Applications of Image Processing in Laparoscopic Surgeries: An Overview

Chapter · January 2018

DOI: 10.4018/978-1-5225-5204-8.ch063

Toktam Khatibi
Tarbiat Modares University

Mohammad Mehdi Sepehri
Tarbiat Modares University

Pejman Shadpour
Iran University of Medical Sciences

Chapter 14
Applications of Image Processing in Laparoscopic Surgeries:
An Overview

Toktam Khatibi
Tarbiat Modares University (TMU), Iran

Mohammad Mehdi Sepehri
Tarbiat Modares University (TMU), Iran

Pejman Shadpour
Iran University of Medical Sciences (IUMS), Iran

Seyed Hessameddin Zegordi
Tarbiat Modares University (TMU), Iran

ABSTRACT

Laparoscopy is a minimally-invasive surgery using a few small incisions on the patient's body to insert the tools and telescope and conduct the surgical operation. Laparoscopic video processing can be used to extract valuable knowledge and help the surgeons. We discuss the present and possible future role of processing laparoscopic videos. The various applications are categorized for image processing algorithms in laparoscopic surgeries including preprocessing video frames by laparoscopic image enhancement, telescope related applications (telescope position estimation, telescope motion estimation and compensation), surgical instrument related applications (surgical instrument detection and tracking), soft tissue related applications (soft tissue segmentation and deformation tracking) and high level applications such as safe actions in laparoscopic videos, summarization of laparoscopic videos, surgical task recognition and extracting knowledge using fusion techniques. Some different methods have been proposed previously for each of the mentioned applications using image processing.

DOI: 10.4018/978-1-5225-2515-8.ch014

INTRODUCTION

The general perspective of the chapter is considering different applications of processing laparoscopic videos to extract the valuable knowledge. The main scope is considering the applications of image processing for laparoscopic videos. For this purpose, only the previous studies proposing methods for processing RGB videos of laparoscopic/endoscopic surgeries are considered. 3D videos like RGB-D images and videos augmented by supplementary data from the external sensors or other types of images such as MRI are not in the scope of this study.

The rest of the paper is organized as follows. Section 2 reviews the applications of image mining in laparoscopy. In section 3, the datasets being used in the previous studies are listed and the discussion is presented in section 4. Finally, section 5 summarizes and concludes the paper.

BACKGROUND

Laparoscopy lies in the category of endoscopic interventions. During the endoscopic interventions, a camera called endoscope is inserted into the patient body to display the internal organs. Endoscopes are divided into two categories including flexible endoscope used for inspecting the esophagus, stomach, small bowel, colon and airways; and rigid endoscope used for a variety of minimal invasive surgeries (i.e., laparoscopy, arthroscopy, endoscopic neurosurgery). The endoscopes have various sizes with a tiny video camera at the tip (Oh et al., 2007). The endoscope in the laparoscopic surgeries is called telescope.

Laparoscopy is a minimally-invasive procedure with a few small incisions on the patient's body. Therefore, the hospital stay and recovery time for the patients after the laparoscopic surgeries is shorter than the open surgeries for similar surgical operations.

The first laparoscopic surgery was performed about 100 years ago on dogs. The first laparoscopic surgery on human was about 30 years ago. Till now, many advancements in laparoscopy have been occurred by introducing robotics and new instruments. It leads to less invasive surgeries (Cwach & Kavoussi,2016).

The surgical tools and the telescope are inserted through the incisions to conduct an operation. The telescope displays the internal organs and can record the surgery as a laparoscopic video (Uecker, Wang, Lee, & Wang, 1995).

Laparoscopy is widely used for diagnosis and treatment of many diseases (Yu et al., 2015). It has many advantages such as reduced patient trauma, reduced pain, decreased rates of infection and sepsis, small scars, reduced hospitalization with improved prognosis, lower rate of returning to the operating room, reduced need for blood transfusion and a quicker recovery (Amodeo, L., J.V., E., & H.R., 2009; Grasso, Finin, Zhu, Joshi, & Yesha, 2009; S. L. Lee et al., 2010; Semerjian, Zettervall, Amdur, Jarrett, & Vaziri, 2015).

But, this procedure has some drawbacks such as a limited view because of two-dimensional imaging, challenging eye-hand coordination, absence of tactile feedback (Schols, Bouvy, Dam, & Stassen, 2013) and surface view of the organs (Ma˚ rvik et al., 2004).

Many types of surgical operations can be performed by laparoscopic procedure. For example, laparoscopic partial nephrectomy (LPN) is a standard therapy method for renal carcinomas (Baumhauer et al., 2008). Laparoscopic cholecystectomy is a surgical procedure to remove the gall bladder (Sánchez-González et al., 2011). Some other surgeries that can be performed via laparoscopic procedure in reduced hospital stays, faster recovery, reduced pain and improved results compared with open surgery include colorectal surgery (Braga et al., 2010; Choi, Lee, Park, & Lee, 2010), cholecystectomy (Navarra, Pozza, Occhionorelli, Carcoforo, & Donini, 1997), appendectomy (D'Alessio, Piro, Tadini, & Beretta, 2002), adrenalectomy (Castellucci et al., 2008), hernia repair (Podolsky, Mouhlas, Wu, Poor, & Curcillo, 2010) and etc.

The laparoscopic surgery consists of three main phases including before inserting the telescope in the patient body, when the telescope lies inside the patient body and after the telescope has brought out from the patient body. The laparoscopic videos can be taken during the second phase. In the first phase, the patient is prepared for the surgery and some activities such as positioning, anesthesia and making the incisions are performed for the patient. The telescope is inserted into the patient body via one of the incisions. At this time, the second phase is started. In the second phase, all of the surgical activities being performed inside the patient's body can be displayed on the monitor via the telescope and recorded simultaneously. This video taken by the telescope in the second phase of the laparoscopic surgery can be processed for further analysis purposes. When the telescope is taken out from the patient's body for the last time and turns off, the second phase terminated and the third phase is started. The third phase continues till the patient is transferred to the recovery room. We want to review the previous studies proposing methods for processing the videos recorded in the second phase of the laparoscopic surgery.

A technical improvement in laparoscopic procedures is the robotic surgery. Robotic technology can provide stable camera with 3D optics and more comfortable position for surgeons. Different surgical robotic systems have been proposed previously such as AESOP, telerobotic Zeus and da Vinci systems. Telerobotic Zeus and da Vinci systems can be used for remote surgery as well. The robotic instruments simulate the surgeon's motion (Ballantyne, 2002).

Some main disadvantages of the laparoscopic surgery are limited precision and poor ergonomics for surgeons in using the instruments (Amodeo et al., 2009) which can be reduced by robotic surgery.

Since, a video of the laparoscopic surgery can be captured and recorded by the telescope, processing this video can be helpful for extracting useful knowledge from the laparoscopic video to answer the surgeons' questions. For example, the surgeons may want to know which surgical tasks are performed in a laparoscopic video and in which style every task is performed, which surgical instruments are inserted into the patient's body and the motion pattern of each instrument. The prerequisite for answering these questions is applying the appropriate image processing algorithms on the laparoscopic video to enhance the video, segment regionally and/or temporally the video, extracting the color, texture and motion based features from the segmented regions in the laparoscopic video. And then, using the extracted features, the high level knowledge can be extracted from the laparoscopic videos. For example, which motion pattern each instrument has and which surgical activity is being performed in each temporal segment of the video by the instruments.

Image processing can be used to improve the performance of laparoscopic robotic surgery. One of the characteristics of robotic laparoscopic surgery is using the robots to hold and manipulate the endoscope to help the surgeon during the lengthy surgery (X. Zhang & Payandeh, 2002).

Visual tracking method is needed to achieve automated instrument localization and endoscope maneuvering in robot-assisted laparoscopic surgery (X. Zhang & Payandeh, 2002).

Depth perception in minimally invasive surgery is reduced because of two-dimensional and limited field of view in laparoscopic videos. Augmented reality uses patient's medical images (US, CT or MRI) to increase the surgeon's vision via 3D visualization of anatomical or pathological structures and the registration of this visualization on the real patient (S. Nicolau, Soler, Mutter, & Marescaux, 2011).

APPLICATIONS OF IMAGE PROCESSING IN LAPAROSCOPY

As mentioned in the previous sections, laparoscopic technique in surgery allows to record surgical videos captured by the telescope.

Laparoscopic video shows the tissues and surgical tasks performed on the tissues by the surgical instruments if the telescope is inside the patient's body. Therefore, many applications can be solved by image processing on laparoscopic videos.

Sometimes, the telescope is outpatient for some reasons. Camera position classification (whether the telescope is inpatient or outpatient), camera motion estimation and compensation, surgical instrument detection and tracking, tissue segmentation and navigation are some elementary problems can be solved by processing the laparoscopic images.

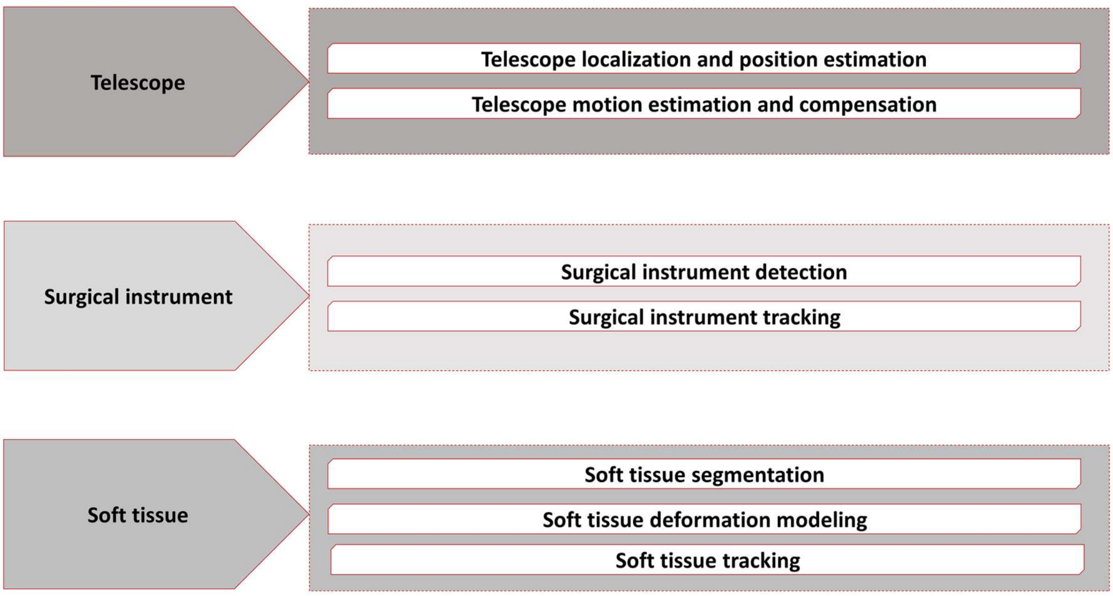
The mentioned applications can obtain the desired information to help solve some more complicated problems such as laparoscopic video summarization, safe action detection, and laparoscopic activity recognition.

There are some complicated problems that cannot be solved only by analyzing the laparoscopic videos. They need more information that must be obtained from other imaging techniques such as MRI, CT, and etc.

3D models can be obtained by augmenting the laparoscopic video with other information such as other imaging techniques or data obtained by external sensors. These models can be used in virtual reality for enhancing the surgical skills of the novice surgeons and mining these datasets is not in the scope of our study.

As illustrated by Figure 1, factors having visible effects on the laparoscopic videos are telescope, surgical instruments and soft tissues. Their related image processing problems are shown in Figure 1 as well.

Figure 1. Factors having visible effects on the laparoscopic videos and their corresponding image processing problems



In this study, we will focus on applications that can be solved by processing the laparoscopic videos (as shown in Figure 2).

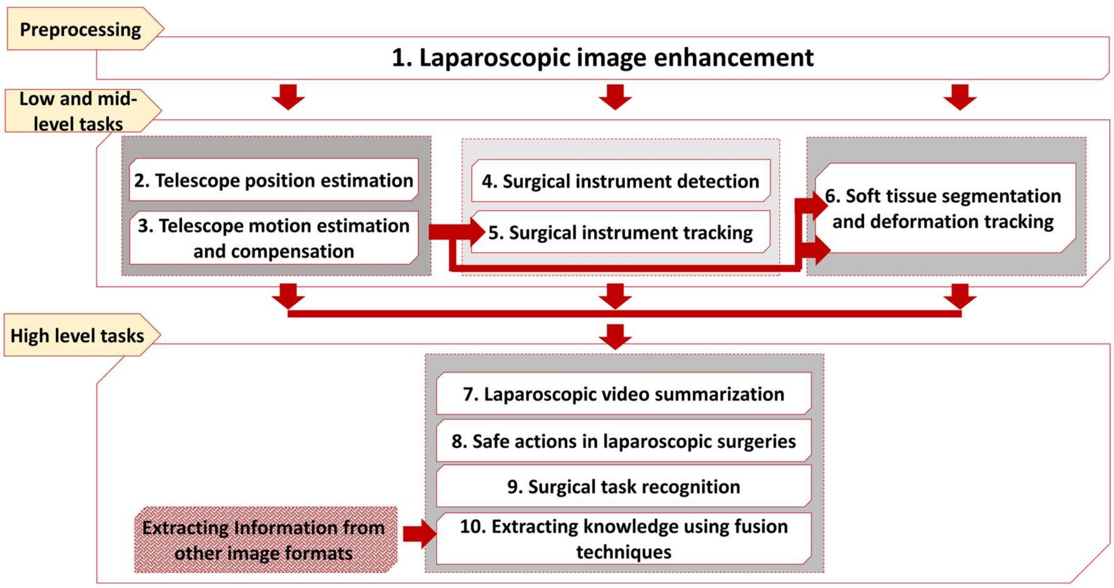
Figure 2 shows different image processing tasks that can be performed on the laparoscopic video. These tasks are classified into pre-processing task, low and mid-level tasks and high-level tasks. Pre-processing task includes image processing algorithms being exploited for enhancing the laparoscopic video frames by reducing noise, improving the contrast and etc.

Usually, before performing every image processing task on the laparoscopic video, it is required to pre-process the laparoscopic video frames. Therefore, the output of the laparoscopic image enhancement is the input of the tasks in the second category in Figure 2.

The second category illustrated in Figure 2 includes three different task types that try to extract features from the telescope, surgical instruments and soft tissues. The first task tries to estimate the position and motion of the telescope which is the camera of the video and tries to compensate the global motion pattern caused by the camera. The second task includes surgical instrument detection and tracking. The prerequisite task for surgical instrument tracking is the telescope motion estimation and compensation. It must be performed to improve the accuracy of estimation of the motion patterns of the surgical instruments. The last task in the second category is related to soft tissue segmentation and deformation tracking. The soft tissue deformation tracking has two prerequisites including the telescope motion estimation and compensation and surgical instrument tracking, because the global motion and the instrument motion can have influence on the neighboring soft tissues.

The tasks of the second category provide the necessary input for the third category tasks. The third category tasks are much more complicated tasks. They try to extract useful knowledge from the laparoscopic videos. These tasks include laparoscopic video summarization, detecting safe/unsafe actions,

Figure 2. The classification of laparoscopic video processing problems considered in this chapter



surgical task recognition and extracting knowledge using fusion techniques, in which the last item needs more inputs which can be collected from other external resources such as CT-scan or magnetic resonance images.

Laparoscopic Image Enhancement

Laparoscopic image enhancement is a preprocessing step that can improve the accuracy of extracting information from the video frames. Previous studies have proposed many different image enhancement methods (Giannarou, Visentini-Scarzanella, & Yang, 2012; Selka et al., 2015).

Enhancing images with the aim of improving the image quality is being performed and includes improving the image contrast, reducing unusual radiation, improving the image brightness, removing noise and etc.

Image enhancement must be performed using fast and simple algorithms for laparoscopic video frames because of their large number. Therefore, many of the previously proposed algorithms for enhancing images cannot be applicable on laparoscopic video frames because of their high computational complexity.

Generally image enhancement methods can be classified into simple pixel-level methods, filter-based methods and pseudo-color based methods. Simple pixel-level methods include intensity transformation and histogram equalization methods. Filter-based methods have used spatial or frequency domain filters to smooth or sharpen the images. Pseudo-color based methods have used false colors or pseudo color spaces.

Unfortunately, because of the limited number of published researches about laparoscopic image processing, the limited methods have been used for laparoscopic image enhancement.

Histogram equalization (HE) has been used to limit tracking failure (Wu, Chen, Liu, Chang, & Sun, 2004), adjust the image contrast and reduce the negative effects of the illumination changes on tracking the tissues (Giannarou et al., 2012). Selka et al. have shown that HE does not improve the tissue tracking performance (Selka et al., 2015).

In some researches, Gaussian smoothing filter has been performed to remove variations and noise, the brightness and contrast has been enhanced using classic brightness/contrast adjustment method and histogram thresholding has been used to remove specular reflections from laparoscopic video frames (Shu & Cheriet, 2005).

Both of the mentioned methods for enhancing the laparoscopic video frames are simple and fast. Therefore, they are applicable on the real laparoscopic video frames. But, they are sensitive to the parameter tuning.

Other than contrast enhancement, another challenging issue affecting laparoscopic image processing steps is specular reflection on the soft tissues. Therefore, a method for segmenting and recognizing specular reflection was proposed in the previous researches. For this purpose, dynamic thresholding technique is used to segment the video frames. Then the closed contours of the segmented regions are analyzed, and three different types of reflections are identified including non-reflection, small reflection with high brightness and reflections with small brightness. The contour shape is used to classify the regions into these three different classes (Marcinczak & Grigat, 2013).

This method is sensitive to parameter tuning as well. Moreover, for setting the dynamic thresholds, it needs the training data.

Telescope Position Estimation

Sometimes, it is important to know whether the telescope shows the inside of the patient's body or outpatient. If the telescope is outpatient, there is no need to process and archive the video frames. Surgical interruptions and problems occurred during the surgery may cause the telescope to exit the patient's body.

For example, when the telescope lens becomes foggy or dirty during the surgery, the surgeon's visibility is reduced and it is required to take out the telescope from the patient's body and clean it.

Sometimes, other problems occurring with the outpatient that may cause long interruptions leading to take out the telescope from the patient's body. For example, CO2 capsule becomes empty and continuing the surgery before providing a new capsule is impossible. In these situations, the telescope is taken out from the patient's body.

The telescope position estimation can be seen as a classification problem in which the video frames are classified into outpatient and inpatient.

Stanek et al. have introduced novel color-based features to detect outpatient frames in endoscopic videos. The proposed features include Mean-red, mean-normalized-red, and accumulated mean-normalized-red. For classifying video frames, a threshold-based method has been used in (Stanek, Tavanapong, Wonga, Oh, & Groenc, 2012), and the results showed that the video frames could be classified based on the proposed features with high accuracy. But, the inpatient images of the laparoscopic surgery may have different color histogram pattern from the endoscopic video frames depending on the target tissue. Therefore, the proposed feature may not be sufficient to classify the laparoscopic video frames to inpatient and outpatient images.

Telescope Motion Estimation and Compensation

Camera motion estimation and compensation method is very important to process the laparoscopic videos. For example, for estimating local motions such as surgical instrument motion or tracking the instrument accurately, global motion can affect the local motion parameters and reduce the accuracy of local motion estimation and tracking. Therefore, camera motion estimation and compensation is a necessary pre-processing step in many applications of laparoscopic video processing.

Many previous studies have considered camera motion estimation and compensation. But, as the best of our knowledge, we have not seen any study considering camera motion estimation and compensation for laparoscopic videos specifically.

Camera motion estimation and compensation method has been studied in many researches generally (B. H. Chen et al., 2016). But, camera motion estimation and compensation in laparoscopic videos is different and more challenging task than camera motion estimation and compensation in indoor or outdoor scenes. Tissue respiratory motion can make camera motion estimation in laparoscopic videos a more challenging task. Moreover, very homogeneous tissue surfaces in very large magnified video frames, local motions of surgical instruments, their interaction with tissues and tissue deformation due to surgical instrument forces are other factors that can increase the complexity of camera motion estimation in laparoscopic videos.

Surgical Instrument Detection

If 3D pose of the surgical tools can be estimated, it will be valuable to extract data for the surgical training evaluation and using the knowledge of surgical instrument position to design new robotic systems. Since the laparoscopic video is recorded in two dimensions, surgical instrument shape and orientation can be useful to get its 3D position (Cano González, Vara Beceiro, Sánchez González, Pozo Guerrero, & Gómez Aguilera, 2008).

Surgical tools are solid objects. In the literature, detection and tracking of solid objects are easier than the soft tissues, because they have specified border and their shapes are less variant than the soft tissues. But, if the orientation of the solid object and its angle to the camera change, its shape may change as well. Therefore, the shape of surgical instrument in laparoscopic video frames depends on its orientation and relative position to the telescope. It makes instrument detection and tracking a challenging image processing task.

Another issue is the perceived color and texture of the surgical tool. Surgical tool tip is usually metallic and silver. Therefore, it is more sensitive to brightness and illumination variations. It may lead to inhomogeneous texture with a variety of colors and different brightness levels on the surface of the tool tip. Therefore, surgical tool segmentation and detection will be challenging.

The surgeon may need to adjust the camera magnification in different video frames. The camera magnification parameter can affect the perceived texture and color of the surgical tool tip.

According to the mentioned challenging issues, many researches have proposed methods for surgical instrument detection and tracking (Voros, Long, & Clinquin, 2007). In this section, some previous studies considering surgical instrument detection are reviewed. In the next section, some previous researches about tracking the surgical instruments are considered.

Segmenting the laparoscopic video frames into homogeneous segments is a necessary prerequisite step for surgical instrument detection. After segmenting the video frame, each segment is analyzed and the segments indicating the surgical instruments are detected. Therefore, detecting the surgical instrument consists of two main steps including image segmentation and then classifying the generated segments into surgical instrument and other regions.

Previously proposed methods for segmenting the video frames to detect the surgical instrument are classified into color-based segmentation methods, segmenting based on the combination of color and texture features, and edge-based segmentation methods.

Color-Based Segmentation Methods

Amini Khoyi et al. have proposed a segmentation algorithm based on combination of saturation and Value components of HSV color space to detect and track the tool tip (Amini Khoyi, Mirbagheri, & Farahmand, 2016).

Some researchers have proposed a fast color-based segmentation method to detect the gray regions in color images of laparoscopic videos. They have assumed that surgical tool's color is seen as gray in images which is not true all of the time. This method is based on recursive thresholds obtained from color histograms in HSI color space. Saturation component is enhanced and used in combination with hue component. Then a region growing approach is used for instrument segmentation. Some shape descriptors are extracted from the gray segmented regions including Fourier descriptors and moments. Then the gray regions are classified into instrument and non-instrument classes based on the extracted features (Doignon, Graebling, & Mathelin, 2005). The main advantage of this approach is its high speed. But, it is based on color features and region shape descriptors. Therefore, it is not robust to illumination variations and partially occlusion.

Some researchers have used artificial color markers on the surgical instruments to detect these markers in the images (Wei, Arbter, & Hirzinger, 1997; Zhao, 2014). It is a very fast approach and is more robust to the illumination variation and surgical tool partially occlusion. But, it may not be applicable in all the surgeries to use artificial markers on the instruments.

Segmenting Video Frames Based on the Combination of Color and Texture Features

Another approach has been proposed for surgical instrument detection in laparoscopic video frames based on color and texture features. At first, the images are segmented twice, one time based on color features and another time based on texture features. Then the segmented regions are merged if they have enough overlap. Finally, the regions having texture and color pattern similar to the surgical instruments are considered as surgical tool regions (Khatibi, Sepehri, & Shadpour, 2013). Using color and texture features simultaneously can reduce the error rate of surgical tool detection due to illumination variation and occlusion.

Edge-Based Segmentation Methods

Some researchers have tried to detect the surgical tools based on edge detection and applying the Hough transform on the detected edges in laparoscopic video frames. They have assumed that the surgical tool border is a combination of straight lines. Finally, motion information has been used to discriminate surgical tools from other structured elements in images (Climent & Mars, 2010). The Hough transform is robust under illumination variations, occlusion and distractions (Climent & Hexsel, 2012).

Finally, when the video frame is segmented, the segmented regions with similar color or motion pattern to the surgical instruments are considered as the surgical instrument regions.

Surgical Instrument Tracking

Instrument tracking can be used for autonomous control of a cameraman robot during the robotic laparoscopic surgery (Amini Khoyi et al., 2016). Prerequisite step for instrument tracking is the image segmentation for detecting the surgical instruments which is described in the previous section.

But in laparoscopic videos, a stable motion estimation is very complicated and challenging task and cannot be guaranteed generally (Öhsen, Marcinczak, Vélez, & Grigat, 2012).

Many different methods were proposed in the previous researches for tracking surgical instruments in the laparoscopic videos via image processing algorithms (Wolf, Duchateau, Cinquin, & Voros, 2011). The previously proposed methods for surgical instrument tracking can be divided into two main categories including tracking-based on image processing and tracking-based on the supplementary data.

Tracking the Surgical Instrument Only Based on Image Processing Methods

Some researchers have proposed an image processing method for tracking surgical instruments without artificial markers. The proposed method could estimate 2D/3D pose of visible instruments in real-time using Frangi filter for detecting edges and tool tips (Agustinos & Voros, 2016).

An adaptive mean-shift Kalman filter has been used to track the tool tip in another study. The filter can increase the accuracy of tracking when the tool tip is invisible in a few frames. This method is robust to the changes of size and shape of the instrument (Sa-Ing, Thongvigitmanee, Wilasrusmee, & Suthakorn, 2012).

In another study, a novel approach has been proposed for surgical instrument detection and tracking. The proposed image segmentation algorithm is based on training a decision tree on color features to classify the regions as surgical instrument or other regions. Then, the corner points of the surgical instruments are extracted and matched in two consecutive video frames. The instrument motion vector is estimated based on the replacement of the matching corner points in two consecutive frames (Khatibi, Sepehri, & Shadpour, 2014). This method is based on color and motion features and all the limitations of the color-based instrument segmentation and local motion estimation holds true, here.

Some researchers have used methods for segmenting and tracking the tools and organs to 3D reconstruct the surgical field (Sánchez-González et al., 2011).

Needle tracking is an assisting prerequisite of suturing which is a complex and difficult task in laparoscopy. Some researchers have used color and geometry-based segmentation to detect the needle (Speidel et al., 2015).

Other researchers have used texture and geometric features from edge-based descriptors as input to a spiking neural network for detecting the instrument. Tracking the instrument has been done using Kalman filter (C. J. Chen, Huang, & Song, 2013).

The mentioned methods have used color, texture, edge-based and/or motion features for surgical instrument tracking. Unfortunately, none of the color, texture and edge features can be extracted accurately in many real laparoscopic video frames, because of illumination variation, magnification changes and other challenging situations described before. Tracking objects based on motion vectors can be performed using filters such as Kalman filter with reasonable accuracy. Therefore, considering the combination of color, texture, edge and motion descriptors can improve the accuracy of instrument tracking. But, it may lead to more computational complexity of the tracking method.

Tracking the Surgical Instrument Based on the Supplementary Data

This category consists of methods using image processing techniques along with other supplementary data obtained by external sensors or other resources for tracking the surgical instruments.

Oropesa et al. have proposed an instrument tracking software based on tool edges, tool insertion point and focal point of camera to extract 3D position of the instruments (Oropesa et al., 2013). This method needs supplementary information and the tool tracking cannot be performed using only laparoscopic video frames. In robotic surgery, some researchers have proposed an approach for tracking surgical instruments to measure the distance between the instrument surface and a laser point illuminated on the tissue. For detecting the laser points, they have used high-pass filters, and then they have segmented the filtered image and used erosion morphologic operator (Krupa, Doignon, Gangloff, & Mathelin, 2002). This method needs laser point illumination which may not be available and safe in all situations.

Adding the supplementary data to the laparoscopic video may improve the accuracy of the surgical instrument tracking. But, it is not possible to use external devices and sensors to provide such supplementary data in many situations. Moreover, the synchronization of supplementary data with video data is another challenging issue that must be addressed.

Tissue Segmentation and Deformation Tracking

Sometimes, tissue deforms due to instrument forces and many other factors. Therefore, tissue deformation tracking is an important issue in laparoscopic video processing.

A preprocessing step for tissue tracking can be tissue segmentation. Tissue segmentation in laparoscopic videos can have many other advantages and is very helpful. For example, in gynecology, tissue segmentation can be performed to localize the uterus to facilitate and analyze the lesion on uterus and its surroundings. Therefore, a threshold-based segmentation has been proposed to segment the tissues, and then the regions are classified by SVM to uterus and other tissues (Yu et al., 2015). The threshold-based method for segmenting the soft tissues may divide a soft tissue into multiple regions because of illumination and contrast variations in different blocks of the image. Therefore, it is not an exact and accurate segmentation method.

Tissue navigation system which uses intraoperative imaging and real-time image processing can enhance the surgeon's perception and can guide surgeon's decisions and to provide some valuable information such as information about the tumor infiltrated tissue and risk structures (Baumhauer et al., 2008).

In the previous studies, some methods have been proposed for tissue deformation tracking (S. L. Lee et al., 2010; Stoyanov & Yang, 2011).

Collins et al. have introduced a method for 3D reconstruction of the tissue surface using the sliding windows approach (Collins, Compte, & Bartoli, 2011). Using fusion techniques and other imaging data such as MRI and/or CT images can improve the proposed method accuracy.

Stoyanov and Yang have proposed optical feature tracking method based on stereo-laparoscopic images. They have defined and used a constrained geometrical surface model that deforms with motion. Their proposed method is robust to occlusions and specular reflection, but it needs stereo camera (Stoyanov & Yang, 2011).

A previous study has proposed some context descriptors for tissue deformation tracking. They have considered tissue tracking as a classification problem and solved it using decision trees. They have argued that the proposed method is robust to drift, occlusion, orientation and scalingvariations (Mountney & Yang, 2012). Since the classifier trains in a supervised manner, the generalization ability for tissue deformation tracking may be reduced and for new deformations that are not seen in the training set, it may fail to recognize them.

Some researchers have proposed a technique to generate rough depth maps from laparoscopic video frames based on SURF features. This method can provide 3D visualization of 2D laparoscopic video (P. Y. Lee, Yan, Hu, & Marescaux, 2015). SURF features can be extracted from the interest points and matching the SURF features of the consecutive video frames can be used for estimating the motion of these points. SURF features are one of the highly accurate features that can be used for tracking the objects. But, it needs the number of the SURF features extracted from the tissue to be sufficient to estimate the tissue deformation with high accuracy.

Laparoscopic Video Summarization

One popular camera system for laparoscopic surgeries is the vision system of the daVinci-Si robotic surgical system (Intuitive Surgical, Sunnyvale, California) with high quality. Its generated video streams has approximately 360 MB of data per second (Ronaghi, Duffy, & Kwartowitz, 2015). This large volume of data shows the importance and necessity of laparoscopic video summarization.

On the other hand, all frames of a laparoscopic video do not contain valuable information. When the surgeons want to archive the laparoscopic videos for further analysis and training purposes, they are not interested in archiving all segments of a lengthy laparoscopic video. Therefore, summarizing a laparoscopic video by keeping only the valuable segments and removing the other segments can have high importance. It can reduce the review time of a summarized video and the required memory for archiving the video. But, the question is which segments of a laparoscopic video are valuable and which segments can be removed without information loss.

Simply, when the telescope is outside the patient's body or when there is no camera motion and/or surgical instrument movement, it can be concluded that no surgical activity is being performed inside the patient's body. Therefore, these video segments can be removed without loss of information.

The previously proposed methods for laparoscopic video summarization can be classified into two categories including key-frame selection and informative segment selection.

Key-Frame Selection

Key-frame selection tries to find the most desired video frames denoting the tissues and surgical activities with high quality. Therefore a video with N video frames may be summarized into K video frames ($K << N$) which every key-frame is selected from one temporal segment of the laparoscopic video. Therefore, the selected video frames are not immediately consecutive.

Some researchers have proposed algorithms for key-frame selection in laparoscopic/endoscopic videos (Öhsen et al., 2012; Wang et al., 2016). Ohsen et al. have scored all combination of frames in a sequence and they have tried to find the optimal solution. Finding the optimal solution have been described as a weighted directed graph problem and Dijkstras Algorithm has been used to find the best selection of frames (Öhsen et al., 2012).

Before extracting key-frames, video must be segmented temporally. Wang et al. have proposed a video segmentation method with an assumption that the semantic content of frames in a segment should change dramatically. Color and texture features are used for this purpose. Color features are the I component histogram in HSI color space, histogram of H and V components in HSV color space, RGB histogram, Normalized RGB histogram, RG histogram, Opponent histogram and HUE histogram. Local binary pattern (LBP) histogram is used for describing the texture. The reason of using histograms as the image features is their robustness against rotation and scaling variations. The similarity between two consecutive frames are calculated as their histogram intersection (Wang et al., 2016).

Another study has proposed a data reduction algorithm for extracting representative video frames (key-frames) from a Wireless Capsule Endoscopy (WCE) video. It can reduce significantly the reading time of WCE videos for experts without any loss of abnormalities. This method is used WCE video segments with fixed length, and features are extracted from video frames of each segment. Then the number of features are reduced using data reduction algorithms. The reduced feature set are clustered and the representative frames are selected from the clusters showing abnormalities in tissues (Iakovidis, Tsevas, & Polydorou, 2010).

In some researches, they have proposed methods for laparoscopic video summarization (Grasso et al., 2009). Grasso et al. have used support vector machines (SVM) for classifying the video frames to identify which frames are near to critical view and which frames are not. They have extracted feature vector of spectral and textural features such as energy, entropy, contrast, correlation and color descriptors. Then the feature selection have been performed using the Jeffrey Divergence with an experimentally derived threshold (Grasso et al., 2009).

In another study, similar video frames taken by a wireless capsule endoscope have been eliminated to minimize the redundancy in the images. For this purpose, enhanced intensity features, optical flow and Speed Up Robust Features (SURF) have been used to detect and reduce near-duplicate images (H. G. Lee, Choi, Shin, & Lee, 2013).

Some researchers have classified endoscopic video frames into informative and non-informative frames (Bashar, Kitasaka, Suenaga, Mekada, & Mori, 2010; Oh et al., 2007). For this purpose, Oh et al. have proposed two different approaches. In the first method, edge-based features are extracted by Canny edge detector and then the feature set are classified to informative, non-informative and ambiguous frames. In the second method, each frame is transformed using Discrete Fourier Transform (DFT) and the texture of the transformed image is analyzed. For this purpose, gray level co-occurrence matrix (GLCM) is extracted from the frequency spectrum image. Then seven texture features including Entropy, Contrast, Correlation, Homogeneity, Dissimilarity, Angular Second Moment and Energy are extracted from GLCM. The feature set is clustered using k-means algorithm into two clusters and each cluster is labeled as informative or non-informative cluster (Oh et al., 2007).

Informative Segment Selection

All of the above methods lay in the first category that tries to find key-frames. Now, the methods belonging to the second category tries to find the informative temporal segments which they will be reviewed.

The second class including informative segment selection tries to segment temporally the laparoscopic video and then classifies the temporal segments into informative and non-informative segments. Finally, only the informative segments are kept.

Temporal segmentation of laparoscopic video is an important prerequisite step for laparoscopic video summarization and identifying which temporal segments can be removed without the loss of information.

Khatibi et al. have proposed a novel approach for temporal segmentation of laparoscopic videos. They have extracted instrument motion parameters and defined temporal segmentation of video as an optimization problem. Then, they have solved this problem using multi-objective genetic algorithm in an unsupervised manner (Khatibi et al., 2014). The generated temporal segments can discriminate different surgical activities and each activity lies in a separate temporal segment. Because the surgical instruments have different motion patterns in different surgical activities. The temporal segments can be used for discriminating different surgical activities in surgical task recognition and finding the most desired activities. The laparoscopic video can be summarized by keeping only the segments corresponding to the most desired surgical activities. It is a more complicated type of the video summarization applications.

Other researches have proposed simpler methods for laparoscopic video summarization (Munzer et al.,2013; Stanek et al.,2012). They want to find inpatient segments and/or high quality segments. They do not distinguish different types of surgical activities.

Munzer et al. have defined three types of unrelated segments of laparoscopic videos including dark segments, segments which telescope is outside the patient's body and blurry segments. They have tried to discriminate these three different classes using features extracted from laparoscopic video. For this purpose, some preprocessing steps have been performed on the video frames including specular reflection removal, ignoring border pixels, low saturated pixels and dark segments. Then, HSV color features have been used. Their reason for choosing HSV color space has been its superiority to RGB color space. HSV color space is near to human perception (Münzer, Schoeffmann, & Böszörmenyi, 2013).

Stanek et al. have proposed an approach to automatically detect informative temporal segments of an endoscopic video and recorded only informative temporal segments. For this purpose, motion and color features have been used. Outpatient video frames have been detected and automatically discarded (Stanek et al., 2012). Their proposed method classifies the temporal segments into inpatient and outpatient segments and records only the inpatient segments using buffers.

In future, more deep researches can be performed for laparoscopic video summarization. For example, surgeons may be interested in keeping only special types of surgical tasks and remove other segments. It needs to recognize the laparoscopic surgical tasks from videos which will be discussed in section 3.4.

Safe Actions in Laparoscopic Surgeries

The surgeon wants to perform a safe surgical activity. For this purpose, the surgeon needs to know whether performing the surgical task by the instruments may hurt the tissues and/or arteries. For example, in cholecyctectomy, cystic artery must be avoided from the injury by the surgical instruments. Otherwise, injury to the cystic artery may lead to uncontrollable bleeding (Lahane et al., 2012).

Therefore, it is required to detect the blood vessels in laparoscopic video frames using image processing techniques. Lack of 3D view in laparoscopic videos makes it difficult to detect blood vessel positions accurately. Some researchers have used color and geometric features to detect blood vessels in laparoscopic video frames. But, this approach cannot recognize vessels occluded by fat or other tissues (Akbari & Kosugi, 2007).

Other researchers have used HSV color space information for detecting blood vessels among the enhanced laparoscopic video frames using multiple regression analysis (Hiroyasu, Tanaka, Hagiwara, & Ozamoto, 2015).

During the surgery, surgical instrument should not perform surgical activities near some critical tissues and vessels. For this purpose, surgical tool tip is detected in video frames using image processing techniques. Then, the fixed-size rectangular block around the surgical tool tip in laparoscopic video frames are drawn and analyzed to verify whether critical tissues or arteries exist there or not. If there is no critical tissue or artery in the block, it has been concluded that the block is safe. The regions around the tool tip are classified as safe or unsafe regions for surgical activities (Lahane et al., 2012).

Another study has proposed chromaticity moments of HSI color space as the features for bleeding and ulcer detection in WCE images and discriminating normal regions and abnormal regions. This classification problem has been solved using a neural network classifier (Li & Meng, 2009).

The mentioned above methods try to find blood vessels and/or discriminate normal/abnormal regions. In the laparoscopic surgery, these are two important applications of image processing for preventing unsafe actions. But, many other situations may exist that can cause the unsafe actions. In the future, the researchers can focus on finding and preventing other situations that may lead to unsafe surgical actions by image processing techniques.

Surgical Task Recognition in Laparoscopic Videos

An important and complex application of laparoscopic video processing is surgical task recognition for activities which are visible in the laparoscopic videos and performed in the patient's body.

Generally, several studies have been performed on human activity recognition in video clips (Moayedi, Azimifar, & Boostani, 2015; Yuan, Zheng, & Lu, 2016). Activity recognition problem is a complicated problem and there are many challenges yet. For example, same action can be performed in different styles and different speeds by different persons. Therefore, in literature, many different descriptors have been proposed for human activity recognition such as motion and shape descriptors (Dou & Li, 2014; Kviatkovsky, Rivlin, & Shimshoni, 2014).

There are some differences between human activity recognition and surgical activity recognition. Actors in the laparoscopic surgeries are surgical instrument. The surgical instrument can be hidden in the neighboring tissues for a few minutes during the surgery. They can cause soft tissue deformation. Therefore, the interaction between the actors and their neighboring tissues makes the surgical activity recognition problem different, more complicated and more challenging than the human activity recognition.

Recent approaches have demonstrated that video data can be used in surgical task recognition with high accuracy (Blum, Feuner, & Navab, 2010; Padoy et al., 2007).

The previously proposed methods for surgical activity recognition can be classified into two classes: the first class includes methods using only the information obtained from the laparoscopic videos and the second class includes the methods using other supplementary resources of data such as data obtained by external sensors being installed on the surgical instruments and capturing the instrument motion pattern or sensors detecting surgeon's eye motion.

Methods Using Only the Information Obtained From the Laparoscopic Videos

Some researchers have proposed an approach for recognition of surgical tasks in eye surgery videos. Short video subsequences have been classified into surgical tasks by fast nearest neighbor classifier. For this purpose, fixed-length feature vector including texture, color and motion features are extracted from each subsequence (Quellec et al., 2014). But, there are some differences between eye surgery videos and laparoscopic videos. Laparoscopic videos indicate different tissues and more variety of surgical activities. Moreover, the eye and its border are clear in eye surgery videos. Therefore, the video processing of laparoscopic surgery has more and different challenges.

Detection and segmentation of surgical motion can be used for evaluating surgical skill, obtaining surgical training feedback and archiving essential aspects of a procedure. Some researchers have used raw motion data to recognize surgical actions (H. C. Lin, Shafran, Yuh, & Hager, 2006). The motion data can be obtained by installing the appropriate sensors on the surgical instrument or by tracking the instruments in the video by image processing algorithms. Therefore, we lay this method in the first class.

Some researchers have proposed using Hidden Markov Models to recognize surgical tasks based on surgeon's motion analysis (Dosis et al., 2005). Some researchers have shown that surgical movement data can be used for laparoscopic surgical skill assessment with high accuracy (Z. Lin, Uemura, Zecca, & Sessa, 2012). But, extracting stable motion features is a challenging and complicated task.

Researchers have proposed generating surgical activity log from laparoscopic video as an information source. For this purpose, surgical instruments have marked and the markers have been tracked to produce a surgical activity log. Then, the workflow model has been extracted and the outliers have been detected using global pair-wise alignment (Bouarfa & Dankelman, 2012).

The Methods Using Other Supplementary Resources of Data

Many studies have used dynamic cues such as time to completion, speed, forces, torque or kinematic data such as robot trajectories and velocities to automatically recognize the surgical skills and gestures (Zappella, Béjar, Hager, & Vidal, 2013). They have assumed that the video has been segmented into video clips and each video clip has denoted a single gesture. Surgical gestures include grabbing the needle, passing the needle and etc. Three different models have been proposed for surgical gesture recognition based on video data and kinematic data. The first model has considered video clip as the output of a linear dynamical system (LDS) and has tried to classify the video clips. In the second one, spatio-temporal features are extracted from each video clip with bag-of-features representation for classification. The third one has used multiple kernel learning (MKL) to combine LDS and BOF approaches (Zappella et al., 2013).

After stating the methods of the two different categories of surgical action recognition, we must say which challenges and opportunities exist here. In laparoscopic videos, the actor is surgical instrument. Therefore, shape and motion descriptors can be extracted from the surgical instruments to recognize different types of surgical activities. Another information that can be used for surgical activity recognition is the interaction between the surgical tools and the soft tissues. Therefore, some extra descriptors must be extracted for describing different types of interaction between the instrument and the tissues.

Some researchers have shown that there is the statistical difference between expert surgeons and residents in terms of: (1) force/torque magnitude, (2) type of tool/tissue interactions and (3) time interval for each tool/tissue interaction and total completion time (Rosen, Solazzo, Hannaford, & Sinanan, 2002). Therefore, surgical activity recognition and recording the time intervals for each surgical task can improve the accuracy of surgical skill assessment methods. But, extracting the shape and motion of the surgical instruments and describing the interactions between the instruments and the soft tissue are very challenging and complex issue.

Extracting Knowledge From Laparoscopic Videos Using Fusion Techniques

Fusion techniques are not in the scope of this study. Therefore, this section is briefly described.

For robotic assisted minimally invasive surgery, knowing the 3D shape of tissues is important for guiding the robots and compensating the tissue respiratory motions (Stoyanov, Elson, & Yang, 2009).

Lack of 3D view in laparoscopic video frames makes it challenging to extract knowledge from laparoscopic videos in many applications. Therefore, it is required to use fusion techniques of surface information of laparoscope and supplementary information obtained from other image formats such as MRI, CT Scan and ultrasound images. For example, depth of blood arteries inside the target organs can be estimated by projecting the ultrasound image over a laparoscopic image (Zenbutsu, Igarashi, & Yamaguchi, 2013).

Many researches have used other images such as ultrasound images(Zenbutsu et al., 2013), MRI, CT scan images(Hayashi, Misawa, Hawkes, & Mori, 2016), multispectral imaging(Y. Zhang et al., 2016) and etc for extracting information from laparoscopic video frames more accurately.

Some papers have used stereoscopic video camera attached to a laparoscope to implement image-based 3D reconstruction of soft tissues (Kowalczuk et al., 2012).

In this chapter, we are interested in extracting knowledge using only laparoscopic video without fusion techniques.

Finally, the summary of many previous studies considered in this chapter is shown in Table 1.

Table 1. Summary of the previous studies considered in this chapter

Authors	Topic	Image Preprocessing	Features	Model
(Shu & Cheriet, 2005)	Laparoscopic image segmentation	Gaussian smoothing Enhancing brightness and contrast Histogram thresholding	intensity	Graph-based and region merging
(Marcinczak & Grigat, 2013)	Specular reflection segmentation	-	Intensity Region boundary shape descriptors	Dynamic thresholding technique Classifying the regions
(Lahane et al., 2012)	Unsafe action detection	Gaussian smoothing	HSV color features	Jeffrey divergence feature selection SVM classifier
(Akbari & Kosugi, 2007)	Blood vessel detection	-	RGB color features	LVQ neural network
(Hiroyasu et al., 2015)	Blood vessel detection	-	HSV color features	Multiple regression analysis
(Li & Meng, 2009)	Bleeding and ulcer detection	-	HSI color features	Neural network classifier
(Doignon et al., 2005)	Surgical instrument detection	-	HSI color features Region boundary shape features: Fourier descriptors and moments	Region growing algorithm
(Khatibi et al., 2013)	Surgical instrument detection	-	Color features Texture features	Novel region merging algorithm
(Climent & Mars, 2010)	Surgical instrument detection	-	Edge-based features Hough transform Motion features	-
(Wei et al., 1997)	Surgical instrument detection	-	Color features	-
(Zhao, 2014)	Surgical instrument detection	-	Color features	-
(Krupa et al., 2002)	Surgical instrument tracking	High-pass filtering Erosion morphologic operator	Color features Region boundary descriptors	-
(Speidel et al., 2015)	Surgical instrument detection	-	Color features Geometry features	-
(C. J. Chen et al., 2013)	Surgical instrument detection and tracking	-	Texture features Geometric features	Spiking neural network Kalman filter
(Amini Khoyi et al., 2016)	Surgical instrument detection	-	HSV Color features	-
(Sa-Ing et al., 2012)	Surgical instrument tracking	-	Color features	Mean-shift Kalman filter
(Oropesa et al., 2013)	Surgical instrument tracking	-	Edge features Tool insertion point	-
(Agustinos & Voros, 2016)	Surgical instrument tracking	-	Edge features Frangi filter	-
(Öhsen et al., 2012)	Video summarization: Key frame selection	-	Motion vectors	Dijikstra algorithm

continued on following page

Table 1. Continued

Authors	Topic	Image Preprocessing	Features	Model
(Wang et al., 2016)	Video summarization: Key frame selection	-	Color histograms: HIS-I histogram, HSV-HV histogram, RGB histogram, NormRGB histogram, RG histogram, Opponent histogram and HUE histogram. Texture features: LBP histogram	Similar-inhibition dictionary selection
(Iakovidis et al., 2010)	Video summarization: Extracting representative video frames	Subsampling video frames Dimensionality reduction of video segments	Non-negative matrix from intensity	Fuzzy clustering
(Grasso et al., 2009)	Video summarization	-	Color features Textural features: energy, entropy, contrast and correlation	Feature selection using Jeffrey Divergence SVM
(Münzer et al., 2013)	Irrelevant scene detection	Specular reflection detection Excluding border pixels dynamically Ignoring low saturated pixels	HSV color features	Fuzzy classification
(H. G. Lee et al., 2013)	Video summarization: eliminating similar frames	HE	Motion features: optical flow SURF Intensity	Egomotion classification
(Bashar et al., 2010)	Video summarization	-	Color features: local moments and HSV histogram Gauss Laguerre transformation (GLT) based multiresolution feature	SVM
(Oh et al., 2007)	Video summarization: informative frame detection	Identifying specular reflection using dynamic thresholding and outlier detection Discrete Fourier Transform	Edge-based features: using Canny Edge Detector Texture features: entropy, contrast, correlation, homogeneity, Angular second moment, energy	Clustering-based frame classification
(Stanek et al., 2012)	Video summarization: irrelevant scene detection	-	Color features Motion features	Novel proposed algorithms for identifying the start frames and end frames of a procedure
(Khatibi et al., 2014)	Video temporal segmentation Surgical instrument tracking	-	Color features Motion features	Decision tree for surgical instrument detection A novel proposed algorithm (based on Multi-objective GA) for temporal segmentation of video
(Yu et al., 2015)	Tissue segmentation	-	Color features	Threshold-based segmentation SVM region classifier
(Yip, Lowe, Salcudean, Rohling, & Nguan, 2012)	Tissue deformation tracking			
(Collins et al., 2011)				
(Stoyanov & Yang, 2011)				
(P. Y. Lee et al., 2015)				
(Dosis et al., 2005)	Surgical task recognition		Motion features	HMM
(Bouarfa & Dankelman, 2012)	Surgical task recognition	-	Motion features	Global pair-wise alignment
(Zappella et al., 2013)			Kinematic data Spatio-temporal features	Video clip classifiers

Datasets

Many researchers have stated that there is no publicly available dataset for laparoscopic image processing. Therefore, we present a list of datasets being used for validation of the previous studies in Table 2.

As listed in Table 1, many researches have used their datasets. One reason may be lack of any publicly available laparoscopic/endoscopic video dataset with ground-truth segmentation. This issue can introduce many limitations in implementing, improving and comparing the proposed works.

DISCUSSION

As mentioned in the previous sections, laparoscopic video processing is a challenging task due to moving background, respiratory motions of tissues, light reflection from the tissue surfaces, inhomogeneous and variable illumination in consequence video frames (Doignon et al., 2005), narrow field of view and lack of depth cue in conventional laparoscopy (Igarashi, Suzuki, & Naya, 2009).

Table 2. Datasets being used for validation of the previous studies for laparoscopic image processing applications

Authors	Research Problem	Dataset
(Marcinczak & Grigat, 2013)	Specular Reflection Segmentation	A dataset of 49 laparoscopic images taken from27 patients containing 269 true specular reflections. Ground truth segmentation of specular reflection has been performed manually.
(Lahane et al., 2012)	Unsafe action detection	900 image as train set 213 image as test set (positive and negative classes)
(Doignon et al., 2005)	segmentation of surgical instruments	Images from 3 video sequences of 500 color images
(Climent & Mars, 2010)	Instrument localization	128 images extracted from a real operation video
(Sa-Ing et al., 2012)	Instrument tracking	simulated videos for different situations
(Zhao, 2014)	3D visual tracking of laparoscopic instruments	a plastic phantom with a test-bench inside
(Collins et al., 2011)	deformable 3D surface reconstruction	Simulating a 3D kidney model comprising 1820 vertices
(Mountney & Yang, 2012)	tracking deforming tissue	simulated and in vivo MIS datasets of the liver, heart and abdomen
(Grasso et al., 2009)	Laparoscopic Video Summarization	378 images randomly selected from five laparoscopic cholecystectomy videos
(Iakovidis et al., 2010)	Endoscopic video summarization	annotated video frames with ground truth information labeled manually
(Oh et al., 2007)	Informative/ non-informative frame classification in endoscopic videos	70 frames were selected as a test set from three colonoscopy videos consisting of 35 informative frames and 35 non-informative frames
(Wang et al., 2016)	Laparoscopic video summarization and key-frame selection	a new gastroscopic video dataset has been proposed from 30 volunteers with more than 400k images
(Zappella et al., 2013)	Surgical gesture classification	Previously presented dataset

Many researches have tried to address some of the mentioned issues in their work. But, there is still need to propose more robust and stable methods for any of the mentioned applications in laparoscopic image processing.

CONCLUSION

In this study, an overview of the applications of image processing in laparoscopic surgeries are presented. The various applications include preprocessing video frames by laparoscopic image enhancement, telescope related applications (telescope position estimation, telescope motion estimation and compensation), surgical instrument related applications (surgical instrument detection and tracking), soft tissue related applications (soft tissue segmentation and deformation tracking) and high level applications such as safe actions in laparoscopic videos, summarization of laparoscopic videos, surgical task recognition and extracting knowledge using fusion techniques.

FUTURE PERSPECTIVES

A proposed direction in the future researches is tracking the surgical instruments and describing their interaction with soft tissues for surgical skill assessment.

Depth estimation is another important challenging issue in laparoscopic video processing (M. Nicolau, James, Benny, Darzi, & Yang, 2005). Another future research can be using depth information in laparoscopic videos. For this purpose, RGB-D cameras can be used as telescope in the laparoscopic surgeries.

Exploiting robotic technology in the laparoscopic surgeries with high popularity brings many advantages for single-port and multi-port laparoscopies such as 3D high-definition optics, better instrument performance and comforting the surgeon (Ramirez, Maurice, & Kaouk, 2016). It brings some new challenges and research opportunities by itself. For example, processing 3D images and using external sensors on the robots can be used. Moreover, detection and tracking the new instrumentation, describing the shape of surgical instruments and shape changes while performing different surgical activities are some other research opportunities that can be considered. Real-time image processing, finding the target region and magnifying it, adding more supplementary data for better describing the target region using the fusion techniques, and guiding the novice surgeons while performing the surgery are other future research directions.

REFERENCES

Agustinos, A., & Voros, S. (2016). 2D/3D Real-Time Tracking of Surgical Instruments Based on Endoscopic Image Processing. In X. Luo, T. Reichl, A. Reiter, & G. M. Mariottini (Eds.), *Computer-Assisted and Robotic Endoscopy* (pp. 90–100). Springer. doi:10.1007/978-3-319-29965-5_9

Akbari, H., & Kosugi, Y. (2007). Neural Network Blood Vessel Detection in Laparoscopy. *International Journal of Bioelectromagnetism*, *9*(1), 37–38.

Amini Khoyi, K., Mirbagheri, A., & Farahmand, F. (2016). Automatic tracking of laparoscopic instruments for autonomous control of a cameraman robot. *Minimally Invasive Therapy & Allied Technologies*, *25*(3), 121–128. doi:10.3109/13645706.2016.1141101 PMID:26872883

Amodeo, A., L., Q. A., J.V., J., E., B., & H.R., P. (2009). 2009, Robotic laparoscopic surgery: Cost and training. *The Italian journal of urology and Nephrology, 61*(2), 121-128.

Ballantyne, G. H. (2002). Robotic surgery, telerobotic surgery, telepresence, and telementoring. *Surgical Endoscopy and Other Interventional Techniques, 16*(10), 1389-1402.

Bashar, M. K., Kitasaka, T., Suenaga, Y., Mekada, Y., & Mori, K. (2010). Automatic detection of informative frames from wireless capsule endoscopy images. *Medical Image Analysis*, *14*(3), 449–470. doi:10.1016/j.media.2009.12.001 PMID:20137998

Baumhauer, M., Simpfendörfer, T., Müller-Stich, B. P., Teber, D., Gutt, C. N., Rassweiler, J., & Wolf, I. et al. (2008). Soft tissue navigation for laparoscopic partial nephrectomy. *International Journal of Computer Assisted Radiology and Surgery*, *3*(3-4), 307–314. doi:10.1007/s11548-008-0216-7

Blum, T., Feuner, H., & Navab, N. (2010). *Modeling and segmentation of surgical workflow from laparoscopic video.* Paper presented at the 13th international conference on Medical image computing and computer assisted intervention, Beijing, China. doi:10.1007/978-3-642-15711-0_50

Bouarfa, L., & Dankelman, J. (2012). Workflow mining and outlier detection from clinical activity logs. *Journal of Biomedical Informatics*, *45*(6), 1185–1190. doi:10.1016/j.jbi.2012.08.003 PMID:22925724

Braga, M., Frasson, M., Zuliani, W., Vignali, A., Pecorelli, N., & Di Carlo, V. (2010). Randomized clinical trial of laparoscopic versus open left colonic resection. *Br J Surg, 97*(8), 1180-1186.

Cano González, A., Vara Beceiro, I., Sánchez González, P., Pozo Guerrero, F. D., & Gómez Aguilera, E. J. (2008). Laparoscopic image analysis for automatic tracking of surgical tools. *Computer Assisted Radiology and Surgery 22nd International Congress and Exhibition (CARS 2008).*

Castellucci, S. A., Curcillo, P. G., Ginsberg, P. C., Saba, S. C., Jaffe, J. S., & Harmon, J. D. (2008). Single port access adrenalectomy. *J Endourol, 22*(8), 1573-1576.

Chen, B. H., Kopylov, A., Huang, S. C., Seredin, O., Karpov, R., Kuo, S. Y., & Hung, P. C. et al. (2016). Improved global motion estimation via motion vector clustering for video stabilization. *Engineering Applications of Artificial Intelligence*, *54*, 39–48. doi:10.1016/j.engappai.2016.05.004

Chen, C. J., Huang, W., & Song, K. T. (2013). *Image tracking of laparoscopic instrument using spiking neural networks.* Paper presented at the Control, Automation and Systems (ICCAS), Gwangju, China. doi:10.1109/ICCAS.2013.6704052

Choi, S. I., Lee, K. Y., Park, S. J., & Lee, S. H. (2010). Single port laparoscopic right hemicolectomy with D3 dissection for advanced colon cancer. *World J Gastroenterol, 16*(2), 275-278.

Climent, J., & Hexsel, R. A. (2012). Particle filtering in the Hough space for instrument tracking. *Computers in Biology and Medicine*, *42*(5), 614–623. doi:10.1016/j.compbiomed.2012.02.007 PMID:22425691

Climent, J., & Mars, P. (2010). Automatic instrument localization in laparoscopic surgery. In H. Bunke, J. J. Villanueva, G. Sanchez, & X. Otazu (Eds.), *Progress in computer vision and image analysis* (Vol. 73). Singapore: World Scientific Publishing Co.

Collins, T., Compte, B., & Bartoli, A. (2011). *Deformable Shape-From-Motion in Laparoscopy using a Rigid Sliding Window*. Paper presented at the MIUA.

D'Alessio, A., Piro, E., Tadini, B., & Beretta, F. (2002). One-trocar transumbilical laparoscopic-assisted appendectomy in children: our experience. *Eur J Pediatr Surg, 12*(1), 24-27.

Doignon, C., Graebling, P., & Mathelin, M. D. (2005). Real-time segmentation of surgical instruments inside the abdominal cavity using a joint hue saturation color feature. *Real Time Imaging*, *11*(5-6), 429–442. doi:10.1016/j.rti.2005.06.008

Dosis, A., Bello, F., Gillies, D., Undre, S., Aggarwal, R., & Darzi, A. (2005). Laparoscopic task recognition using Hidden Markov Models. *Studies in Health Technology and Informatics*, *111*, 115–122. PMID:15718711

Dou, J., & Li, J. (2014). Robust human action recognition based on spatio-temporal descriptors and motion temporal templates. *Optik - International Journal for Light and Electron Optics, 125*(7), 1891-1896.

Giannarou, S., Visentini-Scarzanella, M., & Yang, G. Z. (2012). *Probabilistic tracking of affine-invariant anisotropic regions*. Paper presented at the IEEE Trans Pattern Anal Mach Intell.

Grasso, M. A., Finin, T., Zhu, X., Joshi, A., & Yesha, Y. (2009). *Video Summarization of Laparoscopic Cholecystectomies*. Paper presented at the the AMIA 2009 Annual Symposium.

Hayashi, Y., Misawa, K., Hawkes, D. J., & Mori, K. (2016). Progressive internal landmark registration for surgical navigation in laparoscopic gastrectomy for gastric cancer. *International Journal of Computer Assisted Radiology and Surgery*, *1*(5), 837–845. doi:10.1007/s11548-015-1346-3 PMID:26811079

Hiroyasu, T., Tanaka, N., Hagiwara, A., & Ozamoto, Y. (2015). *Emphasizing mesenteric blood vessels in laparoscopic colon cancer surgery video images.* Paper presented at the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy. doi:10.1109/EMBC.2015.7318780

Iakovidis, D. K., Tsevas, S., & Polydorou, A. (2010). Reduction of capsule endoscopy reading times by unsupervised image mining. *Computerized Medical Imaging and Graphics*, *34*(6), 471–478. doi:10.1016/j.compmedimag.2009.11.005 PMID:19969440

Igarashi, T., Suzuki, H., & Naya, Y. (2009). Computer-based endoscopic image-processing technology for endourology and laparoscopic surgery. *International Journal of Urology*, *16*(6), 533–543. doi:10.1111/j.1442-2042.2009.02258.x PMID:19226356

Khatibi, T., Sepehri, M. M., & Shadpour, P. (2013). SIDF: A Novel Framework for Accurate Surgical Instrument Detection in Laparoscopic Video Frames. *International Journal of Hospital Research*, *2*(4), 163–170.

Khatibi, T., Sepehri, M. M., & Shadpour, P. (2014). A novel unsupervised approach for minimally-invasive video segmentation. *Medical Signals and Sensors*, *4*(1), 53–71. PMID:24695410

Kowalczuk, J., Meyer, A., Carlson, J., Psota, E. T., Buettner, S., Pérez, L. C., & Oleynikov, D. et al. (2012). Real-time three-dimensional soft tissue reconstruction for laparoscopic surgery. *Surgical Endoscopy*, *26*(12), 3413–3417. doi:10.1007/s00464-012-2355-8 PMID:22648119

Krupa, A., Doignon, C., Gangloff, J., & Mathelin, M. D. (2002). *Combined image-based and depth visual servoing applied to robotized laparoscopic surgery*. Paper presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland. doi:10.1109/IRDS.2002.1041409

Kviatkovsky, I., Rivlin, E., & Shimshoni, I. (2014). Online action recognition using covariance of shape and motion. *Computer Vision and Image Understanding*, *129*, 15–26. doi:10.1016/j.cviu.2014.08.001

Lahane, A., Yesha, Y., Grasso, M. A., Joshi, A., Park, A., & Lo, A. (2012). *Detection of unsafe action from laparoscopic cholecystectomy video.* Paper presented at the ACM SIGHIT International Health Informatics Symposium (ACM IHI). doi:10.1145/2110363.2110400

Lee, H. G., Choi, M. K., Shin, B. S., & Lee, S. C. (2013). Reducing redundancy in wireless capsule endoscopy videos. *Computers in Biology and Medicine*, *43*(6), 670–682. doi:10.1016/j.compbiomed.2013.02.009 PMID:23668342

Lee, P. Y., Yan, S. L., Hu, M. H., & Marescaux, J. (2015). *A computed stereoscopic method for laparoscopic surgery by using feature tracking*. Paper presented at the IEEE international conference on Consumer Electronics, Taipai, Taiwan. doi:10.1109/ICCE-TW.2015.7217046

Lee, S. L., Lerotic, M., Vitiello, V., Giannarou, S., Kwok, K. W., Visentini-Scarzanella, M., & Yang, G. Z. (2010). From medical images to minimally invasive intervention: Computer assistance for robotic surgery. *Computerized Medical Imaging and Graphics*, *34*(1), 33–45. doi:10.1016/j.compmedimag.2009.07.007 PMID:19699056

Li, B., & Meng, M. Q. H. (2009). Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments. *Computers in Biology and Medicine*, *39*(2), 141–147. doi:10.1016/j.compbiomed.2008.11.007 PMID:19147126

Lin, H. C., Shafran, I., Yuh, D., & Hager, G. D. (2006). Towards automatic skill evaluation: Detection and segmentation of robot-assisted surgical motions. *Computer Aided Surgery, 11*(5), 220-230.

Lin, Z., Uemura, M., Zecca, M., & Sessa, S. (2012). Objective Skill Evaluation for Laparoscopic Training Based on Motion Analysis. *IEEE Transactions on Bio-Medical Engineering*, *60*(4), 977–985. PMID:23204271

Marcinczak, J. M., & Grigat, R. R. (2013). Closed Contour Specular Reflection Segmentation in Laparoscopic Images. *International Journal of Biomedical Imaging*, 6. PMID:23983675

Marvik, R., Langø, T., Tangen, G. A., Andersen, J. O., Kaspersen, J. H., Ystgaard, B., . . . Nagelhus Hernes, T. A. (2004). Laparoscopic navigation pointer for three-dimensional image-guided surgery. *Surg Endosc, 18*, 1242-1248.

Moayedi, F., Azimifar, R., & Boostani, R. (2015). Structured sparse representation for human action recognition. *Neurocomputing*, *161*, 38–46. doi:10.1016/j.neucom.2014.10.089

Mountney, P., & Yang, G. Z. (2012). Context specific descriptors for tracking deforming tissue. *Medical Image Analysis*, *16*(3), 550–561. doi:10.1016/j.media.2011.02.010 PMID:21641270

Münzer, B., Schoeffmann, K., & Böszörmenyi, L. (2013). *Relevance Segmentation of Laparoscopic Videos*. Paper presented at the 2013 IEEE International Symposium on Multimedia (ISM), Anaheim, CA. doi:10.1109/ISM.2013.22

Navarra, G., Pozza, E., Occhionorelli, S., Carcoforo, P., & Donini, I. (1997). One-wound laparoscopic cholecystectomy. *British Journal of Surgery*, *84*(5), 695. doi:10.1002/bjs.1800840536 PMID:9171771

Nicolau, M., James, A., Benny, P. L., Darzi, A., & Yang, G. Z. (2005). *Invisible Shadow for Navigation and Planning in Minimal Invasive Surgery. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005* (Vol. 3750, pp. 25–32). Springer. doi:10.1007/11566489_4

Nicolau, S., Soler, L., Mutter, D., & Marescaux, J. (2011). Augmented reality in laparoscopic surgical oncology. *Surgical Oncology*, *20*(3), 189–201. doi:10.1016/j.suronc.2011.07.002 PMID:21802281

Oh, J. H., Hwang, S., Lee, J. K., Tavanapong, W., Wong, J., & Groen, P. C. D. (2007). Informative frame classification for endoscopy video. *Medical Image Analysis*, *11*(2), 110–127. doi:10.1016/j.media.2006.10.003 PMID:17329146

Öhsen, U. V., Marcinczak, J. M., Vélez, A. F. M., & Grigat, R. R. (2012). *Keyframe selection for robust pose estimation in laparoscopic videos.* Paper presented at the Medical Imaging 2012: Image-Guided Procedures, Robotic Interventions, and Modeling, San Diego, CA.

Oropesa, I., Sanchez-Gonzalez, P. J., Chmarra, M. K., Lamata, P., Fernandez, A. I., Sanchez-Margallo, J. A., & Gomez, E. J. et al. (2013). EVA: Laparoscopic Instrument Tracking Based on Endoscopic Video Analysis for Psychomotor Skills Assessment. *Surgical Endoscopy*, *27*(3), 1029–1039. doi:10.1007/s00464-012-2513-z PMID:23052495

Padoy, N., Blum, T., Essa, I., Feussner, H., Berger, M., & Navab, N. (2007). *A boosted segmentation method for surgical workflow analysis.* Paper presented at the Medical Image Computing and Computer Assisted Intervention. doi:10.1007/978-3-540-75757-3_13

Podolsky, E. R., Mouhlas, A., Wu, A. S., Poor, A. E., & Curcillo, P. G. (2010). Single Port Access (SPA) laparoscopic ventral hernia repair: initial report of 30 cases. *Surg Endosc, 24*(7), 1557-1561.

Quellec, G., Charrière, K., Lamard, M., Droueche, Z., Roux, C., Cochener, B., & Cazuguel, G. (2014). Real-time recognition of surgical tasks in eye surgery videos. *Medical Image Analysis*, *18*(3), 579–590. doi:10.1016/j.media.2014.02.007 PMID:24637155

Ramirez, D., Maurice, M. J., & Kaouk, J. H. (2016). Robotic Single-port Surgery: Paving the Way for the Future. *Urology*, *95*, 5–10. doi:10.1016/j.urology.2016.05.013 PMID:27211930

Ronaghi, Z., Duffy, E. B., & Kwartowitz, D. M. (2015). Toward real-time remote processing of laparoscopic video. *medical. Imaging*, *2*(4).

Rosen, J., Solazzo, M., Hannaford, B., & Sinanan, M. (2002). Task Decomposition of Laparoscopic Surgery for Objective Evaluation of Surgical Residents' Learning Curve Using Hidden Markov Model. *Computer Aided Surgery, 7*(1), 49-61.

Sa-Ing, V., Thongvigitmanee, S. S., Wilasrusmee, C., & Suthakorn, J. (2012). Adaptive Mean-Shift Kalman Tracking of Laparoscopic Instruments. *International Journal of Computer Theory and Engineering*, *4*(5), 685–689. doi:10.7763/IJCTE.2012.V4.557

Sánchez-González, P., Cano, A. M., Oropesa, I., Sánchez-Margallo, F. M., Del Pozo, F., Lamata, P., & Gómez, E. J. (2011). Laparoscopic video analysis for training and image-guided surgery. *Minimally Invasive Therapy & Allied Technologies*, *20*(6), 311–320. doi:10.3109/13645706.2010.541921 PMID:21247251

Schols, R. M., Bouvy, N. D., Dam, R. M. V., & Stassen, L. P. S. (2013). Advanced intraoperative imaging methods for laparoscopic anatomy navigation: An overview. *Surgical Endoscopy*, *27*(6), 1851–1859. doi:10.1007/s00464-012-2701-x PMID:23242493

Selka, F., Nicolau, S., Agnus, V., Bessaid, A., Marescaux, J., & Soler, L. (2015). Context-specific selection of algorithms for recursive feature trackingin endoscopic image using a new methodology. *Computerized Medical Imaging and Graphics, 40*, 49-61.

Semerjian, A., Zettervall, S. L., Amdur, R., Jarrett, T. W., & Vaziri, K. (2015). 30-Day morbidity and mortality outcomes of prolonged minimally invasive kidney procedures compared with shorter open procedures: National surgical quality improvement program analysis. *Journal of Endourology*, *29*(7), 830–837. doi:10.1089/end.2014.0795 PMID:25646859

Shu, Y., & Cheriet, F. (2005). *Segmentation of Laparoscopic Images: Integrating Graph-Based Segmentation and Multistage Region Merging.* Paper presented at the Computer and Robot Vision.

Speidel, S., Kroehnert, A., Bodenstedt, S., Kenngott, H., Müller-Stich, B., & Dillmann, R. (2015). *Image-based tracking of the suturing needle during laparoscopic interventions.* Paper presented at the Medical Imaging 2015: Image-based tracking of the suturing needle during laparoscopic interventions, Orlando, FL.

Stanek, S. R., Tavanapong, W., Wonga, J., Oh, J. H., & Groenc, P. C. D. (2012). Automatic real-time detection of endoscopic procedures using temporal features. *Computer Methods and Programs in Biomedicine, 108*, 524-535.

Stoyanov, D., Elson, D., & Yang, G. Z. (2009). *Illumination position estimation for 3D soft-tissue reconstruction in robotic minimally invasive surgery.* Paper presented at the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO. doi:10.1109/IROS.2009.5354447

Stoyanov, D., & Yang, G. Z. (2011). *Soft tissue deformation tracking for robotic assisted minimally invasive surgery*. Paper presented at the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN.

Uecker, D. R., Wang, Y. F., Lee, C., & Wang, Y. (1995). Laboratory Investigation:Automated Instrument Tracking in Robotically Assisted Laparoscopic Surgery. *Journal of Image Guided Surgery*, *1*(6), 308–325. doi:10.1002/(SICI)1522-712X(1995)1:6<308::AID-IGS3>3.0.CO;2-E PMID:9080352

Voros, S., Long, J. A., & Clinquin, P. (2007). Automatic Detection of Instruments in Laparoscopic Images: A First Step Towards High-level Command of Robotic Endoscopic Holders. *The International Journal of Robotics Research*, *26*(11-12), 1173–1190. doi:10.1177/0278364907083395

Wang, S., Cong, Y., Cao, J., Yang, Y., Tang, Y., Zhao, H., & Yu, H. (2016). Scalable gastroscopic video summarization via similar-inhibition dictionary selection. *Artificial Intelligence in Medicine*, *66*, 1–13. doi:10.1016/j.artmed.2015.08.006 PMID:26363682

Wei, G. Q., Arbter, K., & Hirzinger, G. (1997). Real-time visual servoing for laparoscopic surgery. Controlling robot motion with color image segmentation. *IEEE Engineering in Medicine and Biology Magazine*, *16*(1), 40–45. doi:10.1109/51.566151 PMID:9058581

Wolf, R., Duchateau, J., Cinquin, P., & Voros, S. (2011). *3D Tracking of Laparoscopic Instruments Using Statistical and Geometric Modeling. In Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011* (Vol. 6891, pp. 203–210). Springer. doi:10.1007/978-3-642-23623-5_26

Wu, C. H., Chen, Y. C., Liu, C. Y., Chang, C. C., & Sun, Y. N. (2004). *Automatic extraction and visualization of human inner structures from endoscopic image sequences.* Paper presented at the SPIE2004. doi:10.1117/12.535880

Yip, M. C., Lowe, D. G., Salcudean, S. E., Rohling, R. N., & Nguan, C. Y. (2012). Real-Time Methods for Long-Term Tissue Feature Tracking in Endoscopic Scenes. In P. Abolmaesumi, L. Joskowicz, N. Navab, & P. Jannin (Eds.), *Information Processing in Computer-Assisted Interventions* (pp. 33–43). Springer Berlin Heidelberg. doi:10.1007/978-3-642-30618-1_4

Yu, Z., Liu, J., Ding, H., Lu, M., Deng, M., Huang, Y., & Li, K. et al. (2015). An Image-Based Method of Uterus Segmentation in Gynecologic Laparoscopy. *Journal of Medical Imaging and Health Informatics*, *5*(4), 819–825. doi:10.1166/jmihi.2015.1463

Yuan, Y., Zheng, X., & Lu, X. (2016). A discriminative representation for human action recognition. *Pattern Recognition*, *59*, 88–97. doi:10.1016/j.patcog.2016.02.022

Zappella, L., Béjar, B., Hager, G., & Vidal, R. (2013). Surgical gesture classification from video and kinematic data. *Medical Image Analysis*, *17*(7), 732–745. doi:10.1016/j.media.2013.04.007 PMID:23706754

Zenbutsu, S., Igarashi, T., & Yamaguchi, T. (2013). Development of Blood Vessel Depth Displaying Method for Laparoscopic Surgery Guidance. *Journal of Medical Imaging and Health Informatics*, *3*(1), 101–106. doi:10.1166/jmihi.2013.1134

Zhang, X., & Payandeh, S. (2002). Application of Visual Tracking for Robot-Assisted Laparoscopic Surgery. *Journal of Robotic Systems*, *19*(7), 315–328. doi:10.1002/rob.10043

Zhang, Y., Wirkert, S. J., Iszatt, J., Kenngott, H., Wagner, M., Mayer, B., . . . Maier-Hein, L. (2016). *Tissue classification for laparoscopic image understanding based on multispectral texture analysis.* Paper presented at the Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling.

Zhao, Z. (2014). Real-time 3D visual tracking of laparoscopic instruments for robotized endoscope holder. *Bio-Medical Materials and Engineering*, *24*, 2665–2672. PMID:25226970

KEY TERMS AND DEFINITIONS

Global Motion Compensation: Trying to reverse the camera motion to improve the accuracy of the local motion estimation.

Image Enhancement: Trying to improve the image quality.

Image Segmentation: Dividing the image into several homogeneous regions.

Laparoscopy: A minimally-invasive surgical procedure using a camera to display the inside patient's body.

Soft Tissue Deformation Tracking: Trying to estimate the parameters of the soft tissue deformation model.

Surgical Activity Recognition: Trying to detect which surgical activity is being performed in the laparoscopic video.

Telescope: A camera inserted into the patient's body.