

Vision-Assisted Automatic Detection and Segmentation of Robot-Assisted Surgical Motions

Henry Lin (hcl@cs.jhu.edu), Izhak Shafran, David D. Yuh, Gregory D. Hager (hager@cs.jhu.edu)

The Johns Hopkins University and The Johns Hopkins Medical Institutions

Problem

Surgical training and evaluation has traditionally been an interactive and slow process in which interns and junior residents perform operations under the supervision of a faculty surgeon. This method of training lacks an objective means of quantifying and assessing surgical skills. Economic pressures to reduce the cost of training surgeons and national limitations on resident work hours have created a need for efficient methods to supplement traditional training paradigms. While surgical simulators aim to provide such training, they have limited impact as a training tool since they are generally operation specific and cannot be broadly applied.

Recent approaches to surgical skill evaluation have had success but still lack manual interpretation, flexibility, scope and/or accuracy. Evaluating surgical skill is a complex task, even for a trained faculty surgeon. To achieve automatic recognition of even elementary motions that occur in a simplified surgical task, the system would require complex machine learning algorithms, and, potentially, a large number of parameters. With the advent of robot-assisted minimally invasive surgical system, such as Intuitive Surgical's da Vinci, the ability to record quantitative motions and video data opens up the possibility of creating simple descriptive, mathematical models to recognize and analyze surgical training and performance.

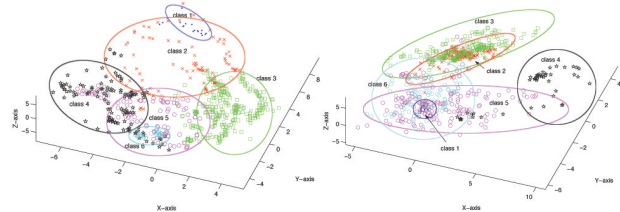
Method

Our group in the Engineering Research Center for Computer Integrated Surgical Systems (CISST-ERC) has been investigating the problem of recognizing simple elementary motions that occur in a suturing task performed on the da Vinci robot. We have attempted to differentiate the source of each motion - by an expert surgeon or an intermediate surgeon. We divide the task into functional modules (as described below) akin to other pattern recognition tasks such as automatic speech recognition. We extend our current research to include a computer vision component, which we describe as vision-assisted motion analysis.

1. **Local feature extraction** – Surgical motion seldom abruptly changes from one gesture to another. Information from adjacent input samples is useful in improving the accuracy and robustness of recognizing a surgical motion. We concatenate feature vectors at time t with neighboring samples to give temporal influence.
2. **Feature normalization** – The range and units of values of within each feature vector are different (i.e. position and velocity). We normalize the mean and variance of each dimension so that the difference in dynamic range does not hurt performance of the classifier.
3. **Linear discriminant analysis** – When the features corresponding to different surgical motions are well separated, the accuracy of the recognizer can be considerably improved. Our experiments have shown linear discriminant analysis as a transformation that

effectively does this. The dimension reduction has a range of three to six.

4. **Bayes classifier** – We implement a Bayes classifier as the optimal discriminant function to classify the elementary motions.
5. **Computer vision** – Using stereo video data synchronized with the feature API data recorded from the daVinci robot, we can use visual cues and vision analysis to help determine the motion at a particular time unit, and to verify motion transition points.
6. **Validation study** – A set of eight elementary motions are defined in the motion set. The expert surgeon data consisted of 15 suture trials with each trial consisting of four throws and lasting at least 400 time units. Each feature vector consists of 72 features. A 15-fold cross-validation paradigm is used to maintain integrity.



The result of LDA reduction of 6 motion classes in three dimensions. The expert surgeon's motions (left) separate more distinctly than the intermediate surgeon's (right).

Results

We have implemented the first four components and preliminary results from our validation study consistently exceed 90% motion recognition rate for expert surgeon data. Varying the LDA reduction dimension and number of motion classes does affect the recognition rate. We expect the rate of recognition to increase when we add the computer vision component to the system. The computer vision component can provide a different technical perspective on surgical motion recognition.

Conclusion and Discussion

The preliminary results have shown that linear discriminant analysis is a robust tool in assisting surgical motion recognition. With promising results, we want to increase our recognition rate and statistical significance by adding a computer vision component to our feature and recognition processing, and also other statistical methods such as hidden Markov models. The system is not confined in its scope and can be applied to any surgical task, as long as elementary motions can be defined.

References

- [1] H. Lin, I. Shafran, T. Murphy, A.M. Okamura, D. Yuh, G.D. Hager, "Automatic Detection and Segmentation of Robot-Assisted Surgical Motions," MICCAI, 2005.
- [2] J. Rosen, et al. Task decomposition of laparoscopic surgery for objective evaluation of surgical resident's curve using hidden Markov model. In: Computer Aided Surgery. Volume 7(1). (2002) 49-61.