内視鏡下鼻内手術トレーニングシステムによる手術スキル評価手法

Surgical Skill Evaluation in Endoscopic Sinus Surgery Training System

山内 康司(正会員)^{a*}, 山下 樹里^a, 森川 治^a, 橋本 亮一^a, 持丸 正明^a,
福井 幸男^{ab}, 宇野 廣^c, 横山 和則^d

^a産業技術総合研究所, ^b筑波大学電子・情報工学系, ^c(株)高研, ^d耳鼻咽喉科@南大通り

Yasushi Yamauchi^a, Juli Yamashita^a, Osamu Morikawa^a, Ryoichi Hashimoto^a, Masaaki Mochimaru^a,
Yukio Fuku^{ab}, Hiroshi Uno^c, Kazunori Yokoyama^d

^aNational Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

^bInstitute of Information Sciences and Electronics, University of Tsukuba, Tsukuba, Japan

^cKoken Ltd., Tokyo, Japan

^dENT Clinic@Tsukuba South Avenue, Tsukuba, Japan

Abstract

In most of existing surgical training systems, task completion time and error ratio are common metrics for surgical skill evaluation. To avoid unnecessary and injurious force to the tissue, surgeons must know for themselves how much force they give in handling surgical tools. Our goal is to develop an endoscopic sinus surgery (ESS) training system that quantitatively evaluates the trainee's surgical skills. In this paper, we present an ESS training system with force sensors for surgical skill evaluation. Our experiment revealed that the integral of the force data can also be one of the useful metrics for surgical skill evaluation.

Key words

Surgical training, Surgical skill, Endoscopic sinus surgery.

1. 序論

内視鏡下手術は腹部外科をはじめとして各科に急速に普及しているが、鼻科においても、これまでの侵襲の高い鼻外手術に対して内視鏡を用いた鼻内手術(一般にEndoscopic Sinus Surgery, ESS)が浸透しつつある。内視鏡下手術は、患者に対しては低侵襲と入院期間の短縮というメリットがある一方で、外科医に対しては狭い内視鏡の視野内での正確な作業

を要求する。そのため外科医の熟練度が低い場合やアプローチが難しい患部の場合には、患部の処置が不十分となったり、患部以外の正常な組織を誤って傷つけることによる合併症が問題となる。このような問題を機器側で解決するために、内視鏡下手術を支援する新しい手術器具や手術ロボット、ナビゲーションシステムなどが開発されており、その有効性も示されてきた。しかしながらどのような高度な機器を用いようと、医師が機器を操作する限り、治療成績は医師の知識と技術力にかかってくることは言うまでもない。これまでの内視鏡下手術トレーニングでは模型や動物、献体が用いられてきたが、近年では適切な実験動物の確保も難しく、また献体による練習も機会が限られている。従って臨床で実際の患者に対応しながら熟練医から手ほどきを受けるの

*産業技術総合研究所 人間福祉医工学研究部門
〒305-8566 茨城県つくば市東1-1-1 中央第6

受付2002年3月27日; 採択2002年6月4日

が実態であるが、習熟に時間もかかり、未熟練医による不適切な処置も大きな問題となる。このため実際の患者以外の手段で手術に習熟するための手術トレーニングシステムの開発が行なわれている。

手術トレーニングシステムでは、手術スキルの定量的かつ客観的な評価方法が求められる。手術トレーニングシステムには、その対象とする臓器や手術操作によってさまざまなシステムが開発されているが、近年のいわゆるバーチャルリアリティ技術を応用した手術トレーニングシステムのほとんどは¹⁾³⁾臓器の触覚や手術の過程を「模倣する」ことに主眼が置かれており、手術スキル評価については述べられていない。これら手術「シミュレーション」システムは、訓練者の技術の進捗状況を把握するための定量的なデータを提供するとはいえず、手術「トレーニング」システムとは分けて考えられるべきである。手術スキル評価で最も用いられる指標は、訓練者に一定のタスクを行なわせ、その遂行時間や成功率を評価するものであり⁴⁾⁵⁾、多くの手術トレーニングシステムでも用いられている⁶⁾¹¹⁾。これに加えて、O'Toole ら¹²⁾のように手術中の力覚を重要視し、被訓練者の操作力を手術スキルの指標の一つとして取り入れる試みもあり、指標の多様化が始まっている。

われわれは内視鏡下鼻内手術の研修医レベルを対象とした、操作力計測を取り入れたトレーニングシステムを開発している。腹腔鏡下手術に比べて鼻腔内の術者の作業空間ははるかに狭く、その訓練においては手術器具と鼻内との過度な干渉を避けつつ、いかにスムーズに器具操作をするかが大きなポイントとなっている。既報の内視鏡下鼻内手術トレーニングシステムには³⁾⁹⁾¹⁰⁾このような干渉の度合いを手術スキルの評価指標の一つとして取り入れているものは無い。そこで本研究では模型を用いた内視鏡下鼻内手術トレーニングシステムを開発するだけでなく、被訓練者の手術スキルの評価指標として、タスク遂行時間とともに手術操作力が指標として成立するかを示すとともに、各評価指標の特徴について考察する。

2. 内視鏡下鼻内手術トレーニングシステム

われわれの開発した内視鏡下鼻内手術トレーニングシステムを **Fig. 1** に示す。本システムは頭部ダミー、圧力センサ、内視鏡装置、位置センサ、ビデオモニタリングシステムおよびコンピュータシステムよりなる。

Fig. 1 The ESS training system consists of a dummy of the head, force sensors, endoscope, position sensor, video monitoring system and computer systems.

Fig. 2 Force sensors are fixed beneath the platform of dummy (red and blue circle). These sensors detect the forces perpendicular to the desk.

2.1 頭部ダミー

ダミーは頭蓋骨標本から正確に型を取ったウレタン樹脂によりできている。ダミーは頭蓋骨モジュールおよび左右の鼻腔モジュールからなり、将来的には複数の鼻腔モジュールを交換して使用することが可能となっている。鼻腔モジュールを実際の鼻内に似せるために、軟組織相当の部分をシリコンで肉付けし、さらに着色を施した。本ダミーのCT撮影を行ったところ、臨床相当の撮影条件下で顕著なアーチファクトは見出されず、平均的なCT値は250HUであり、空気と明確に分別することができた。このことから本ダミーは画像誘導鼻内手術のトレーニングにも応用できることが分かる。

2.2 圧力センサ

手術操作力を計測する従来研究のほとんどでは、力センサやトルクセンサを手術器具側(多くの場合鉗子)に設置しているが⁸⁾¹²⁾¹⁴⁾⁻¹⁸⁾、本研究では手術器具とダミーの干渉によって生ずる力を大局的にとらえることを主眼としていることから、ダミーを固定する台の下に左右に二つの薄型圧力センサ(共和電業製PS-5KA)を設置した(Fig. 2)。圧力センサの出力は5Hzでデジタイズされ、コンピュータ(日本電気PC-9801RX)で逐次記録し、ディスプレイ上にグラフとしてリアルタイムに表示される。較正実験では、圧力センサの出力値75unitが100gfに相当することが確認されている。

操作力が全くかからない状態でセンサの出力が双方0になるように調整してあるため、両センサの出力が正であれば、机に対して垂直な成分の力がか

Fig. 3 A scene of training from VCR. The endoscopic view (lower right) is also shown for subject. The force output is shown on the computer display (upper right).

かっていることとなる。一方のセンサの出力が正でもう一方が負であれば、主に左右にねじるようなモーメントがかかっていることになる。

2.3 内視鏡、位置センサおよびビデオシステム

内視鏡システムは30度斜視の硬性鏡(新興光器製作所製、直径4mm)、CCD撮像装置および光源からなる。内視鏡画像は14インチのCRTモニタ(ソニー製)上に表示される。内視鏡先端の位置からタスク遂行時間を決定するために、内視鏡のCCDユニットに光学式位置センサ(Hybrid Polaris, Northern Digital Inc.)のプロープを備え付けた。

ダミー、内視鏡およびモニタ類は、一つのビデオカメラの視野内に収まるように配置され、操作の様子はビデオカメラにて録画される(Fig. 3)。

3. 実験

実験では被験者に、「評価タスク」および「練習タスク」の二つのタスクを交互に行なわせた。評価タスクはダミーの左鼻腔内で行なわせた。練習タスクは評価タスクと類似しているが、右鼻腔内で行なわせた。練習タスクは本システムによる「練習過程」と位置付けて、被験者のパフォーマンスは各評価タスクにおいて評価することとした。同じタスクを長時間繰り返すと、そのタスクのみに特化した運動学習が行なわれるが、左右の鼻腔形状は異なっているほか、後述のターゲット位置も両タスクでは異なっており、内視鏡の導入方法や鉗子の把持方法が同一ではないため、評価タスクでは練習タスクにおける運動学習の影響はほとんど受けないと考えられる。

3.1 タスク

評価タスクは、左鼻腔内の前頭洞孔をターゲットとし、30度斜視鏡を用いて確認し、その次に鉗子の先端をターゲットに接触させるというものである。前頭洞孔は鼻腔内の上方に位置するため、前方の被写体しか観察できない直視鏡では観察が難しい。斜視鏡では通常上方を観察する状態で挿入されるため、前頭洞孔は斜視鏡を回転させることなく、鼻腔内に挿入したそのままの状態を観察可能である。内視鏡を右手に持ちターゲットを視野内に捉えたと、内視鏡は一旦鼻腔内から出して左手に持ち直し、右手に鉗子を持ちそれを若干内視鏡に先行する形で挿入し、内視鏡で位置を確認しながら鉗子の先端をターゲットに接触させる。以上の過程は各評価タスク中1回行なわれる。

練習タスクは、評価タスクとはターゲット位置が異なる以外はおおむね同様のタスクである。練習タスクでのターゲットは右鼻腔内の上顎洞孔である。斜視鏡挿入時に上方を観察する状態で挿入するのは評価タスクと同様であるが、上顎洞孔は側方に位置するため、しかるべき位置で内視鏡の鏡筒を右側に90度回転させる必要がある。ターゲットを内視鏡の視野内に捉えた後、内視鏡は回転させたままの状態ですぐ鼻腔内から出し、鉗子に持ち替えて評価タスクと同様に鉗子の先端をターゲットに接触させる。以上の過程を各練習タスク中、5分間繰り返す。

3.2 被験者

被験者は内視鏡操作経験の無い3人の非医療従事者と、2人の研修医である。研修医はいずれも内視鏡下鼻内手術の経験を有するが、実験当時の臨床経験数は20例以下であった。いずれの被験者も、過去に本システムを用いた経験はない。比較のため、内視

鏡下鼻内手術を200例以上経験した熟練医にも評価タスクを行なわせた。

手術タスクおよび練習タスクの間、圧力センサの出力(5Hz)および内視鏡先端位置(10Hz)を記録するとともに、内視鏡画像および被験者の操作風景はビデオカメラにより同時に録画された。タスク遂行時間は、内視鏡の先端が鼻腔内に存在する時間として定義され、位置センサによって計測された。

3.3 実験過程

実験に先立ち、被験者に対して実験目的の説明を行ない、実験の同意を得た上で、鼻腔内の解剖学的構造の学習を行った。ビデオ学習を用いて内視鏡と鉗子の操作方法の基礎を学習した後、実際に短時間操作練習を行なわせた。引き続いて以下のように5タスクを行なわせた。各被験者の全実験時間はおよそ1時間であった。

- 1) 評価タスク(第1回): ビデオを用いて熟練医から評価タスクの解説を受けた後に評価タスクを行なう。タスク遂行中は位置と圧力データを取得する。
- 2) 練習タスク(第1回): 続いて、ビデオを用いて熟練医から練習タスクの解説を受け、後に練習タスクを5分間繰り返して行なう。タスク終了後は、被験者自身の練習タスクの操作の模様を記録した録画ビデオを見せ、続いて熟練者による解説ビデオを見せる。これは、被験者に対するフィードバック学習として行なわせるものである。
- 3) 評価タスク(第2回): 第1回と同様。
- 4) 練習タスク(第2回): 第1回と同様。
- 5) 評価タスク(第3回): 第1回と同様。

Fig. 4 Time evolutions of the output of force sensors of a resident before the training tasks (a), after the training tasks (b), and of an expert (c). The first evaluation task needed over two minutes and excessive force occurred. After the training tasks (b), task completion time was shorter and the force was lower than before.

Fig. 5 Force indices for each evaluation tasks: a) maximum value, b) average value and c) integration of the absolute values from two force sensors. V1, V2 and V3 are non-medical volunteers, and M1 and M2 are residents. All data are normalized by that of the expert.

4. 結果

Fig. 4 に、研修医の被験者の一人におけるa)第1回評価タスク、b)第3回評価タスク、およびc)熟練者が評価タスクを行った際の圧力センサの出力値を示す。出力の単位はA/Dコンバータによりデジタル化された出力電圧である。

練習前、すなわち第1回評価タスクでは、**Fig. 4a**に示すように大きく非周期的な操作力が観察される。第2章で示したとおり、これは被験者が模型に対して100unit以上の大きな「押し」あるいは「ひねる」力を加えていたことを示す。一方で熟練者では(**Fig. 4c**)このような突発的な大きな操作力はほとんど観察されない。

これに対して練習後、すなわち第3回評価タスク(**Fig. 4b**)では、最初的评价タスク時よりも明らかに

操作力は小さく、操作にかかる時間も短縮されており、圧力センサの出力が練習効果を反映することが示唆される。

次に、この圧力センサの出力を手術スキルの指標とするには、何らかのスカラ値のインデックスの導入が必要となる。そこでわれわれは両センサの出力の絶対値に対して、最大値・平均値・積分値の3つの値を算出し、インデックスとしての妥当性を検討することとした(**Fig. 5**)。すべての値は熟練医の結果の値によって正規化することとした。臨床的な意味付けとしては、最大値は瞬間的に加えた力を、平均値はタスク中に臓器に平均的に与えつづけた力を、積分値はタスク中に臓器に与えた力の総和を意味する。**Fig. 5**に示すとおり、全ての値に関して熟練者のそれが最も低く(1.0)、熟練者が臓器に対して全体的に不必要な力を加えていないことが明らかである。

Fig. 6では、タスク遂行時間について熟練者の値で正規化した値を示す。操作力も遂行時間も、練習タスクを繰り返すことで値が低下するように見えるが、例外も見受けられる。

Fig. 6 Completion time of each evaluation tasks. Times are normalized by that of the expert (24.6 s).

5. 考察

力の3つのインデックスのうち、積分値は練習を重ねることによって顕著に減少している(ノンパラメトリック: Friedmanテスト($p < 0.05$)およびWilcoxon's coding rankテスト($p < 0.05$)). 平均値と最大値に関しても若干減少傾向が見られるが、積分値と異なり統計的な有意差が見られなかった。積分値のディメンジョンは平均値や最大値と異なり、時間と力の積であり、必然的にタスク遂行時間と操作

力平均の両方の傾向を兼ね備えているといえる。また臨床的にも積分値は操作力による臓器への影響の総和を意味することから、内視鏡下鼻内手術の手術スキルを評価するにあたって操作力の積分値は極めて有効な指標であると考えられる。

Fig. 4 および Fig. 5 を見ると、熟練医と研修医および非医療従事者との相違も明らかである。もちろん、熟練医はどの被験者と比較しても良好な結果を収めている。研修医については、最初の評価タスクでは明らかに手術器具とダミーとの干渉が大きかったが、2回の練習タスクを経る間に操作力の積分値が下がり、最後には熟練医と同じレベルに達している。一方で非医療従事者では、練習の間に積分値は低下するように見えるものの、熟練医と同じレベルに達したとはいえない。被験者数が少ないため参考程度にとどめるべきであるが、研修医と非医療従事者との間では操作力の積分値に著しい差がある(分散分析, $p < 0.05$)。本研究は指標の妥当性を検討するため、短時間の学習中のスキルを観察したものだが、今後学習プログラムも含めた手術トレーニングシステムとしての有効性を示すためには、最終的には数日から数か月にわたる長期のスキル評価も必要である。

本研究の実験では、練習タスク中の被験者の練習風景を撮影したビデオ(被験者の操作の様子、内視鏡画像、圧力センサからの出力値)を練習タスク終了後に被験者が見ることで、被験者のフィードバック学習としている。本研究の実験では用いられていないが、本システムには圧力センサの出力値のレベルに応じたビーブ音を発する機構が備わっており、練習中のリアルタイムのフィードバック学習として用いることが可能である。本研究では位置センサをタスクの開始と終了を決定するためだけに用いられているが、手術器具の位置もまた手術スキル評価の定量的な指標の一つとなりうるので¹⁹⁾、今後は手術器具の位置を含めたスキル指標も検討する。

操作力の検出方法については、本研究は操作力の指標としての評価を行なう段階であり、2自由度で十分であると判断した。操作力の指標としての有効性が本研究で示されたため、今後は6自由度全ての操作力を計測すれば、力のベクトルを含めたより細かい評価が可能となろう。

また2.2節で示したように、操作力計測に関する従来研究のほとんどでは手術器具側に力センサを設置している。これは切断操作や把持操作といった局所的な手術操作を対象としていたり、模型でなく生体や術中の操作力の計測を目的としているからである

が、当然力センサを設置した特殊な手術器具が必要となる。これに対して、ダミー側にセンサを設置できれば、手術器具を選ばないトレーニングが可能となるため、幅広いトレーニングメニューに対応できる。また、内視鏡下鼻内手術のように術野が狭く手術器具間(例えば内視鏡と鉗子)の衝突が避けられない場合、手術器具側で力を計測したのでは、器具同士との衝突によって生じた反力も検出してしまい、臓器との干渉によって生ずる反力との区別がつかなくなる。従って本研究のように、手術器具を生体になるべく干渉させないように操作することを目的とした手術トレーニングシステムでは、ダミー側への力センサの設置のメリットは大きい。

多くの先行研究ではタスクの成功率とタスク遂行時間が用いられている。成功率は、タスクの成績が「成功」と「失敗」に明確に判定できる事象には向いているが、成績が二元化できない事象(例:操作の迅速さや正確さ)は評価できず、またタスクの成功率が非常に高い、あるいは低すぎるタスクの評価にも向いていない。一方でタスク遂行時間は、手先の器用さを操作の迅速さで定量化するものであり、操作の質の指標として有用であるように見える。しかしながらタスク遂行時間のみを指標にすると、練習者はタスク完了を優先するあまり操作の丁寧さを犠牲にする恐れがある。従ってタスク遂行時間は、実際の手術でそのタスクに迅速性が最優先される場合や、手術器具と周辺臓器との干渉がさほど生じない場合(例:腹腔鏡操作)を除けば、補完的な指標とするのが適切である。力については、本研究のように手術器具の挿入・移動時に臓器との干渉を防ぐ必要のあるタスクには好適であり、鼻腔のみならず消化管や呼吸器・循環器系といった内腔が狭い臓器内の操作の評価に有用である。またマイクロサージェリやロボット手術のように、手術器具からの力覚フィードバックが直接得られないタスクにも有用な評価指標となる。例えば鉗子の把持力や縫合糸の張力といった、ある適正な範囲の力が必要となるタスクで、適正值からの誤差といった指標が用いられることとなる¹⁹⁾。手術器具の位置情報については、位置決め精度が求められる手術タスクの評価だけでなく、運動軌跡の解析によるタスクのスムーズさの指標としても期待できる²⁰⁾。

このように、手術スキルの評価指標に定石はなく、評価対象タスクによって臨床的・人間工学的に意味のある指標選択が求められる。

6. 結論

内視鏡手術のトレーニングを促進するため、われわれは操作力を検出する圧力センサを有した鼻腔ダミーによるトレーニングシステムを開発し、基礎的な手術操作を元にしたトレーニングプロトコルによる操作実験により、練習による手術スキルの向上度の指標化について検討した。その結果圧力センサの出力値の絶対値の積分値が、手術操作力の評価の有力な評価指標であることが示された。

文献

- 1) Kaufmann C, Zakaluzny S, Liu A. First Steps in Eliminating the Need for Animals and Cadavers in Advanced Trauma Life Support. Lecture Notes in Computer Science 2000;1935:618-623.
- 2) Gomes MPSF, Davies BL. Computer-Assisted TURP Training and Monitoring. Lecture Notes in Computer Science 2000;1935:669-677.
- 3) Edmond JCV, Heskamp D, et al. ENT Endoscopic Surgical Training Simulator. Studies in Health Technology and Informatics 1997;39:518-528.
- 4) Rosser Jr JC, Rosser LE, Savalgi RS. Objective evaluation of a laparoscopic surgical skill program for residents and senior surgeons. Arch Surg 1998;133:657-661.
- 5) Hanna GB, Shimi SM, Cuschieri A. Task performance in endoscopic surgery is influenced by location of the image display. Ann Surg 1998;227:481-484.
- 6) Jordan JA, Gallagher AG, McGuigan J, McClade K, McClure N. A comparison between randomly alternating imaging, normal laparoscopic imaging, and virtual reality training in laparoscopic psychomotor skill acquisition. Am J Surg 2000;180:208-211.
- 7) Gallagher AG, McClure N, McGuigan J, Crothers I, Browning J. Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST VR). Endoscopy 1999;31:310-313.
- 8) Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. Am J Surg 1998;175:482-487.
- 9) Rudman DT. Functional endoscopic sinus surgery training simulator. Laryngoscope 1998;108:1643-1647.
- 10) Weghorst S. Validation of the Madigan ESS simulator. Studies in Health Technology and Informatics 1998;50:399-405.
- 11) Yamashita J, Yamauchi Y, et al. Real-Time 3D Model-Based Navigation System for Endoscopic Paranasal Sinus Surgery. IEEE Biomed Eng 1999;46:107-116.
- 12) O'Toole RV, Playter RR, Krummel TM, Blank WC, Cornelius NH, Roberts WR, et al. Measuring and developing suturing technique with a virtual reality surgical simulator. J Am Coll Surg 1999;189:114-127.
- 13) Yamauchi Y, Suzuki M, et al. A Training System for Endoscopic Sinus Surgery with Skill Evaluation. International Congress Series 2001;1230:1150.
- 14) Rosen J, Hannaford B, MacFarlane MP, Sinanan MN. Force controlled and teleoperated endoscopic grasper for minimally invasive surgery--experimental performance evaluation. IEEE Trans Biomed Eng 1999;46:1212-1221.
- 15) Emam TA, Hanna GB, Kimber C, Dunkley P, Cuschieri A. Effect of intracorporeal-extracorporeal instrument length ratio on endoscopic task performance and surgeon movements. Arch Surg 2000;135:62-65.
- 16) Hanna GB, Shimi S, Cuschieri A. Influence of direction of view, target-to-endoscope distance and manipulation angle on endoscopic knot tying. Br J Surg 1997;84:1460-1464.
- 17) Hanna GB, Drew T, Clinch P, Shimi S, Dunkley P, Hau C, et al. Psychomotor skills for endoscopic manipulations: differing abilities between right and left-handed individuals. Ann Surg 1997;225:333-338.
- 18) 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;48:579-591.
- 19) Salcudean SE, Ku S, Bell G. Performance measurement in scaled teleoperation for microsurgery. Lecture Notes in Computer Science 1997;1205:789-798.
- 20) 山内康司, 山下樹里, 持丸正明, 福井幸男, 横山和則. 内視鏡下副鼻腔手術のためのナビゲーションシステムにおける操作性の比較評価. 日本バーチャルリアリティ学会論文誌 1998;3(4):221-227.