International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

A Survey on Surgical Phase Recognition Approaches

Sisira C Sunny[1], Sabitha Raju[2]
[1] PG Student, [2] Asst. Professor, Department of Computer Science, Viswajyothi College of Engineering and Technology, Vazhakulam

Abstract: Laparoscopic surgeries are famous due to its numerous advantages. Video based monitoring is incorporated among such surgeries. These are really useful for the surgeons for performing the surgery. Surgical tool detection and phase or workflow recognition are the major issues in medical field. These are really useful for training surgeons and clinicians. It has also applications in robot-assisted surgery. Since a video camera is inserted along with the tools, no additional setups are needed. These are the reasons behind the popularity of tool detection and workflow recognition tasks. This paper presents some of the common phase recognition approaches. The advantages and disadvantages of these methods are also discussed.
Keywords: Workflow Recognition, Phase Recognition, Robot-Assisted Surgery, Minimally Invasive Surgery

I. INTRODUCTION

Medical field is one of the fields where tremendous innovations are incorporated. Many medical procedures have been digitalized. This reduces the complexity of medical procedures. Digital strategies are highly acceptable in this area.

Several new technologies are nowadays incorporated to make the Operating Room (OR) much more efficient. Many kind of surgeries like cataract, neurological, and laparoscopic surgeries use a video camera as an observation tool. Surgical phase recognition and tool detection are challenging problems in this context.

Surgical phase recognition aims at automatically recognizing the phases of a surgery. One of the most expensive stage is training the surgeons. It is not always possible to accommodate an experienced surgeon with each trainee. So the best way is to monitor them. So for this, the surgical phase or task performed must be tracked. It can reduce the risks associated with the surgery. It can also decrease the chances of faulty procedures. Surgical workflow recognition can also give automated assistance to clinical staff. It can be also used to automatically index the video databases. Surgical workflow recognition also contributes to robot-assisted surgeries. Monitoring clinicians can easily track malfunctions.

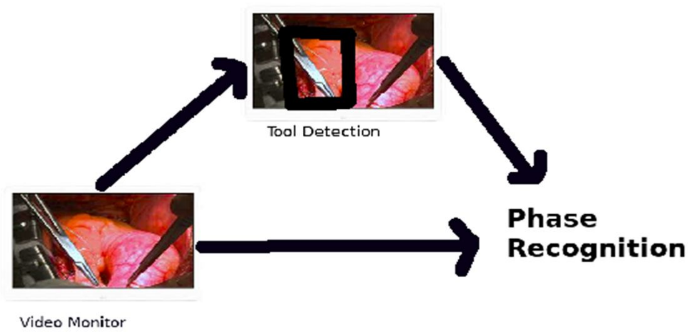

Fig. 1 Phase Recognition

II. RELATED WORKS

M. S. Holden et al. [1] proposed a workflow segmentation algorithm. Clinical trainees normally study and practice things under the guidance of an expert. But this is tedious and it needs lots of effort. To overcome this limitation, a computer-assisted training system can be used. The proposed algorithm [1] works in this scenario. First the algorithm is trained using a training data set and then it is tested using a test dataset. The proposed algorithm [1] first remove the noise in the tool tracking process with the use of a Gaussian Filter. To map tool path that is in time to vector, orthogonal transformation is used. Since orthogonal transformation convert the data in a higher dimension, it may create problems in the data mining task. So principal component analysis is used to map the data from higher to a lower dimension. K-means clustering algorithm is used to map data into class labels. To validate the algorithm, Ultrasound-guided epidural procedures and Lumbar Puncture Procedure are used.

The workflow is as follows [1]:
1) Finding the insertion point
2) Finding the insertion angle
3) Finding the insertion depth
4) Verifying the target and
5) Retracting the needle

Surgical gesture detection in robot-assisted MIS is a challenging task. The approach proposed by H. C. Lin et al.[2] encompasses automatic techniques to solve this challenge. The da Vinci API data is used. Eight particular gestures are selected for the experiments. Proposed technique [2] first perform local feature extraction from the API data. Abrupt changes in the surgical motion are very rare. So adjacent input samples can be utilized. Then the features are normalized to a range. Linear discriminant analysis is performed for increasing the accuracy of recogniser. A Bayes classifier is built.

Experts are always trying to make operating rooms much more efficient. Real-time workflow identification systems will be an essential element of operating rooms in the future. The paper proposed by N. Padoy et al. [3] address this. In this paper [3], two models are constructed for phase recognition. Average surgery is one among them. This model is based on a training dataset synchronised using DTW. The other one is a Hidden Markov Model (HMM). HMMs are used to model the probabilistic properties of a training dataset [3]. With the help of these models off-line segmentation and on-line recognition of surgical phase is completed. This approach can be used for phase recognition in any endoscopic surgeries.

The workflow detection method [4] proposed by O. Dergachyova et al. consist of 4 stages. It detects surgical workflow with the help of video data and instrument usage signal. MICCAI 2015 EndoVis dataset is used for validating the experiments [4]. Surgical Process Modelling (SPM) is used for implementing the idea. To describe the input data visual description of frame and tool presence are taken into account. Colour, shape and texture aspects of the images are examined. Colour representation is done in the form of histograms. Shape description is done by Discrete Cosine Transform (DCT) and Histograms of Oriented Gradients (HOG). The texture is represented by Local Binary Patterns (LBP) histograms. AdaBoost classifier is used for classifying the image samples. Hidden semi-Markov Model (HsMM) is used for modelling the temporal aspect. The matrices used for validation are Average transitional delay (ATD), Noise level (NL), Coefficient of transitional moments (CTM)
and Application-dependent scores (AD-scores).

There may be a single person doing the activity in an operating room or one or two persons in collaboration with each other. So recognizing the activity in such a scenario is a challenge. The approach [5] proposed by J. E. Bardram et al. is a sensor based one. The platform consists of sensors that are embedded in operating room as well as instruments. It also has body-worn sensors for the clinicians to track the tools used. For identifying the activity, the information to be tracked include location of the clinician, location of the patient, location of objects in the table, and use of objects and tools by the surgeon. To track the location, Ubisense real time location tracking system (RTLS) [5] is used. This will locate a person wearing a tag and returns the x, y, z, coordinates. Because of the reduced accuracy of RTLS system, a buffer is introduced. All the instruments contain passive RFID tags. Tables contain RFID readers. The third sensor doesn’t contain any wires. It is a palm-based sensor to track the instrument usage. The sensor data is then sampled and synchronised. Then it is transformed to features. Here Decision Tree classifier is used.

Nowadays, the operating rooms became more rich and complex. A lot of sensors like microscopes and endoscopes are deployed. The information obtained from the video data can be integrated with information from other sensors to model surgical procedures. F. Lalys et al. proposed a framework[6] that can be used for other types of surgeries because it is highly adaptable. But this paper make use of this in cataract surgeries. Cataract surgery usually remove the damaged lens of the eye and insert an artificial one. In this [6], the first step is the pupil segmentation. Visual cues are always associated with pupil. Hence region around the pupil may bias the segmentation procedure. So, here, the pupil is the ROI (Region Of Interest). Pupil is identified using the colour difference, because it is the darkest region and the remaining regions are light. To segment the pupil, first a mask is created by applying smoothing first. Then, some image processing operations are applied to the input image. Then, Hough transform is applied to detect the circle in eye. The third step is used to detect the most probable circular zone if there is a failure in Hough transform. Histogram intersection principle is used to extract colour-oriented information. This step uses a training dataset with positive and negative images. Bag-of-visual-words (BVW) representation is used for texture. BWV treat images as a collection of patches. Then some representative patches are selected and for each of them a visual descriptor vector is evaluated. The resulting distribution is used to characterize the image. To detect key points SIFT, SURF , Harris and STAR are tested. In this paper, mainly SURF is used. Instrument detection is a challenging task here because most of the instruments have same shape. Therefore, two methods are proposed. One is to just detect the presence of instrument. Second is to detect and categorize specific instruments. For detecting specific instruments, Viola Jones object detection framework is used. Then, AdaBoost classifier is built. For instrument presence detection, they used a pixel based approach that uses the colour difference between instrument and the eye. Noise is removed using connected component method. By applying this mask, we get ROI. For description and segmentation, SVM classifier with BVW is used. An alternate method is also proposed. HMM is used for time-series data modelling. Here the output of SVM is the observations of HMM and the states are different phases. Dynamic time warping is also used.

The method [7] proposed by G. Forestier et al. uses decision tree classifier. Surgeries are basically a sequence of activities done by a surgeon in an operating room. Here, the triplet representation of surgical activities are used; ie., action, anatomical structure, instrument . Activities performed by both hands and the use of microscope are noted.

e.g.,{(cut, scissors, muscle)r, (hold, retractors, muscle)l, false}

A training dataset is constructed using a set of known phases. Previous activities are analysed, then the probability density functions (PDF) of the predictions are combined. Experiments for phase prediction are conducted based on single activity, local context, noisy data and among cluster of surgeries.

Analysis of human-activity from videos have gained interest in computer-vision nowadays. This is used in medical monitoring and surveillance. Here activities are considered as workflow or phases. Temporal dependencies between phases is an important challenge. Here action is a fundamental element in scene. Several actions together are called activities. Meaningful group of activity is called phase. Phase occur in repetition. Set of phases are called workflow [8]. Phase recognition aims at detecting the occurrence of phase. For this, construct a model first. The next problems are off-line and on-line recognition. A model called Workflow-HMM is proposed by N. Padoy et al. in this paper [8]. This is a statistical model [8]. The proposed approach [8] consists of a two-level hierarchy. It is used to model dependencies between different phases and dependencies within each phase. WHMM is modelled using labelled and unlabelled observations. Model parameters are initialized first. There are multiple cameras. Hence we get 3D-grids. Features that describe spatial distribution of motion are extracted, eg: patient entry can be identified by a strong distribution at the OR door. Phases are identified using motion patterns. For computing the feature, first reconstructed volume is splitted into cells and for each cell a 3D histogram is computed. Resulting flow vectors are then quantized.

A lot of cases were reported where patients die due to medical errors. To prevent these kind of errors something has to be done. Hence risk identification seems to be really important. The conversation between surgeons and staffs in operating room contain vital information about the surgical task performed. T. Suzuki et al.[9] proposed an audio recording and analysis system. Video cameras and microscopes ate fitted all over the room. Here it is 6, i.e., 6 channels. The videos from these channels are synchronized. Seven computers are deployed.

The data form each channel goes to a single computer. The remaining one is the server. Microphones record audio information. A software is developed which consists of three modules- input module, analysis module and output module. If the movie file size is larger it indicate that some anomaly happened or a phase is happening. In the case of recorded sound, too much or too less of conversation indicate a phase change.

Robot assisted MIS has more benefits than the traditional surgical procedures. Three methods for surgical video classification are proposed [10] by L. Zappella et al.

The first one is linear dynamical systems (LDSs) -based. It is used to model the features' time series that are extracted from each video. Gesture classification training uses metrics in LDSs. Different features, and different metrics that are important in the context of LDS are evaluated.

The second one is a bag-of-features (BoF) approach. Here a bag of words is learned. Each clip of video is then plotted with a histogram. Classifiers are then trained. An intense analysis of all the components and their variations are discussed. The third approach integrates the two previously mentioned approaches. It uses multiple kernel learning (MKL).

Several studies are conducted in the area of laparoscopic surgery as it is one of the most challenging image processing area. Surgical workflow recognition has gained interest because of its numerous applications. Unlike other image processing areas, laparoscopy is much more challenging because of dynamic nature of camera and blood stains in lens. Tools used in the surgery can be utilized for finding phase of surgery.

The approach proposed byA. P. Twinanda et al. [11] consists of a novel method for phase recognition based on Convolutional Neural Network (CNN). This uses Alexnet as the basis. This enables automatic learning of features from laparoscopic videos. It also uses tool presence signals to guide phase recognition.

TABLE I
COMPARISON OF METHODS

Sl No.	Method	Advantage	Disadvantage
1.	M. S. Holden et al.[1]	•Tasks in the workflow are repeated and can be occurred in any sequence. •Sequence of tasks that are unknown are segmented. •Real-time workflow segmentation.	•No feedback to the user.
2.	H. C. Lin et al.[2]	•Automatic Technique. •Perform surgical gesture detection and segmentation.	•Surgical skill levels are not incorporated.
3.	N. Padoy et al.[3]	•Statistical model for surgical workflow constructed from Dynamic Time Warping (DTW) and Hidden Markov Model (HMM).	•Not suitable for complex workflows.
4.	O. Dergachyova et al.[4]	•Automatic segmentation and recognition of surgical phases and instrument presence detection in real time.	•Standard visual features are applied.
5.	J. E. Bardram et al.[5]	•High classification accuracy. •Ignoring data from RTLS didn't affect the accuracy of classification.	•Overhead in the case of wearing sensors.
6.	F. Lalys et al.[6]	•High accuracy (94%). •Adaptable.	•Doesn't include surgeon's gestures.
7.	G. Forestier et al.[7]	•Accepted amount of precision with limited amount of data. •Low complexity of decision tree in terms of computation allows its integration with low-power devices.	•Only four phases are concerned.
8.	N. Padoy et al.[8]	•Due to generality of features, it can be used to other workflow.	•Online results are slightly worse than offline. •Online recognition of phases with varying length seems to be difficult.
9.	T. Suzuki et al.[9]	•Simple. •Easy to implement.	•Less accuracy. •Cannot identify the phase name.
10.	L. Zappella et al.[10]	•The unique combination outperform all other existing techniques.	•Low-level visual features are used. •Not a generalized model.
11.	A. P. Twinanda et al.[11]	•Good accuracy	

III.CONCLUSIONS

Nowadays, workflow recognition systems have been receiving acceptance worldwide. Surgical workflow or phase recognition is a challenging problem. There are several methods that are proposed to solve these problems. Some works are proposed to just find out workflow changes. They doesn't tries to find out the exact name of workflow or its intentions. Most of the works are designed to solve the training cost associated with clinical training procedures. Majority of works are mainly based on training and testing strategies. The main idea behind workflow recognition is the instrument detection. Because all the surgical phase make use of surgical tools. So, the tools are a crucial factor that uniquely identifies a workflow. All of the works uses some kind of noise removal strategies to enhance the results. To accurately define the sequence of workflow, most of the approaches uses HMM and it's variants. Several hand crafted features are used for training and testing. Very few methods uses location tracking and sensor systems to find out the surgical phase. But they end up in huge overhead. These approaches are somehow context sensitive. Some works are for endoscopic surgeries. Some other works are for cataract surgeries and so on. Very few approaches uses 3D modelling procedures. These works are really complex. But the main problem is to find out a real-time solution that accurately detect and recognize surgical phases by making use of unique strong features. Some of these methods works really well.

REFERENCES

[1] M. S. Holden, T. Ungi, D. Sargent, R. C. McGraw, E. C. Chen, S. Ganapathy, T. M. Peters, and G. Fichtinger, "Feasibility of real-time workflow segmentation for tracked needle interventions," IEEE Transactions on Biomedical Engineering, vol. 61, no. 6, pp. 1720–1728, 2014.

[2] H. C. Lin, I. Shafran, T. E. Murphy, A. M. Okamura, D. D. Yuh, and G. D.Hager, "Automatic detection and segmentation of robot-assisted surgical motions," in International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 802–810, Springer, 2005.

[3] N. Padoy, T. Blum, S.-A. Ahmadi, H. Feussner, M.-O. Berger, and N. Navab, "Statistical modeling and recognition of surgical workflow,"Medical image analysis, vol. 16, no. 3, pp. 632–641, 2012.

[4] O. Dergachyova, D. Bouget, A. Huaulm ́e, X. Morandi, and P. Jannin, "Automatic data-driven real-time segmentation and recognition of surgical workflow,"International journal of computer assisted radiology and surgery, vol. 11, no. 6, pp. 1081–1089, 2016.

[5] J. E. Bardram, A. Doryab, R. M. Jensen, P. M. Lange, K. L. Nielsen, and S. T. Petersen, "Phase recognition during surgical procedures using embedded and body-worn sensors," in Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on, pp. 45–53,IEEE, 2011.

[6] F. Lalys, L. Riffaud, D. Bouget, and P. Jannin, "A framework for the recognition of high-level surgical tasks from video images for cataract surgeries," IEEE Transactions on Biomedical Engineering, vol. 59, no. 4, pp. 966–976,2012.

[7] G. Forestier, L. Riffaud, and P. Jannin, "Automatic phase prediction from low-level surgical activities,"International journal of computer assisted radiology and surgery , vol. 10, no. 6, pp. 833 841, 2015.

[8] N. Padoy, D. Mateus, D. Weinland, M.-O. Berger, and N. Navab, "Workflow monitoring based on 3d motion features," in Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on,pp. 585–592, IEEE, 2009.

[9] T. Suzuki, Y. Sakurai, K. Yoshimitsu, K. Nambu, Y. Muragaki, and H. Iseki, "Intraoperative multichannel audio-visual information recording and automatic surgical phase and incident detection," in Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, pp. 1190–1193, IEEE, 2010.

[10] L. Zappella, B. B ́ejar, G. Hager, and R. Vidal, "Surgical gesture classification from video and kinematic data,"Medical image analysis, vol. 17,no. 7, pp. 732–745, 2013.

[11] A. P. Twinanda, S. Shehata, D. Mutter, J. Marescaux, M. De Mathelin, and N. Padoy, "Endonet: A deep architecture for recognition tasks on laparoscopic videos,"IEEE transactions on medical imaging, vol. 36, no. 1,pp. 86–97, 2017