general, there is a lack of meaningful evaluation and comparison of different methods. Lacking a benchmark or at least a clear agreement on a protocol to evaluate these approaches, progress in the field of automated camera positioning or viewpoint automation cannot be objectively assessed. However, at least qualitatively, the experimental results referred above clearly suggest that autonomous robotic camera assistance in MIS has a realistic chance to

- improve the endoscopic view (positioning, image stability, independence from available camera assistant)
- reduce the OR staff required for simple laparoscopic interventions and free the assistants for other tasks in complex ones
- allow the surgeon to focus on his main tasks, without increasing his cognitive load for camera control.

At the same time, there are two issues with current approaches. They are either

- too constrained to represent the minute differences in positioning required for an optimal view or
- require extensive manual modeling of know-how for each surgical task in each intervention type.

Table 2.2: Overview of previous approach on automation of endoscope positioning. The number under class corresponds to those in Fig. 2.39. "VS" stands for Visual Servoing.

Year	Authors	Robot	Class	Method	Ref.
1994	Lee and Uecker et al.	AESOP (2.1.2)	1	Markerless VS	[177], [178], [179]
1996	Casals et al.		6	Task-dependent VS modes	[209], [210]
1996	Wang et al.	AESOP (2.1.2)	8	Action sequences with VS	[215], [216]
1997	Wei and Omote et al.	AESOP (2.1.2)	3	VS with color markers with explicit rules	[196], [194], [195], [197]
1998	Kudo et al.		1	VS with color markers	[180]
2001	Nishikawa et al.	Custom (cf. LARS 2.1.4)	3	Face control	[198], [199], [200], [201]
2002	Zhang et al.	Custom	1	VS with monochrome markers	[181]
2004	Ko and Kwon et al.	KaLAR	8	Step-dependent voice or VS with color markers	[217], [218], [219]

2006	Nishikawa et al.	Custom (cf. LARS 2.1.4)	6	VS with pre-operatively defined zoom ratio	[211]
2006	Sivadasan	AESOP (2.1.2)	1	Eye gaze; Joystick	[182]
2006	Voros et al.	ViKY (2.1.9)	3	Markerless VS with 3D from geometry	[202], [203]
2007	Bourger et al.	AESOP (2.1.2)	1	Model-free VS	[183]
2008	Nishikawa et al.	Custom (cf. LARS 2.1.4)	5	VS with learned angle-distance relation	[212]
2009	Polski et al.	SOLO-ASSIST (2.1.11)	1	Electromagnetic servoing	[185]
2010	Munduri	ZEUS (2.1.23)	3	Kinematics-based servoing	[204]
2010	Noonan et al.	Custom	1	Gaze control	[188]
2010	Osa et al.	Endo[PA]R	1	VS with color markers	[186]
2011	Weede et al.	LWR IV; RX90	8	VS with instrument path prediction	[220], [221]
2012	Song et al.	(Webcam)	3	VS with color markers	[205]

2013	Bauer et al.	SOLO-ASSIST (2.1.11)	8	Step-dependent voice or VS with color markers	[222]
2013	King et al.	(Webcam)	1	VS with color markers	[189]
2013	Lie and Zhang et al.	CoBRA-Surge (2.1.23)	1	Gaze control	[190], [191]
2013	Yu et al.	Custom Telema-nipu-lator	3	Kinematics-based servoing	[206]
2013	Fujii et al.	KUKA LWR IV	3	Gaze gestures and control	[207]
2014	Agustinos et al.	ViKY (2.1.9)	6	Step-dependent VS modes with 3D from geometry	[214]
2014	Rivas-Blanco et al.	Barret WAM	6	Surgical-Action-dep. VS; Offset learning	[213]
2014	Zhao et al.	ViKY (2.1.9)	3	VS color mark-ers & 3D from geometry	[208]
2015	Li et al.	CoBRA-Surge (2.1.23)	1	Fuzzy Gaze control	[193]

http://www.springer.com/978-3-658-14913-0

Learning Dynamic Spatial Relations
The Case of a Knowledge-based Endoscopic Camera Guidance Robot
Bihlmaier, A.
2016, XV, 267 p. 120 illus., Softcover
ISBN: 978-3-658-14913-0