Published in final edited form as:

IEEE Trans Biomed Eng. 2010 February ; 57(2): 479–487. doi:10.1109/TBME.2009.2031982.

A Step Towards Identification of Surgical Actions in Mastoidectomy

Uttama Lahiri

Department of Mechanical Engineering, Vanderbilt University, Nashville, TN 37235-1592 USA
(phone: 615-438-3502; uttama.lahiri@vanderbilt.edu)

Robert F. Labadie

Department of Otolaryngology, Vanderbilt University Medical Center, Nashville, TN, USA
(robert.labadie@Vanderbilt.Edu)

Changchun Liu

The Mathworks Inc., Natick, USA (changchun.liu@mathworks.com)

Ramya Balachandran

Department of Otolaryngology, Vanderbilt University Medical Center, Nashville, TN, USA
(ramya.balachandran@Vanderbilt.Edu)

Omid Majdani

Department of Otolaryngology, Medical University of Hannover, Hannover, Germany
(omid.majdani@gmx.de)

Nilanjan Sarkar

Department of Mechanical Engineering and the Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235-1592 USA
(nilanjan.sarkar@vanderbilt.edu)

Abstract

Mastoidectomy is a core surgical procedure in otologic surgery. It is believed that the procedure is performed by different surgeons with some variability. However, it is also believed that all surgeons use a finite number of fundamental surgical actions to complete the procedure. To determine how a surgeon performs a mastoidectomy, we sought to identify the fundamental surgical actions (called Action Primitives, APs) and determine the transition boundaries between those APs. Our motivation for this work is both to delineate the APs necessary to complete a mastoidectomy and to optimize and potentially automate the surgical process. In this paper, we present a new approach to developing methods for parsing raw data (position and orientation of the surgical tool and end-effector force) into a sequence of surgical tasks. The overall objective is to deconstruct the surgical procedure into a series of Action Primitives (APs). This paper presents results from our initial investigation on detecting transition boundaries and identifying action primitives involved in mastoidectomy.

Index Terms

Action Primitive (AP); identification of APs; Mastoidectomy; transition boundaries between APs (i.e., segmentation)

I. Introduction

Ear surgery is undertaken to treat a number of ear disorders including chronic middle ear infections, cholesteatoma (epithelial cyst within the middle ear), conductive hearing loss, vertigo (dizziness), and sensorineural hearing loss. The central component of these surgeries is the mastoidectomy—a procedure that can take upwards of 2 hours. Each year about 100,000 mastoidectomy surgeries are performed within the United States [1].

During mastoidectomy, the temporal bone, the bone that encases the ear, is systematically drilled away to either remove diseased tissue or provide access to the middle and/or inner ear. The procedure is akin to an archeological dig where it is vital to preserve certain noble structures that reside within the temporal bone. These structures include the facial nerve (injury results in paralysis of the face) [2], the middle ear (malleus, incus and stapes – injury results in conductive hearing loss), the inner ear (injury results in permanent hearing loss and vertigo), the floor of the cranial vault (injury results in leakage of cerebrospinal fluid) [3] and the internal jugular vein and carotid artery (injury results in blood loss which may be life threatening). The level of accuracy needed to prevent damage to these structures is in the order of 1mm [4]. However, the fundamental actions that surgeons use to complete the surgery appear to be repeatable among surgeons with variation in order and style.

The mastoidectomy procedure performed by different surgeons can vary widely [36]. In an endeavor to standardize and automate the major components of mastoidectomy, it is first necessary to understand how a surgeon performs mastoidectomy and determine the basic sequence of surgical actions necessary to complete the procedure. The knowledge thus gained would be helpful for training of novice surgeons and comparing their skill sets with that of an experienced surgeon. Perhaps more clinically significant, by defining the actions that comprise a procedure, we may be better able to optimize the order of the actions and potentially program a robot to perform at least a part of surgical procedure in the future.

Robots have been used in operating theaters for over 20 years. Examples include active-constraint robots such as, the Acrobot (The Acrobot Company Limited, London, UK) used in preparing bone surfaces for positioning prostheses in total knee replacement surgery [5 – 6], master-slave telemanipulators such as the da Vinci Surgical System® (Intuitive Surgical; Sunnyvale, CA), used in laparoscopic procedures [7], and autonomous robots such as the ROBODOC® (Integrated Surgical Systems; Davis, CA) used to perform total hip replacements [8 – 9]. In otologic surgery, to date, robots have been used only for surface milling of the temporal bone (to a depth of 4.5 mm) to create a receiving well for the internal processor of a cochlear implant device [10]. One of the barriers preventing further integration of robots into otologic surgery is that the surgical skills used in these types of procedures are not well quantified.

Recent efforts to evaluate surgical skill in other arenas have been successful. One such example is in the area of high-level surgical modeling where there is evidence that statistical models derived from recorded force and motion data can be used to classify surgical skill level (novice or expert) with a classification accuracy approaching 90% [11 – 12]. However,

these works did not recognize the transition boundaries between the different surgical actions and identify these actions that were used to perform the surgery. Detection of transition of surgical action and identification of these actions involved in a surgery is still in its infancy and remains an active research area.

The objective of this paper is to model a mastoidectomy procedure as a compilation of basic surgical actions, which we call action primitives (APs). We are motivated by our belief that this in turn could be used to train a robot to carry out similar surgical steps in the future. In this paper, we present a novel technique for detection of transition of the different APs and identification of those APs involved in mastoidectomy based on the real-time force, position and orientation data of the surgical tool—a drill—during the procedure. We present experimental results based on datasets obtained from performing mastoidectomy on temporal bones of cadavers. The results are validated against hand segmentation and labeling of the sequence of operations provided by the surgeons. In Section II, we present the design rationale underlying the mapping of the surgical actions. Section III discusses the materials used and experimental procedure. Here, we also discuss the algorithms used in the detection of boundaries between the distinct surgical activities and subsequent identification. The scope of the results is discussed in Section IV. In Section V, we present an overall discussion. Finally, Section VI summarizes the contributions of this work and describes a scope of future work.

II. DESIGN RATIONALE FOR MAPPING OF SURGICAL PROCEDURE IN MASTOIDECTOMY

The main objective of this research is to detect the transition boundaries between surgical actions as well as identify these actions performed during mastoidectomy. We propose that the continuous data stream can be broken down into Action Primitives (APs) based on the surgical drill's position, orientation, and end-effector forces (i.e., force exerted by the drill tip on the tissue). The APs are defined such that they are distinguishable according to their unique roles in the surgery and, when strung together, form a complete set of actions necessary to perform the surgery. There are other instances in literature when a similar approach is used for pattern recognition based on functional modules (e.g., robot-assisted suturing task [13]).

Our hypothesis is that while variability exists between surgical procedures with different goals (e.g., cochlear implantation, removal of cholesteatoma), and between surgeons (who may have different style), that all surgeons use a finite armamentarium of surgical actions to complete a mastoidectomy procedure. Our goal, then, is to extract unique characteristic features based on the acquired signals for each AP. Based both on a guidebook [14] and extensive observations within the operating room, we have modeled a mastoidectomy procedure as a compilation of three basic surgical actions that we call AP1, AP2 and AP3. It is conceivable that there may be other APs in mastoidectomy; however we begin our investigation with the three basic APs. AP1 is “rough boundary exploration” accomplished by removing the cortical bone and subsequent superficial layers of the trabecular mastoid bone. This action is followed by AP3, which is “fine boundary exploration”, and then by

AP2, which is “obliteration of tissue within the boundaries”. We hypothesize that these actions can then be strung together such that they create a unique set describing the surgical procedure.

In this work, we approach the problem from a data-driven perspective where our goal is to detect the transition boundaries between the three APs defined above and also identify these APs from force, position, and orientation data acquired during the surgery. A continuing challenge in biology today is the need to integrate large quantities of experimental data into quantitative and testable descriptions of system behavior. Due to their size and complexity, the data generated through large-scale interrogations are generally recognized uninterpretable without the use of computational methods for data reduction, analysis and modeling [15]. Data-driven models borrow heavily from Artificial Intelligence (AI) techniques and are based on a limited knowledge of the modeling process, thereby relying on the data describing the input and output characteristics. In contrast to physically-based modeling (or knowledge-driven modeling) that tries to explain the underlying process, the data-driven models are able to make abstractions and generalizations of the process from the data itself and play a complementary role to physically-based models [16].

The first step to analyze the surgical procedure using the data-driven approach is to extract relevant features from the surgical data that can be used to detect the boundary between the APs as well as identify each AP. This step enables characterizing each surgical step. The next step is to deconstruct the surgical procedure into a series of APs that comprise the complete procedure.

As a result, we begin by parsing the surgical procedures into APs by extracting relevant features (characterizing each AP) from the raw surgical data (Section III B.1). Once these features are extracted, we then use them to detect the transition time of one AP to another. The problem of knowing when one AP changes to another (boundary detection) can be cast as a segmentation problem in a continuous stream of data that changes its features based on the surgical actions. To achieve segmentation of the recorded signals from surgical procedures, we use Bayesian Information Criterion (BIC) method to learn the unique characteristics of each AP (Section III B.2). BIC is a model selection criterion that has been widely used in video [17] and speech segmentation [18 – 21], which bears some resemblance to our problem. The idea here is that once we train the BIC algorithm on the selected feature set of each AP, it can automatically segment a continuous data stream (fed to the algorithm) into a sequence of APs. Once the boundaries marking the transition between the APs are detected, the next step is to identify and label each AP (Section III B.3) based on the extracted feature set. We use an Artificial Neural Network (ANN) based approach for this purpose. While there are other pattern recognition tools available, we use ANNs because of their proven ability in other applications requiring signal-based classification (e.g., audio signal classification and tagging of different audio portions [22 – 23]) to detect complex patterns within data sets [24 – 25], handle non-linearities, and provide satisfactory real-time predictive accuracy. The schematic of the overall process is given in Fig. 1.

III. EXPERIMENT, ANALYSIS AND RESULTS

A. Materials and Methods

A.1) Temporal bones—Fifteen cadaveric temporal bone specimens were utilized for this study. These specimens were harvested from formalin-fixed cadavers used in our institution's anatomy lab. No criteria were used to select the specimens for this study. Our intent was that these 15 specimens represent a randomly chosen distribution of anatomy (e.g., well-pneumatized versus poorly-pneumatized specimens). To prevent tissue breakdown, the temporal bones were stored in a freezer. Prior to drilling, each frozen specimen was defrosted to room temperature. As per the standard for temporal bone drilling, each specimen was affixed in a temporal bone holder (Fig. 2) consisting of three pins which were screwed into the bone. These pins were rigidly affixed to a metal bowl (the temporal bone holder) which then sits on a surface.

A.2) Force / torque measurement—The temporal bone holder was affixed on top of a force and torque measurement tool, the F/T DAQ Multi-Axis Force/Torque Sensor System (ATI Industrial Automation, Apex, NC). This sensor system has a resolution of 0.003125N (for force components along the x and y directions), 0.00625 N (for force component along the z direction), and 0.0001875Nm (for torque components along the x , y and z directions). The sensor was interfaced to a laptop using the analogue-digital-interface card “PC-Card DAS 16/330” (Measurement Computing Corp., Norton, MA). A custom MATLAB (version 2008a, MathWorks Inc., Natick, MA) application was developed for recording the force and torque values during the drilling process. The sensor with the temporal bone holder was affixed rigidly to a sturdy table to avoid any movement during data collection. During drilling, care was taken not to move the temporal bone holder. With this rigid body setup, we ensured that the force measured by the sensor was the force exerted by the surgical drill on the temporal bone specimen. An arm rest – not physically connected to the table surface – was used to avoid having the weight of the surgeon's arm confound the data.

A.3) Position and Orientation Data Measurement—For tracking the position and orientation of the tip of the surgical drill during the mastoidectomy, a Polaris Spectra optical tracking system (Northern Digital Inc; Waterloo, Ontario, Canada) with 0.25mm localization accuracy was used. We used an eMax2 High Performance Instrument System (The Anspach Effort, Inc., Palm Beach Gardens, FL) surgical drill unit. Markers recognized by the tracking system were attached to the drill (Fig. 3) to allow us to track the tip of the drill during the procedure. Calibration was performed prior to beginning of the surgery to determine the drill tip location relative to the marker locations. A custom MATLAB program residing on another computer was then used to collect and save the drill-tip locations and orientations.

A.4) Experimental Procedure—In an endeavor to develop a computer-controlled robot to perform a mastoidectomy, it is necessary to understand how the mastoidectomy is performed by surgeons. This will also aid in training novice surgeons first in a cadaveric laboratory and subsequently in the operating room. Two experienced otologic surgeons (a.k.a. faculty level surgeons) were asked to (a) perform each of the three APs separately and (b) perform a cortical mastoidectomy procedure by stringing together a sequence of APs. We

analyzed actions of surgeons during drilling of mastoid in cadaveric specimens (Fig. 4) by recording position, force and orientation data. The data was used to (a) characterize each AP; (b) detect transition boundaries between APs; and (c) identify each AP.

The surgeon touched an anatomical landmark (the Spine of Henle) three times at the start of data acquisition as a zero-time reference point. The data from this step was used to synchronize the data from the force/torque sensor and the tracking system. All procedures were performed under microscopic view, and this view was videotaped for offline analysis. Fig. 5 shows the images of the different stages of surgery for a typical temporal bone.

A.5) Data Analysis and Labeling—The videotape was analyzed using Final Cut Express 4 (Apple, Cupertino, CA) video editing software. Time for different AP steps as well as breaks during the procedure were identified and noted by the operating surgeon after completion of the cortical mastoidectomy. This time data was then used during the analysis of the acquired data.

B. Feature Processing Algorithms Used

B.1) Data Preprocessing, Feature Extraction and Selection of reduced Feature Set Characterizing the Surgical Action Steps—First, the characteristic APs had to be defined. To do this we used the data recorded when surgeons performed a characteristic AP. The sampling frequency for data acquisition was 50 Hz. To eliminate transients associated with each AP, particularly at the edges of the discontinuities, we removed, on average, a window of 200 samples each at the beginning and also at the end of each characteristic AP to ensure analysis over the steady state portion of the raw data segments.

In an attempt to determine the distinguishing characteristics associated with each AP, the sensory data was processed to extract features (Fig. 6). The preliminary feature set consisted of 63 features as detailed in Table 1. Some of these features were noted to be redundant in characterizing the surgical actions. Out of these extracted features, we chose those features that best discriminated the different surgical actions. The goal of feature processing is to remove redundant input features while retaining the information essential for recognizing the actions with high accuracy [13]. Dimensionality reduction of a feature set is a common preprocessing step used in pattern recognition, classification applications, and in compression schemes. Principal component analysis (PCA) is one of the popular methods used for this purpose [26 – 30]. The relative importance of different features can be determined by performing Correlation Coefficient Component Analysis which in turn helps to reduce false detections [31 – 35]. The preliminary set of features was then further analyzed to extract those features that had the maximum discriminating ability. Now that each AP had a characteristic feature set, we undertook analysis of the data set (i.e., complete string of APs comprising a cortical mastoidectomy) to (i) detect boundaries between APs and (ii) identify each AP as discussed below.

B.2) Boundary Detection between Two APs—As discussed in Section II, we used BIC to detect the boundaries between the different APs used in surgery. We will describe briefly how BIC works.

Let us denote $X = x_i \in R^d$, $i = 1, 2, \dots, N$ as the sequence of vectors in which there is at most one segment boundary. We wish to consider if there is a position $b \in (1, N)$ that is the boundary of different processes (APs), say, AP1 and AP3 generating the two segment outputs $x_1 \dots, x_b$ and $x_{b+1} \dots x_N$, respectively. The decision rule to check and locate the boundary is:

$$\Delta BIC_b = \frac{1}{2} (N \log |\Sigma| - b \log |\Sigma_1| - (N - b) \log |\Sigma_2|) - \frac{1}{2} \lambda \left(d + \frac{1}{2} d (d+1) \right) \log N$$

$$\hat{b} = \arg \max_{1 < b < N} \Delta BIC_b \text{ and } BIC_b \geq 0$$

The variable Σ denotes the covariance matrix of the extracted feature vectors. Σ_1 and Σ_2 are the covariance matrices of the features of the first and the second segment. The d and λ represent feature dimension and penalty weight factor, respectively.

In order to detect multiple segmentation boundaries, a moving window is considered that sweeps through the stream. We start with a window size of 50 samples. Then the window size is extended if no boundary exists and a new window is started from the detected boundary as the next window.

The reduced feature set for boundary detection is achieved by applying PCA and Correlation Coefficient Component Analysis. The PCA was performed on the preliminary feature set (63 features, Table 1), which resulted in 6 features (Table 2) as the principal components. Additionally, a Correlation Coefficient Component Analysis was also performed on the preliminary feature set (Table 1) and 5 more features (Table 2) with correlation coefficients less than 0.2 (on a 0 to 1 scale) were chosen. Thus these 11 features (Table 2) comprised the feature set that was used to detect the transition boundary from one AP to another using BIC. Note that the same 11 features were used to detect all transitions (AP1:AP3 and AP3:AP2) analyzed in this work. The output of the BIC was a temporal transition boundary in each case.

B.3) Identification of APs—As discussed in Section II, we used ANN to identify and label the APs involved in surgery. As can be seen from Fig. 7, we trained 3 independent ANNs (Binary ANN for AP1, AP2, and AP3 that provides identification of AP1, AP2, and AP3, respectively).

Each ANN is designed as a 6 layered structure having an input layer, 4 hidden layers (20-20-20-14), and an output layer. The design is based on a pilot dataset that takes into consideration the trade-off between the least mean-square error and the computation complexity. We used Logsigmoidal approximation functions that equip the ANNs with the ability to handle nonlinearities in the datasets. Instead of using a multi-input, multi-output ANN where only one ANN is employed to identify the three APs from a given input set, we used a modular approach where each modular ANN serves to identify one of the three APs so as to develop a greater noise immunity and ability to handle inconsistencies in the feature sets of the three APs. Here, we use three independent modular multi-input single-output ANNs, where each of the three ANNs is dedicated to identify AP1, AP2 or AP3.

The reduced feature set for AP identification consisted of six principal component features along with three force features (Table 3). This feature set was applied for training the ANNs. Detailed analysis (based on pilot datasets) indicated that the 9 features maximize the performance of the ANN in identifying APs. These features were then used as training inputs to the Binary ANNs. In our present work, the feature set that was used to identify each AP using ANN was the same for all the APs. Their combination using the binary ANN also remained the same throughout our analysis. In response to the characteristic features of the different APs, the trained binary ANNs give an output of 1 or 0 (i.e., when it identifies an AP, it gives an output of 1; otherwise it outputs 0). For example, in case the input feature vectors corresponded to AP1 segment of the surgery, then the 'Binary ANN for AP1' gave an output of 1 (Fig. 6), thereby leading to the prediction of AP1, whereas, the 'Binary ANN for AP2' and the 'Binary ANN for AP3' gave an output of 0. Similar outputs were recorded for the identification of AP2 and AP3 of the surgery.

Each of these Binary ANNs was trained off-line using the above-mentioned reduced feature set as input and the AP label as the output. The three trained ANNs were then used to identify different APs using the reduced feature set of the test data. The same input feature vectors from the test sets were applied in parallel to the three binary ANNs (Fig. 7). The prediction of the ANN (i.e., whether the test data set corresponded to AP1 or AP2 or AP3) was compared against surgeons own labeling of the corresponding AP.

IV. RESULTS

We used 15 different test cases and the results were validated against hand segmentation and labeling of the sequence of operations provided by the surgeons. In order to estimate the predictive accuracy, a standard cross-validation approach, leave-one-out (LOO), was used. In this method, we used "remove one data set - train on the rest of the data sets - segment and identify APs" procedure for each data vector in our matrix.

A. Results of Boundary Detection

Our test configuration consisted of continuous stream of raw surgical data for APs in the sequence AP1:AP3:AP2 (i.e., rough boundary exploration-followed by-fine boundary exploration-followed by-oblation of tissue within boundaries). While we are cognizant of the possibility that a surgeon may perform a procedure in a different order than this, we chose a standardized approach for this initial study.

Table 4 presents the performance results of the BIC algorithm in boundary detection. It can be seen that the BIC algorithm was successful in detecting the boundaries between all APs. For AP3:AP2 boundary, we had only 14 test data sets because during drilling of one of the temporal bone specimens, the middle portion of the mastoid was broken during AP1 and AP3 stages, thereby removing any recordable obliteration step (AP2).

B. Results of Identification

Table 5 shows the performance results of ANNs in identifying the APs. It can be seen that the trained ANNs were able to identify all of the AP3 and AP1 test cases. The ANNs could also identify 12 out of 14 test cases for AP2. Thus, the trained ANNs were able to detect the

rough boundary and fine boundary exploration actions of the surgery very accurately while missing only 2 cases for the obliteration of tissue action.

C. Analysis of Results

As described above, the BIC method could detect all boundaries with reasonable time accuracy. However, it gave 3 false positives. These were: AP2 of Dataset 10, AP2 of Dataset 7, and AP1 of Dataset 8. On the other hand, the ANN method identified all AP1s and AP3s but missed 2 AP2s. The missed AP2s were: AP2 of Dataset 10 and AP2 of Dataset 7. We also noted that even though ANN correctly identified AP1 of Dataset 8, the identification was marginal. Thus, it is clear that the same segments of data were problematic for both of these algorithms. Upon further evaluation of the data, we found that several features were beyond 1-standard deviation away in all these datasets as shown in Table 6.

From Tables 2 and 3, we find that the reduced feature set comprising the six principal components is employed by both the boundary detection and the AP identification algorithms. These features are the mean, variance and maximum of the third and the fourth quaternion. To understand the physical significance of the orientation components that contribute to the principal distinguishing characteristics of the different basic actions (APs) involved in the surgery, we studied the video captured during the process (as mentioned in Section III A.5). During AP1, the surgical action involves drilling of the temporal bone specimen in order to explore the rough boundary. Then during AP3, the action comprises some drilling accompanied with mainly milling of the mastoid to execute the fine boundary exploration step. Finally, during AP2 stage of the surgery, the action mainly involves milling to obliterate the tissue within the boundaries. A careful examination of the recorded video revealed that the orientation angle of the drill tool with respect to the temporal bone surface varied during performing of the different APs by the surgeon. This variation accounts for different drill tool orientation angles on a three dimensional coordinate system, thereby indicating a justification of the orientation components to emerge as the principal distinguishing features characterizing the different APs. Now, considering the three test cases (as mentioned in Table 6), we studied the drill tool orientation patterns while conducting the different APs. The orientation data plots indicated anomalies during the AP2 action step for Test Case 7, AP1 action step for Test Case 8, and AP2 step for Test Case 10.

We therefore believe that these datasets were significantly different from the other representative datasets. The cause for this anomaly is unknown at this time. We will explore more datasets in the future to further characterize the APs. Even though we had 3 missing cases, we still believe that the above results show the success of the approach in automatically finding the boundaries and identifying APs during mastoidectomy.

V. DISCUSSION

In this study we sought to decompose a complex surgical procedure – a mastoidectomy – into a string of well characterized surgical actions (APs). To do this we (a) characterized surgical actions by extracting the features characteristic of those finite actions (AP1, AP2, AP3), and then (b) had surgeons perform a complete procedure (a cortical mastoidectomy)

and used the discriminating feature set characterizing each AP to (i) detect boundaries between fundamental surgical actions (APs) and (ii) identify each action (AP).

The surgeons performed mastoidectomy surgery on 15 cadaveric temporal bones (Section III A.1) and force, position, and orientation data were collected during the procedure. The surgeons also hand labeled the different action sequences (AP1, AP3 and AP2) after the surgery (Section III A.5). Then we extracted a preliminary set of 63 features characterizing each AP. This was followed by extraction of the reduced feature set (11 for detecting transition boundaries between the APs as shown in Table 2, and 9 for identifying each AP as shown in Table 3) possessing maximum discriminating ability. This dimensionality reduction was facilitated by performing PCA and Correlation Coefficient Component Analysis. Regarding boundary detection, the BIC method could detect all boundaries with reasonable time accuracy (within approximately a couple of seconds). Our preliminary investigation shows that, the pertinent features for boundary detection (Table 2) did not include force. While excluding force may seem counter intuitive, we realized during video analysis that the surgeon makes orientation changes before changing force. Regarding AP identification, we used three independent modular multi-input and single-output binary ANNs. The discriminating features (Table 3) when applied as inputs to the ANNs were able to identify all AP1s and AP3s, while missing 2 cases of AP2s.

Building on these initial results, we now plan to apply this process to more complex variation of a mastoidectomy (e.g. labyrinthectomy). We will start with data acquired in the temporal bone lab and move to data acquired in the operating room using a similar set-up to that shown in Fig. 2, where the force transducer is the sole support of the patient's head thus collecting pertinent data. This will allow us to decompose a human operator's (surgeon's) actions into a list of APs. Once we accomplish this decomposition, we will implement human strategies into our robotic setup which consists of a programmable Mitsubishi Industrial Robot RV-3S (Mitsubishi Electric Corporation, Tokyo, Japan). Initial robotic strategies will be based on our results that orientation of the surgical instrument best delineates boundaries between the APs and identifies each AP. We will then investigate further using force feedback to fine-tune rough surgical actions to emulate human actions.

VI. CONCLUSION

This paper discusses the feasibility of a novel approach which entails feature extraction followed by boundary detection and identification of different APs using a trained BIC algorithm and ANNs. Initial results based on experiments with cadaveric temporal bones are promising. Further extensive investigation will potentially enable us to capture the surgical strategy (as characterized by the APs) and the surgeon's skills (as characterized by the features) that could be emulated by a robot in the future. We hope that the proposed approach will open up the possibility of establishing a standard operating strategy that a robot can use to perform the surgery.

ACKNOWLEDGMENT

We thank Stephan Baron, Jason Mitchell and Kevin Fite for their inputs to this project. We also express our thanks to National Institute of Biomedical Imaging and Bioengineering for funding our research.

This work was supported by the grant R21 EB006044-01A1 from the National Institute of Biomedical Imaging and Bioengineering.

REFERENCES

- [1]. French LC, Dietrich MS, Labadie RF. An estimate of the number of mastoidectomy procedures performed annually in the United States. *Ear Nose Throat J.* May; 2008 87(5):267–270. [PubMed: 18572781]
- [2]. Green JD, Shelton C, Brackmann DE. Surgical Management of Iatrogenic Facial Nerve Injuries. *Otolaryngol Head Neck Surg.* Nov; 1994 111(5):606–610. [PubMed: 7970799]
- [3]. Kerr JT, Chu WKF, Bayles SW. Cerebrospinal fluid rhinorrhea: Diagnosis and Management. *Otolaryngol Clin North Am.* Aug; 2005 38(4):597–611. [PubMed: 16005720]
- [4]. Schipper J, Aschendorff A, Arapakis I, Klensner T, Teszler CB, Ridder GJ, Laszig R. Navigation as a quality management tool in cochlear implant surgery. *The Journal of Laryngology & Otology.* Oct.2004 118:764–770. [PubMed: 15550181]
- [5]. Cobb JC, Davies BL, Harris SJ, Hibberd RD, Lin WJ, Middleton R. Active Compliance in Robotic Surgery—the Use of Force Control as a Dynamic Constraint. *Proc Inst Mech Eng (H).* 1997; 211:85–92.
- [6]. Jakopcic M, Rodriquezy BF, Harris SJ, Gomes P, Cobb J, Davies BL. The Hands-On Orthopaedic Robot “Acrobot”: Early Clinical Trials of Total Knee Replacement Surgery. *IEEE Trans Robot Autom.* 2003; 19:902–911.
- [7]. Guthart, GS.; Salisbury, JK. The Intuitive™ Telesurgery System: Overview and Application. *Proc IEEE Conf Robot Autom;* 2000. p. 618-621.
- [8]. Paul, HA.; Mittlestadt, B.; Bargar, WL.; Musits, B.; Taylor, RH.; Kazanzides, P.; Zuhars, J.; Williamson, B.; Hanson, W. A Surgical Robot for Total Hip Replacement Surgery. *Proc IEEE Conf Robot Autom;* 1992. p. 606-611.
- [9]. Honl M, Dierk O, Gauck C, Carrero V, Lampe F, Dries S, Quante M, Schwieger K, Hille E, Morlock MM. Comparison of Robotic-Assisted and Manual Implantation of a Primary Total Hip Replacement. *J Bone Joint Surg Am.* 2003; 85-A:1470–1478. [PubMed: 12925626]
- [10]. Federspil PA, Geisthoff UW, Henrich D, Plinkert PK. Development of the First Force-Controlled Robot for Otoneurosurgery. *Laryngoscope.* Mar; 2003 113(3):465–471. [PubMed: 12616198]
- [11]. Rosen J, Hannaford B, Richards CG, Sinanan MN. Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. *IEEE Trans Biomed Eng.* 2001; 48(5):579–591. [PubMed: 11341532]
- [12]. Richards C, Rosen J, Hannaford B, Pellegrini C, Sinanan M. Skills evaluation in minimally invasive surgery using force/torque signatures. *Surgical Endoscopy.* 2000; 14:791–798. [PubMed: 11000356]
- [13]. Lin, HC.; Shafran, I.; Murphy, TE.; Okamura, AM.; Yuh, DD.; Hager, GD. Automatic Detection and Segmentation of Robot-Assisted Surgical Motions. *Med Image Comput Comput Assist Interv Int Conf;* 2005. p. 802-810.
- [14]. Nelson, RA. Temporal Bone Surgical Dissection Manual. 1982.
- [15]. Yang W, Johnson GL, Gomez SM. Data-driven modeling of cellular stimulation, signaling and output response in RAW 264.7 cells. *Journal of Molecular Signaling.* 2008; 3:11. [PubMed: 18498628]
- [16]. Solomatine, DP. Data-driven modeling : paradigm, methods, experiences. *Proc. 5th Int. Conference on Hydroinformatics;* Jul. 2002 p. 757-763.
- [17]. Baillie, M.; Jose, JM. An Audio-based sports video segmentation and event detection algorithm. *Conference on Computer Vision and Pattern Recognition Workshop;* Jul. 2004 p. 110
- [18]. Tritschler, A.; Gopinath, R. Improved speaker segmentation and segments clustering using the Bayesian Information Criterion. *Proceedings of the 6th European Conference on Speech Communication and Technology;* Sep. 1999 p. 679-682.
- [19]. Zhou, B.; Hansen, J. Unsupervised Audio Stream Segmentation and Clustering via the Bayesian Information Criterion. *Inter. Conf. Spoken Language Processing;* Oct. 2000 p. 714-717.

- [20]. Akbacak M, Hansen JHL. Environmental Sniffing: Noise Knowledge Estimation for Robust Speech Systems. *IEEE Transactions on Audio, Speech and Language Processing*. Feb; 2007 15(2):465–477.
- [21]. Chen S, Gopalakrishnan P. Speaker, Environment and Channel Change Detection and Clustering via the Bayesian Information Criterion. *Proc. Broadcast News Trans.* Feb.1998 6:127–132.
- [22]. Scheirer, E.; Slaney, M. Construction and Evaluation of a Robust Multifeature Speech/Music Discriminator. *Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*; Apr. 1997 p. 1331-1334.
- [23]. Meinedo, H.; Neto, J. A stream-based audio segmentation, classification and clustering pre-processing system for broadcast news using ANN models. *Proceedings of the 9th European Conference on Speech Communication and Technology, (INTERSPEECH '05)*; Sep. 2005 p. 237-240.
- [24]. Lin CS, Chiu JS, Hsieh MH, Mok MS, Li YC, Chiu HW. Predicting hypotensive episodes during spinal anesthesia with the application of artificial neural networks. *Comput Methods Programs Biomed.* Aug.2008
- [25]. Goles E, Palacios AG. Dynamical Complexity in Cognitive Neural Networks. *Biol Res.* May; 2008 40(4):479–485. [PubMed: 18575680]
- [26]. Alzate, C.; Suykens, JAK. Image Segmentation using a Weighted Kernel PCA Approach to Spectral Clustering. *Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing*; Apr. 2007 p. 208-213.
- [27]. Qian D, Chang C, Chein I. Segmented PCA-based compression for hyperspectral image analysis. *Proceedings of the SPIE*. 2004; 5268:274–281.
- [28]. Cohen, I.; Tian, Q.; Zhou, XS.; Huang, HT. Feature Selection Using Principal Feature Analysis. *Proceedings of the fifteenth International Conference on Information Processing (ICIP '02)*; 2002.
- [29]. Pohl, KM.; Warfield, SK.; Kikinis, R.; Grimson, WEL.; Wells, WM. Coupling Statistical Segmentation and PCA Shape Modeling. *Proc. MICCAI 2004: Seventh International Conference on Medical Image Computing and Computer Assisted Intervention, Rennes / St-Malo; France: Springer-Verlag; 2004. p. 151-159.vol. 3216 of Lecture Notes in Computer Science*
- [30]. Drugman, T.; Gurban, M.; Thiran, JP. Relevant Feature Selection for Audio-Visual Speech Recognition. *IEEE 9th International Workshop on Multimedia Signal Processing (MMSP)*; Oct. 2007 p. 179-182.
- [31]. Wang D, Terman D. Image Segmentation Based on Oscillatory Correlation. *Neural Computation.* May; 1997 9(4):805–836. [PubMed: 9161023]
- [32]. Aggarwal, N.; Prakash, N.; Sofat, S.; Mittal, A. Temporal Video Segmentation using Cross Correlation. *Visual Information Engineering, IET International Conference*; Sep. 2006 p. 7-11.
- [33]. Passonneau, RJ.; Litman, DJ. Intention-Based Segmentation: Human Reliability and Correlation with Linguistic Cues. *Proceedings of the 31st annual meeting on Association for Computational Linguistics*; 1993. p. 148-155.
- [34]. Volker R, Tilman L. Adaptive Feature Selection in Image Segmentation. *Pattern Recognition.* 2004; 3175:9–17.
- [35]. Roterman, Y.; Porat, M. Progressive Image Coding using Regional Color Correlation. *4th EURASIP Conference*; Jul. 2003 p. 65-70.
- [36]. Kimberley BP, Fromovich O. Flexible Approach to Tympanomastoidectomy. *Otololaryngologic Clinics of North America.* Jun.1999 32:585–595.

Figure 1.
Broad overview of the rationale used

Figure 2. Temporal bone specimen fixed in the temporal bone holder mounted on a force/torque sensor system.

Figure 3.
Markers attached to the drill for tracking the drill tip position during drilling.

Figure 4.
A surgeon performing Mastoidectomy

Figure 5.
Images of different stages of Surgery for a typical Temporal Bone

Figure 6.
 Functional Block Schematic of the Boundary Detection and AP Identification

Figure 7.
Real-time Prediction of Neural Networks (trained off-line)

TABLE 1**PRELIMINARY FEATURE SET**

Feature Description	No. of Features	Cumulative no. of Features	Feature measured by
Mean, variance and maximum of force vector along the x, y, and z directions	3×3	9	Force vectors by F/T SS
Mean, variance and maximum of moment vector along the x, y, and z directions	3×3	18	Moment vectors by F/T SS
Mean, variance and maximum of position along the x, y, and z directions	3×3	27	Position vectors by PSOTS
Mean, variance and maximum of speed along the x, y, and z directions	3×3	36	Derived from position vectors
Mean, variance and maximum of acceleration along the x, y, and z directions	3×3	45	Derived from position vectors
Mean frequency of force vector along the x, y, and z directions	1×3	48	Derived from force vectors
Mean of displacement along the x, y, and z directions	1×3	51	Derived from position vectors
Mean, variance, and maximum of orientation (quaternions)	3×4	63	Orientation vectors by PSOTS

Note:

F/T SS - Force/Torque Sensor System (Detailed in Section III A.2).

PSOTS - Polaris Spectra Optical Tracking system (Detailed in Section III A.3).

TABLE 2**REDUCED FEATURE SET FOR BOUNDARY DETECTION**

Component Description	Feature Description (Input feature set to BIC)	No. of Features	Cumulative no. of Features
Principal Component	Mean, variance and maximum of orientation (3 rd and 4 th quaternion)	3×2	6
Correlation Coefficient	Variance of position, speed and acceleration along x direction	3×1	9
Component	Mean of position along y direction	1×1	10
	Mean of orientation (2 nd quaternion)	1×1	11

TABLE 3**REDUCED FEATURE SET FOR AP IDENTIFICATION**

Component Description	Feature Description (Input feature set to ANNs)	No. of Features	Cumulative no. of Features
Principal Component	Mean, variance and maximum of orientation (3 rd and 4 th quaternion)	3×2	6
Other features	Maximum of force along the x, y, and z directions	1×3	9

TABLE 4**PERFORMANCE OF BIC IN BOUNDARY DETECTION**

Hand-labeled boundary by surgeon	Test Data Sets (no.)	Test Cases in which boundary is detected (no.)	Test Cases in which boundary is not detected (no.)	Boundary Detection Mean Error (sec.)	Test Cases in which False Positives are detected (no.)	Test Cases in which False Negatives are detected (no.)
AP1:AP3	15	15	0	1.71	1	0
AP3:AP2	14	14	0	2.39	2	0

TABLE 5

PERFORMANCE OF TRAINED ANNs IN IDENTIFYING THE APs

Hand-labeled AP (segment)	Test Data Sets (no.)	Test Cases Identified (no.)	Test Cases not Identified (no.)	Test Cases Misidentified
AP1	15	15	0	0
AP3	15	15	0	0
AP2	14	12	2	0

TABLE 6**TEST CASES INDICATING ANOMALY**

Data Set	No. of Features outside 1 Standard Deviation	
	Using BIC Algorithm	Using ANN Algorithm
Test Case 10_AP2	4	6
Test Case 7_AP2	6	6
Test Case 8_AP1	4	4