

修士論文

機械学習による気管挿管手技の熟練度判定

指導教員 野村由司彦 教授

平成 29 年度

三重大学大学院工学研究科

博士前期課程 機械工学専攻

416M148 森井 秀幸

目次

第1章	序言.....	4
1.1	研究背景.....	4
	既存研究.....	5
第2章	提案手法.....	6
2.1	装置.....	7
2.2	動作データ前処理.....	10
2.2.1	データ計測.....	10
2.2.2	時系列データの切り出し.....	11
2.2.3	窓移動法.....	11
2.2.4	特徴量の検討.....	11
2.3	学習器.....	12
2.3.1	DNN.....	12
2.3.2	SVM.....	14
2.3.3	判定方法.....	14
第3章	実験.....	16
3.1	計測.....	16
3.2	フェーズ分け.....	16
3.3	手技の所要時間.....	17
3.4	実験条件.....	18
3.5	結果.....	19
3.5.1	窓幅,学習器の比較.....	19
3.5.2	フェーズ毎の評価.....	25
第4章	考察.....	26

4.1	窓幅,学習器の比較.....	26
4.2	フェーズ毎の評価.....	26
第5章	結言.....	28
謝辞		29
参考文献		30

第1章 序言

1.1 研究背景

医療手技の気管挿管は全身麻酔時等における気道確保のため必要不可欠な手技であるが、その習熟は熟練医の主観に基づく評価で指導が行われている。この指導の一般化や効率化のためには手技を客観的に評価できる指標が求められる。Brydges らは医療の手技の「結紮」手技について、所要時間や手先の移動距離に基づく評価を研修段階で行い教育を進めていくのが効率的であると提案している[1]。

気管挿管の教育においては、人の頭部を模したマネキンのシミュレータを利用した訓練が実施されている。これによって、初学者は繰り返し実習をおこなえるようになった。こうした装置を利用する場合、まず熟練者がシミュレータ等を利用して実演をしながら、言葉や身振りで気管挿管手技の重要な動作を説明する。しかし、こうした手法では挿管手技のポイントを十分に理解させることができていないことが多い。

葛岡らは、気管挿管手技に用いる喉頭鏡の教示のため、手の外側にかぶせる外装型力覚提示装置を作成し手の外側から力覚を提示することによって、喉頭鏡の動かし方を教示する方式を提案した[2]。また、MICOTO Technology は気管挿管手技のトレーニングができる医療教育用シミュレータロボット mikoto は咽頭後壁等へのセンサ付加による生体反応を再現し、各種センサの反応強度や手技時間によって客観的評価を実現している[3]。

しかし、力覚やセンサの反応強度によって評価する場合、よりよい評価を実現するための身体の動かし方までは教示できない。また、実際の患者を想定する場合気管挿管手技は必ずしも定まった動きではなく、患者との相対的な挙動を測定することが求められ、シミュレータでは困難である。そこで、医師側の全身モーションキャプチャで熟練者

と初学者の動作を比較し,深層学習で熟練度に関連する身体動作を解明し手技を客観的に評価する手法を提案する.

既存研究

医療に関連した手技に限らず楽器の演奏,スポーツ,職人技といった身体的技能の習熟の過程でも同様に評価できる指標が求められる.そのためにはまず身体動作を客観情報化する必要がある.[2]はモーションキャプチャを用いて陶芸に関する身体技能である「菊練り」について頭部,肩,肘,手の位置情報をもとに動作の周期性や各部位の協調性を分析した.[5]は光学式モーションキャプチャによりピアノ演奏時の指の各関節の角度や速度を計測しそれらの情報をもとに効率的な打鍵を行う際の運動を解析した.[6]は医療手技である縫合手技の動作を医療支援ロボット「da Vinci」によって計測し,両腕の関節の位置情報から評価を行っている.[7]はテニスのサーブをエネルギーフローという観点から解析するため光学式モーションキャプチャで運動解析を行っている.また,機械学習を用いて動作を分類する研究もおこなわれている.

これらの研究では運動解析を行うにあたって人が角速度や加速度などの情報を行っているがその情報は膨大な量となっている.そこで機械学習を用いれば人が対処できない量の情報でも処理することができる.モーションキャプチャから得られたデータを入力とし機械学習を用いて日常動作を判別する研究[8]では,5つの加速度センサを身体に取り付け,一定時間内の加速度の平均や二乗和を特徴量として教師有学習を行い84%の精度で20種類の動作を判別している.[9]はHMMを用いて片脚と腰に取り付けた慣性センサから133ms間の三軸の角速度,加速度,加速度の平均を特徴量として歩く,走る,階段の上り下りの動作判別を90%ほどの精度で行っている.

しかし,機械学習を用いて行われている身体動作は動作種が明らかに違うものであり,熟練度の判別のような動作の微妙な違いを指摘するような研究は行われていない.本研究では機械学習を用いた気管挿管手技の熟練度を判定するシステムを提案する.

第2章 提案手法

臨床麻酔科医：expert（麻酔科専門医以上 13 人）と novice（気管挿管の経験が浅い初期研修医 15 人）の気管挿管動作時の全身関節の角速度を取得し，教師あり機械学習により expert であるか novice であるかの判定を行う．窓移動法によって一定時間ずつを切り取りデータセットとした．学習には DNN (Deep Neural Network) と SVM (Support Vector Machine) を用い性能を比較する．

2.1 装置

Perception Neuron® (Noitom 社)



Figure 1: Perception Neuron 装着時の様子

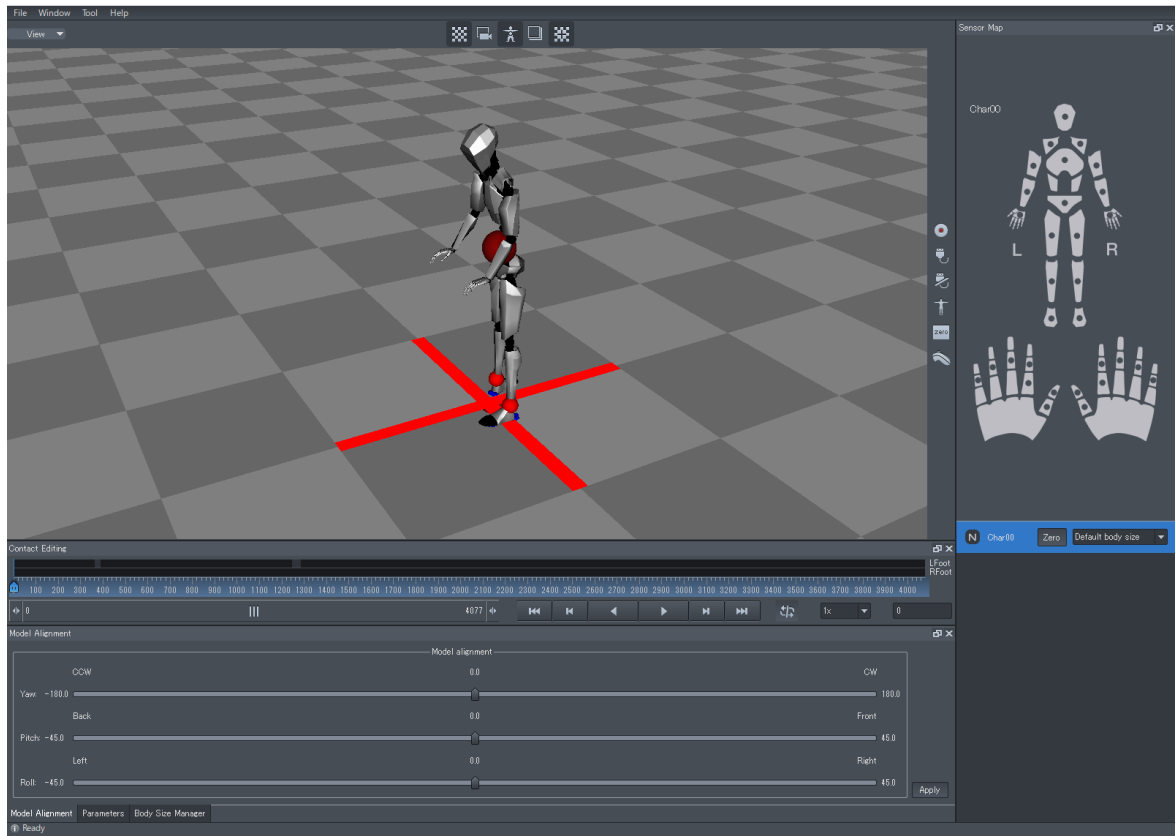


Figure 2:AxisNeuron システム画面

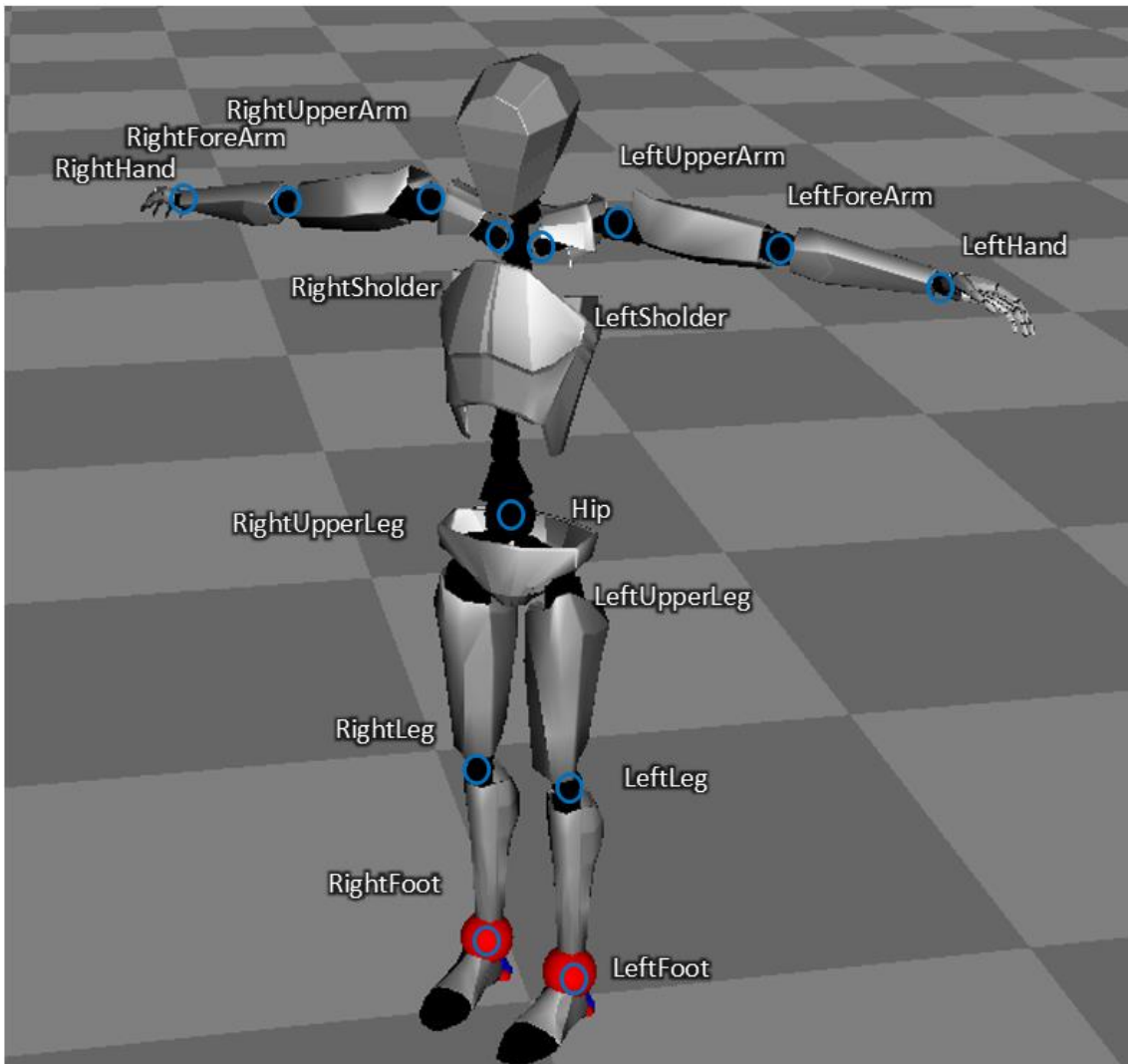


Figure 3 : 角速度取得関節

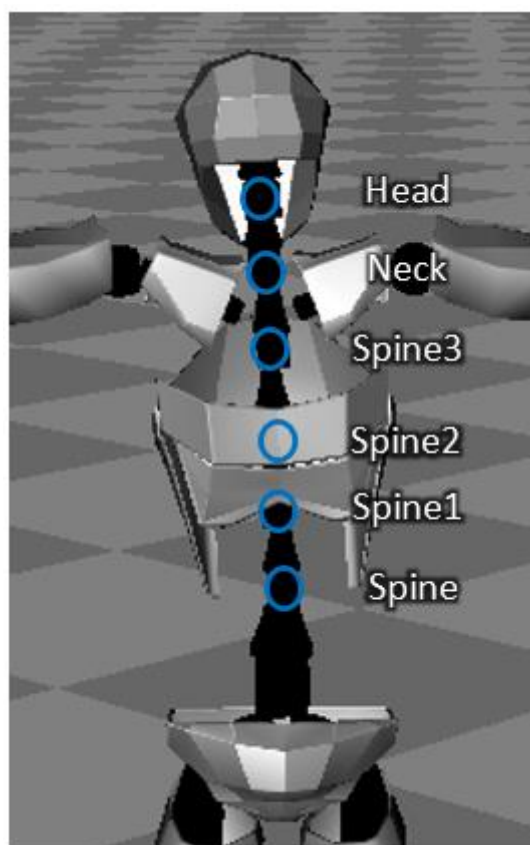


Figure 4 : 角速度取得関節 (背面)

2.2 動作データ前処理

2.2.1 データ計測

対象者それぞれ Figure. 1 に示す, 全身 21 関節の位置, 姿勢が測定できるモーションキャプチャ装置 : Perception Neuron (Noitom 社)を装着し, 気管挿管練習用マネキン : AirSim (日本ライトサービス社)にマッキントッシュ喉頭鏡を用い, 気管チューブ(スタイレット有)を 3 回ずつ挿管する動作を 120 フレーム/秒で取得した.

2.2.2 時系列データの切り出し

学習器に入力するデータとして純度を高めるため,気管挿管手技の開始点と終了点をそれぞれ喉頭鏡挿入の瞬間,チューブ挿入完了の瞬間とし,計測した動作データの中から手技を行っている部分だけを手動で切り取った.

2.2.3 窓移動法

学習器に入力するデータを作成する際に,窓移動法を用いる.結果として手技の1試行に対し複数回の判定を行うことになる.今回の窓移動法で窓を切り取る際,0.25秒ずつずらしながら入力データ作成を行った.窓幅の決定については,気管挿管手技に対して適切な窓幅は未知であったため,0.5秒,1.0秒,1.5秒,2.0秒の4通りを試した.

2.2.4 特徴量の検討

モーションキャプチャによって得られるデータとして,全身21関節の角速度,位置,速度,加速度がある.今回使用するモーションキャプチャ慣性センサを使用したもので,位置や速度は仮定された骨格の大きさをと各関節の角速度をもとに算出していることもあり,実際の値に対してズレが大きくなる.そこで本研究では信頼度の高い角速度を採用し,さらに座標系の取りかたの影響をなくすため三軸合成角速度 ω のみを各関節を代表する特徴量として取り扱うこととした.

$$\omega = \sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2}$$

2.3 学習器

2.3.1 DNN

今回用いる DNN の構造は入力層,隠れ層 2 層,出力層の計 4 層からなるものでから成るもので,活性化関数には `relu` 関数を用いた.過学習を防ぐため `dropout` と L2 正則化を採用した.`Dropout` の調整パラメータ p (ニューロンの活性化の割合)は[10]を参考に,隠れ層 1 層目を 0.8,残りの隠れ層では 0.5 とした.これらの設計には Python と DNN を実装するためのライブラリである `Chainer` を用いた.

Table 1:DNN 構成 (窓幅 0.5[s])

	Unit 数	p
入力層	1260	0.8
隠れ層	756	0.5
隠れ層	756	0.5
出力層	2	0.0

Table 2:DNN 構成 (窓幅 1.0[s])

	Unit 数	p
入力層	2520	0.8
隠れ層	1512	0.5
隠れ層	1512	0.5
出力層	2	0.0

Table 3:DNN 構成 (窓幅 1.5[s])

	Unit 数	p
入力層	3780	0.8
隠れ層	2268	0.5
隠れ層	2268	0.5
出力層	2	0.0

Table 4:DNN 構成 (窓幅 2.0[s])

	Unit 数	p
入力層	5040	0.8
隠れ層	3024	0.5
隠れ層	3024	0.5
出力層	2	0.0

2.3.2 SVM

SVMの実装については、Python 機械学習専用のライブラリである scikit-learn を用いた。今回使用するのは RBF カーネルを用いた SVM であり、ハイパーパラメータは誤分類に対するペナルティの大きさを制御する C と、RBF カーネルのバンド幅を制御する γ である。パラメータ調整に関してはグリッドサーチにより適当な候補値の組み合わせを設定する。候補値として $C:[1, 10, 100, 1000]$, $\gamma:[0.01, 0.001, 0.0001]$ を用意し、最適化を行った。

2.3.3 判定方法

手技 1 試行に対し被験者が熟練者であるか初学者であるかの判定する方法について Figure5 に示す。被験者が熟練者であるかどうかの判別は、熟練者度をもとに行う。“熟練者と判定された窓”の個数を E とし、“初学者と判定された窓”を N とすると、熟練者度は以下のようにして求められる。

$$\text{熟練者度} = \frac{E}{E + N}$$

熟練者群に対する熟練者度の平均値を E_M とし初学者群に対する熟練者度の平均値を N_M とし、被験者の熟練者度に対する熟練者であるか初学者であるかの閾値を以下の式で求めた。

$$\frac{E_M + N_M}{2}$$

この閾値より熟練者度が高ければ、被験者は熟練者と判断される。

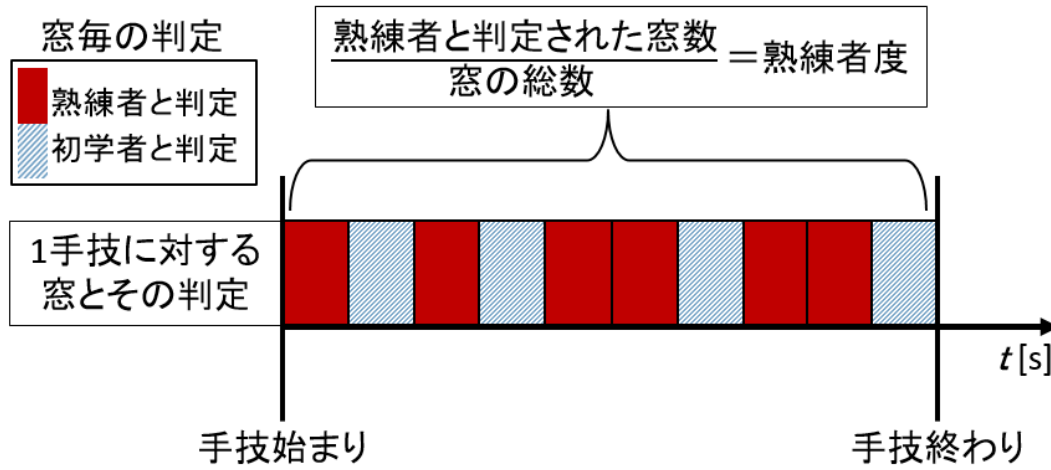


Figure 5 : 手技に対する熟練者度とその算出方法の例

第3章 実験

3.1 計測

三重大学附属病院の臨床麻酔部の協力を得て、麻酔科専門医以上 13 人と気管挿管の経験が浅い初期研修医 13 人に対し計測を行った。



Figure 6 : 計測時の様子

3.2 フェーズ分け

時々刻々の判定と手技動作を関連付ける為、喉頭鏡挿入時から気管チューブ挿入完了までを専門医の目視により 3 段階のフェーズに分けた。

フェーズ1：喉頭鏡挿入から咽頭展開まで

フェーズ2：咽頭展開から挿管チューブ挿入開始まで

フェーズ3：挿管チューブ挿入開始から挿入完了まで

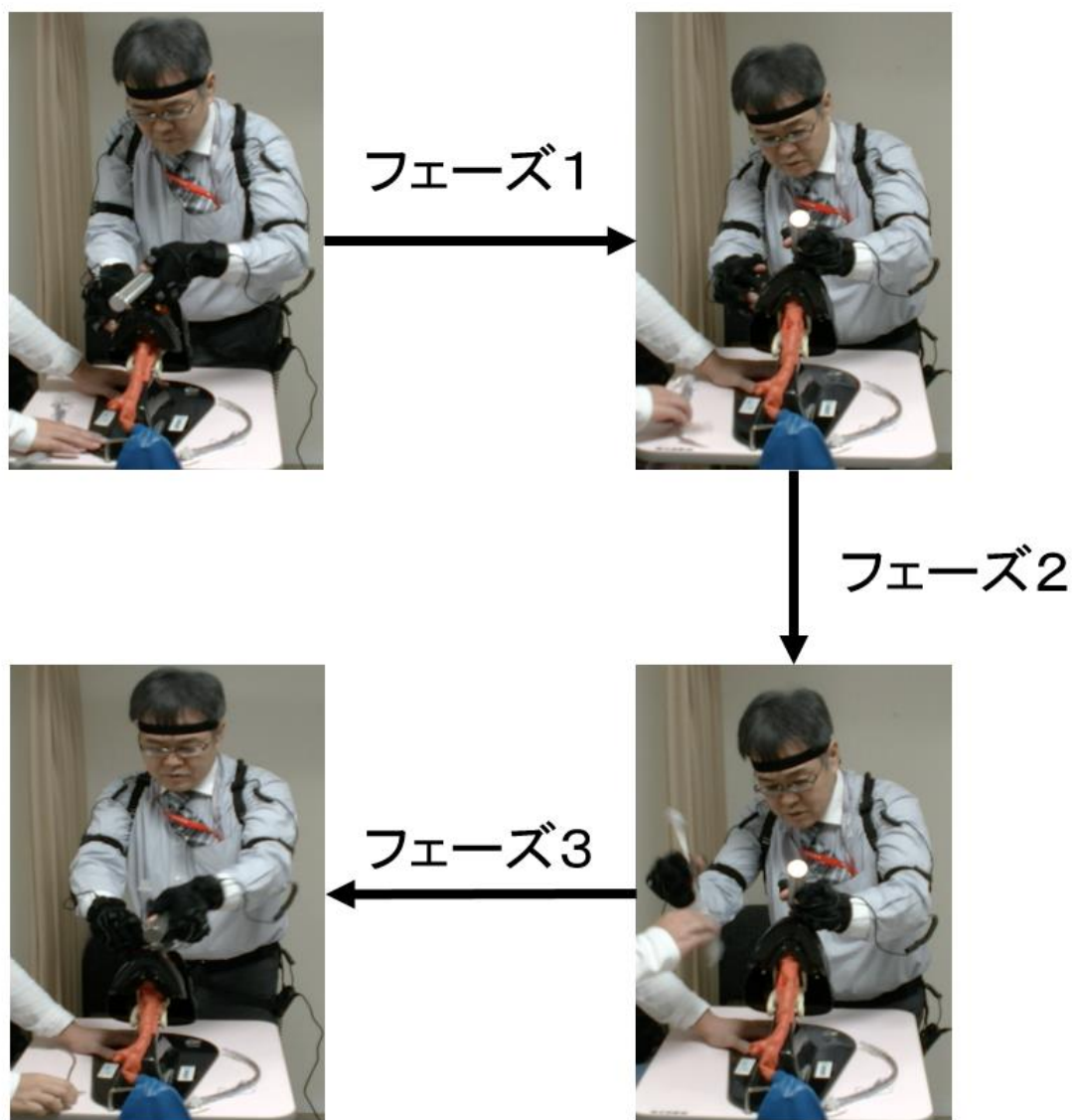


Figure 7：フェーズ分割
{喉頭鏡挿入（左上）,咽頭展開（右上）,挿管チューブ挿入開始（右下）,挿入完了（左下）}

3.3 手技の所要時間

気管挿管手技1 試行当たりの所要時間を以下のグラフに示す.

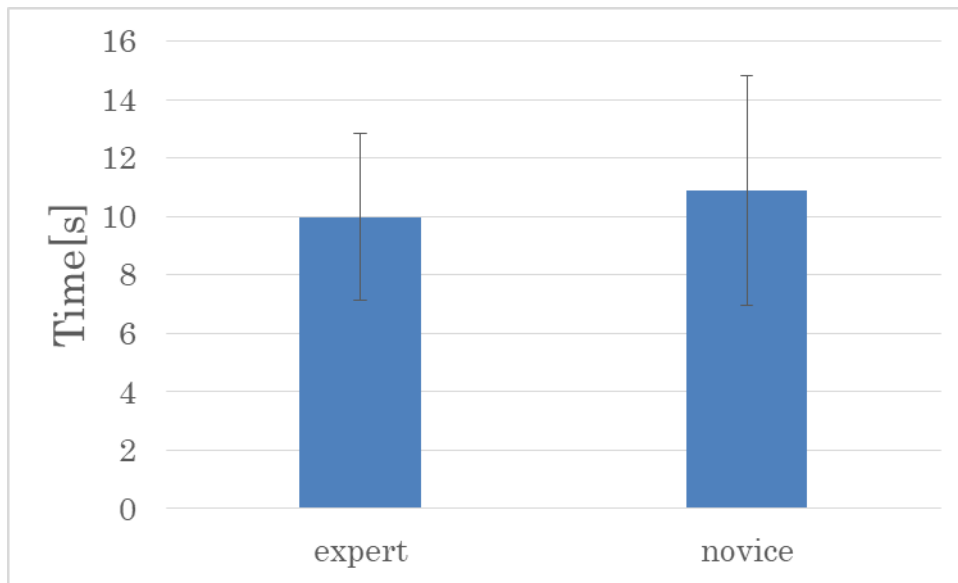


Figure 8 : 気管挿管手技 1 試行当たりの所要時間 (エラーバーは標準偏差)

従来研究[1][2]では,手技に対する所要時間を熟練度の判断基準として用いている.しかし,今回の取り扱う気管挿管手技において手技の所要時間ではなく手技の丁寧さを評価したい.本研究では窓移動法によって手技に対する評価から時間による影響を受けないようにしている.

3.4 実験条件

実験条件を Table5 に示す.Table5 に示した実験条件でさらに学習器(DNN,SVM)の比較を行う.

Table 5 : 学習時における実験条件

因子	水準
被験者	2 水準 : 熟練者,初学者
窓幅	4 水準 : 0.5,1.0,1.5,2.0[s]

3.5 結果

3.5.1 窓幅,学習器の比較

Figure9,10 はそれぞれ窓幅 0.5s の時の交差検証一回目の DNN の学習結果である.各窓に対する判定結果を時系列に並べたグラフであり,行は一回ずつの気管挿管の試行に対応しており右端に “ * ” マークがついているものがテストデータ,残りは訓練データとなっている.1 試行当たりの窓の総数で熟練者と判定された窓の個数を割ったものをその試行における “熟練者度” とする(Figure9,10 グラフ右端).SVM の学習についても同様に行う.

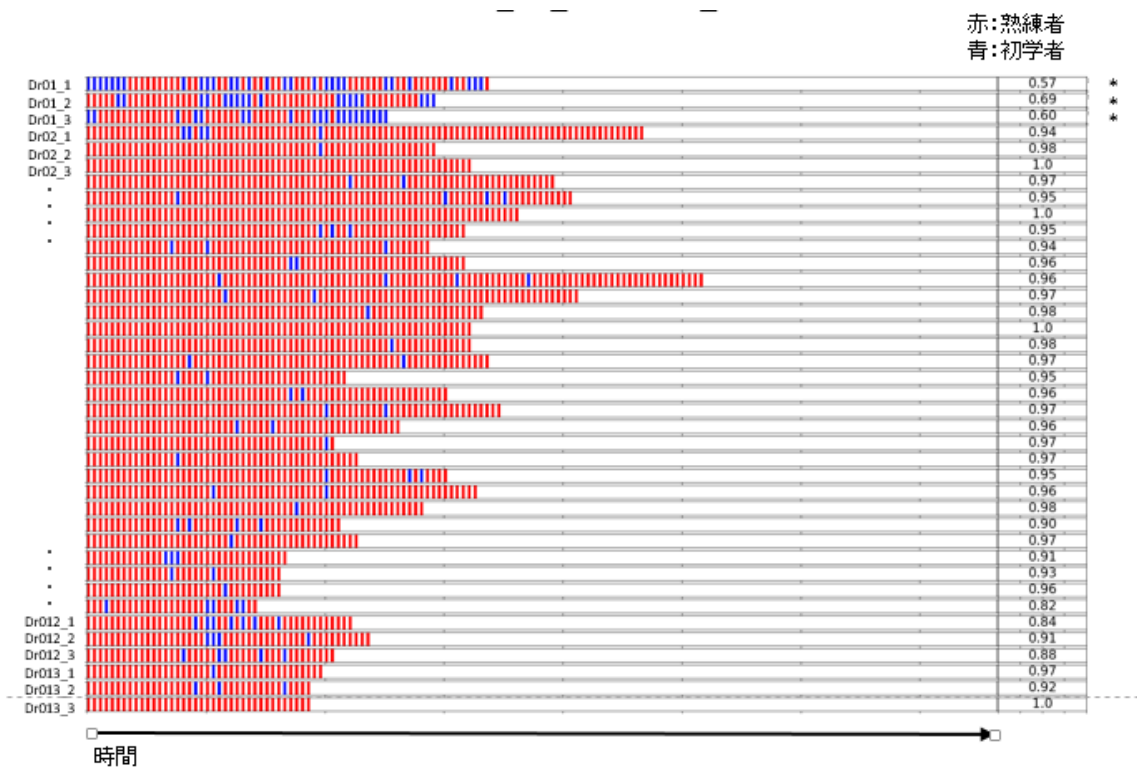


Figure 9 : DNN による交差検証 1 回目の熟練者に対する判定

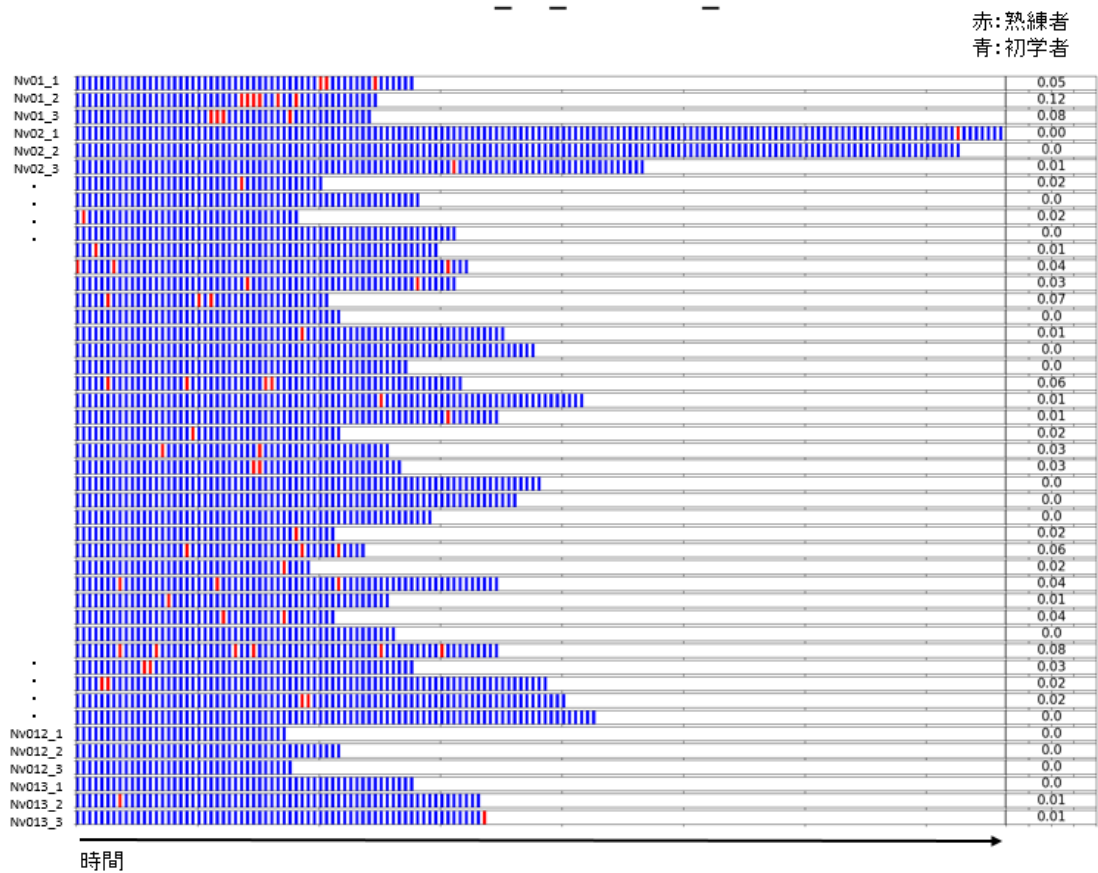


Figure 10 : DNN による交差検証 1 回目の初学者に対する判定

28 人分の被験者のデータを leave one out 交差検証により学習し,窓幅 0.5,1.0,1.5,2.0[s]に対し熟練者と初学者の間の熟練度の比較を行った.Figure 11,12 はそれぞれ DNN と SVM による判定結果である.

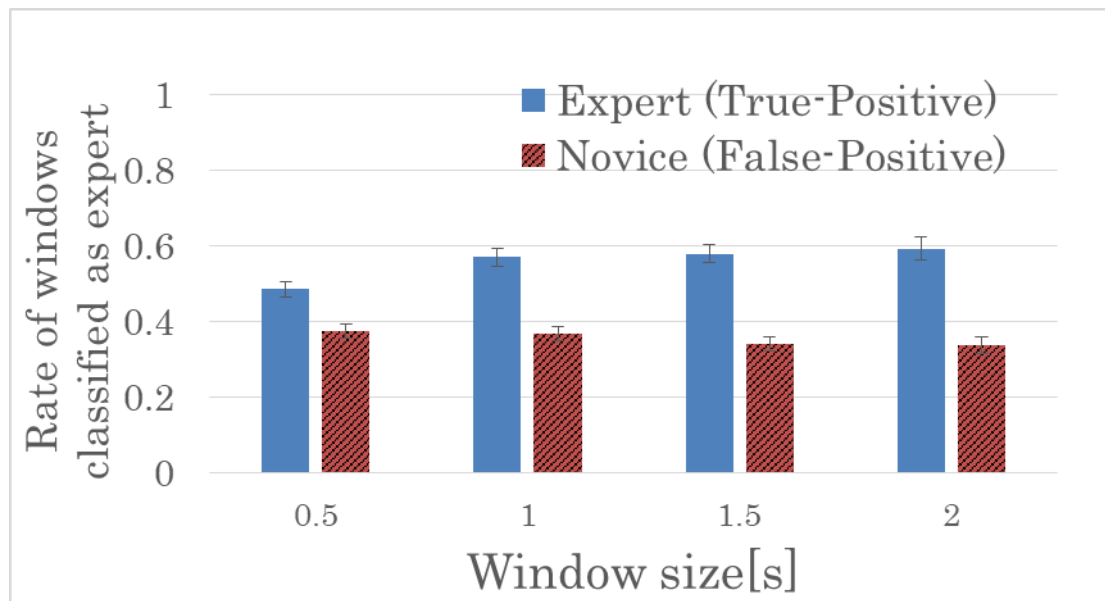


Figure 11 : DNN による学習結果,窓幅に対する熟練者,初学者の判定

分散分析（多因子実験）を適用した結果，被験者の熟練者度への因子効果(熟練者群,初学者群における熟練者度の母平均の違い)には有意差が認められた(検定統計量 $F(1, 180)=154.65, p=0.00$).

窓幅 0.5[s]の熟練者度の被験者因子（熟練者か初学者による違い）における違いを調べるため t 検定を適用した結果，有意差が認められた(検定統計量 $t(82)=4.06, p=1.12 \times 10^{-4}$). 窓幅 1.0[s]の熟練者度の被験者因子（熟練者か初学者による違い）における違いを調べるため t 検定を適用した結果，有意差が認められた(検定統計量 $t(82)=6.61, p=3.62 \times 10^{-9}$). 窓幅 1.5[s]の熟練者度の被験者因子（熟練者か初学者による違い）における違いを調べるため t 検定を適用した結果，有意差が認められた(検定統計量 $t(82)=7.82, p=1.56 \times 10^{-11}$). 窓幅 2.0[s]の熟練者度の被験者因子（熟練者か初学者による違い）における違いを調べるため t 検定を適用した結果，有意差が認められた(検定統計量 $t(82)=6.62, p=3.46 \times 10^{-9}$).

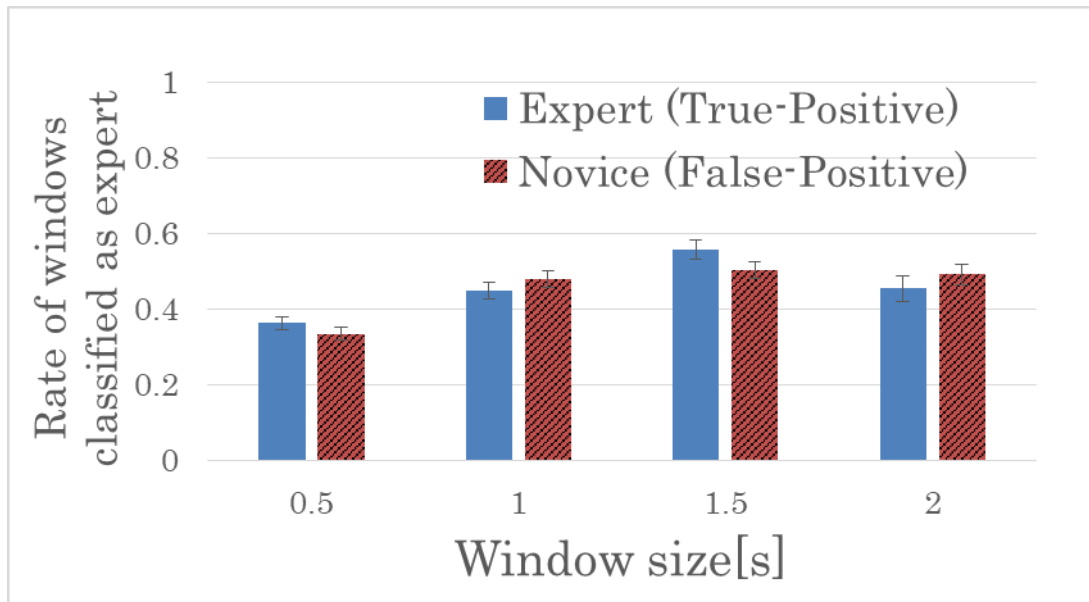


Figure 12 : SVM による学習結果,窓幅に対する熟練者,初学者の判定

分散分析（多因子実験）を適用した結果，被験者の熟練者度への因子効果(熟練者群,初学者群における熟練者度の母平均の違い)には有意差が認められた(検定統計量 $F(1, 180)=0.045, p=0.83$).

窓幅 0.5[s]の熟練者度の被験者因子（熟練者か初学者による違い）における違いを調べるため t 検定を適用した結果，有意差は認められなかった(検定統計量 $t(82)=1.18, p=0.24$). 窓幅 1.0[s]の熟練者度の被験者因子（熟練者か初学者による違い）における違いを調べるため t 検定を適用した結果，有意差は認められなかった(検定統計量 $t(82)=-0.97, p=0.33$). 窓幅 1.5[s]の熟練者度の被験者因子（熟練者か初学者による違い）における違いを調べるため t 検定を適用した結果，有意差は認められなかった(検定統計量 $t(82)=1.62, p=0.11$). 窓幅 2.0[s]の熟練者度の被験者因子（熟練者か初学者による違い）における違いを調べるため t 検定を適用した結果，有意差は認められなかった(検定統計量 $t(82)=-0.84, p=0.40$).

判定結果より,1番良く分類できていた窓幅 1.5[s]の時の結果について,熟練者度に対する閾値を定め,1試行に対し熟練者であるか初学者であるかの判定を行った.閾値は熟練者の熟練者度の平均値と初学者の熟練者度との中間値を取ることによって算出した.

$$\frac{0.580 + 0.342}{2} = 0.461$$

0.461 を閾値と定め,以下のような結果を得た(Table 7).

Table 6 : 閾値による判定

		判定結果	
		熟練者	初学者
正解	熟練者	33	6
	初学者	7	38

被験者一人に対し3回分の試行データがある.熟練者群の総試行数は39回,初学者群は45回である.ここで,2クラスの分類問題に対する評価値としてF値を用いる.F値は再現率と適合率をもとに算出する.再現率は実際に熟練者であったものに対して,熟練者であると判定されたものの割合であり,次の式で求められる.

$$\text{再現率 } E_R = \frac{T_p}{T_p + F_n}$$

適合率は熟練者と判定したものに対して,実際に熟練者であるものの割合であり,次式で求まる.

$$\text{適合率 } E_p = \frac{T_p}{T_p + F_p}$$

E_R と E_p が共に大きな値をとった時に大きくなり、総合的な性能を示す F 値は

$$F\text{値} = \frac{2 \cdot E_R \cdot E_p}{E_R + E_p}$$

で定まる。 F 値の算出結果を Table 8 に示す。

Table 7 : F 値算出結果

再現率	0.85
適合率	0.83
F 値	0.84

3.5.2 フェーズ毎の評価

窓幅の比較で,最も正確に分類ができていたのは窓幅 2.0[s]で学習器に DNN を用いた条件であった.この条件でフェーズ毎の評価を Figure 134 に示す.

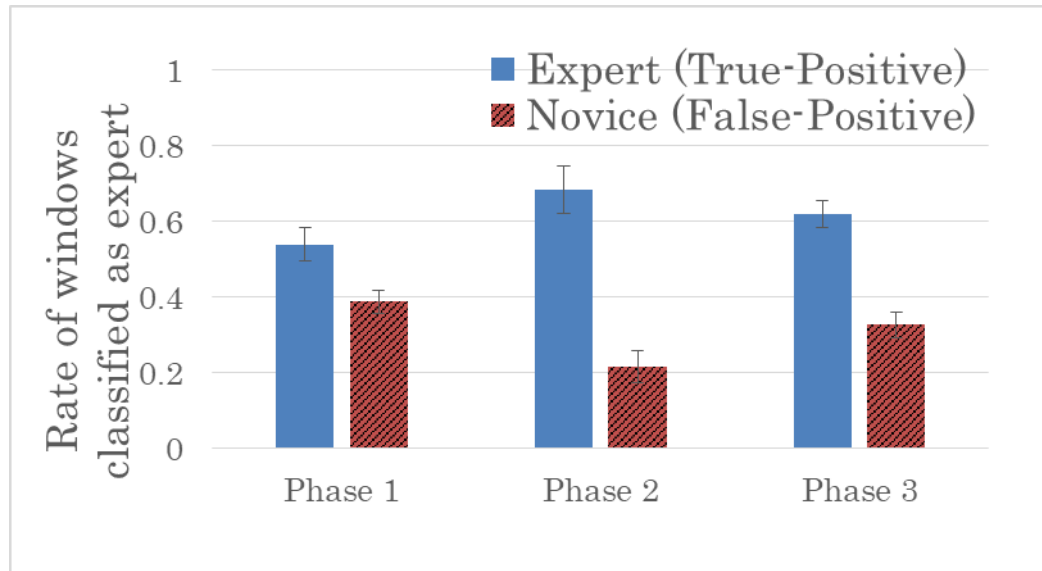


Figure 14 : 窓幅 2.0[s]の時の,DNN による学習時の熟練者度 (エラーバーは標準誤差)

分散分析 (多因子実験) を適用した結果, 被験者の熟練者度への因子効果(熟練者群,初学者群における熟練者度の母平均の違い)には有意差が認められた(検定統計量 $F(1, 135)=81.9, p=4.12 \times 10^{-17}$).

T 検定を適用した結果, フェーズ 1 間での被験者の熟練者度への因子効果(熟練者群,初学者群における熟練者度の母平均の違い) には有意差が認められた(検定統計量 $t(82)=3.67, p=4.36 \times 10^{-4}$). フェーズ 2 間での被験者の熟練者度への因子効果(熟練者群,初学者群における熟練者度の母平均の違い)に有意差が認められた(検定統計量 $t(82)=5.80, p=1.19 \times 10^{-7}$). フェーズ 3 間での被験者の熟練者度への因子効果(熟練者群,初学者群における熟練者度の母平均の違い)に有意差が認められた(検定統計量 $t(82)=6.99, p=6.76 \times 10^{-10}$).

第4章 考察

4.1 窓幅,学習器の比較

Figure 9,10 のグラフから,DNNの方がSVMより精度よく判定できていることが分かった.これは学習モデルの違いから,SVMはどこかで境界面を引くことによって分類できるようなデータセットに適しているのに対し,DNNは境界を引くことが困難なデータセットにも対応出来たということが考えられる.

また,DNNでは窓幅を大きく取った方が精度が良くなっているのに対し,SVMではそのような傾向は見られずむしろ精度が下がる結果となった.窓幅は小さければ「瞬間的な速度」を表し,大きくなるに伴い「動き」を表すものとする.DNNで窓幅を大きく取ったときに精度が良くなるのは「動き」の側面を反映しながらうまく分類が出来ていたからであり,気管挿管手技が「速度」で見分けられるほど単純ではないという事を表す結果となった.

4.2 フェーズ毎の評価

フェーズ1~3それぞれでの熟練者群,初学者群の評価を示したFigure14から,フェーズ2において熟練者と初学者の熟練者度の差が顕著に表れていることがわかる.フェーズ2は咽頭展開から挿管チューブ挿入開始までに該当するが,この間の動作において左手の喉頭鏡の固定が重要となってくることから,初学者は比較的左手の固定がおろそかになっているのではないかと考える.

次に差が大きいのはフェーズ3である.フェーズ3は挿管チューブ挿入開始から挿入完了までに該当する.実際にチューブを挿入している時間であり,このフェーズでは

チューブを挿入する右手の動きとそれに伴う患者の咽頭をのぞき込む頭部の動きや左手の喉頭鏡の固定が重要となってくる。

また、各フェーズでの熟練者度の差は、フェーズ 2,3,1 の順に大きかったが、一番重要だと考えられるのは実際にチューブを挿入しているフェーズ 3 であると考えられる。このことから、各フェーズ間における判定にそれぞれ重み付けすることによって、熟練者、初学者であるという判定をより正確に行えるようになるという可能性を示す結果となった。

第5章 結言

本論文では,手技教育において客観的な評価基準に基づき判定する手段を提案した.被験者は経験年数の長い熟練者と,研修医 1~2 年目の初学者であり,手技動作を 120Hz で全身 21 関節分にわたり取得し,データセットとした.その中の各関節の角速度を元に SVM,および深層学習により熟練者,初学者を識別する手法を提案した.

窓移動法で切り取った合成角速度の時系列データを入力とし,機械学習により窓毎に熟練者であるか初学者であるかの判定を行い,試行に対する熟練者度を算出した.さらに熟練者群と初学者群それぞれの熟練者度をもとに閾値を定め,判別を行った.

また,窓幅を変えながら DNN と SVM を比較することによって,今回の気管挿管手技のような熟練者と初学者で単純に分類できないような動作に対しても窓幅を大きく取り DNN で判定することによって分類ができた.このことから DNN は速度や加速度の合計値や平均値のような単純な特徴量よりむしろ大局的な「動き」をもとに分類していると言える結果となった.

謝辞

本研究を進めるにあたり，終始懇切丁寧な御指導，御助言を賜りました三重大学工学部 野村由司彦教授，ならび同医学部附属病院 坂本良太助教に厚く御礼申し上げます。この場をお借りして謹んで感謝の意を表します。また，気管挿管手技動作の計測を行うにあたり，意見，協力をいただきました三重大学大学院医学系研究科・医学部分子病態学 島岡要教授，同医学部附属病院臨床麻酔部 宮部雅幸教授，亀井政孝准教授，板倉庸介助教，研修医の方々にも深く御礼申し上げます。

実験の協力や助言をいただきましたプロセス解析研究室の皆様と家族に心より感謝いたします。本研究は JSPS 科研費 JP15K16257 の助成を受けたものです。

参考文献

- [1] Brydges, R.; Kurahashi, A.; Brümmer, V.; Satterthwaite, L.; Classen, R. & Dubrowski, A. (2008), 'Developing criteria for proficiency-based training of surgical technical skills using simulation: changes in performances as a function of training year', *Journal of the American College of Surgeons* **206**(2), 205--211.
- [2]阿部真美子; 山本知幸; 藤波努 & others (2003), '技能修得における身体動作のモーションキャプチャを用いた解析', *第 65 回全国大会講演論文集* **2003**(1), 351--352.
- [3]葛岡英明; 矢野博明; 渋谷良太; 山下淳; 大田祥一; 竹内保男 & others (2011), '気管挿管教育のための外装型力覚提示装置の提案 (< 特集> 教育・訓練・協調)', *日本バーチャルリアリティ学会論文誌* **16**(4), 597--605.
- [4] mikoto[医療用シミュレータロボット] 株式会社 MICOTO テクノロジー
<http://www.micotech.jp/mikoto.html>
- [5] Goebel, W. & Palmer, C. (2013), 'Temporal control and hand movement efficiency in skilled music performance', *PloS one* **8**(1), e50901.
- [6] 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.

- [7] Martin, C.; Bideau, B.; Bideau, N.; Nicolas, G.; Delamarche, P. & Kulpa, R. (2014), 'Energy flow analysis during the tennis serve: comparison between injured and noninjured tennis players', *The American journal of sports medicine* **42**(11), 2751--2760.
- [8] Bao, L. & Intille, S. (2004), 'Activity recognition from user-annotated acceleration data', *Pervasive computing*, 1--17.
- [9] Zhu, C. & Sheng, W. (2009), Human daily activity recognition in robot-assisted living using multi-sensor fusion, in 'Robotics and Automation, 2009. ICRA'09. IEEE International Conference on', pp. 2154--2159.
- [10] Srivastava, N.; Hinton, G. E.; Krizhevsky, A.; Sutskever, I. & Salakhutdinov, R. (2014), 'Dropout: a simple way to prevent neural networks from overfitting.', *Journal of machine learning research* **15**(1), 1929--1958.