

[研究助成 (C)]

臨床技能の自動採点による医学教育サポートシステムの開発

Development of Support System for Medical Education Training by Automatic Scoring of Clinical Skills

2197007



研究代表者 早稲田大学 総合機械工学科 助教 菅 宮 友莉奈
 (助成金受領時：早稲田大学大学院 先進理工学研究科 博士課程)

[研究の目的]

本研究は医学生に臨床技能訓練の機会を提供することによって、学習効率を高めることを目的としている。

技能学習は繰り返し練習や、正否を知ることによって定着するとされている。しかし、臨床技能訓練では、正否の判断は指導医が直接行っているため、すべての学生が好きな時間で自習を行うことは現実的ではない。

そこで、正否が自動でフィードバックされ、繰り返し自習が可能となることで、医療における臨床能力の質の向上が可能となると考え、採点のフィードバックが可能な診察の自習システムの構築を目的とした。

本研究で対象とする臨床技能としては、多くの学生に必要であり、学習者すべてにマンツーマンで指導を行うことが難しい行為を選定した。現在、全医学部が共用試験を実施している。共用試験とは『基礎・臨床の知識の総合的理解と問題解決能力をコンピュータを用いた客観的試験(CBT)で評価』し、『態度、臨床技能を客観的臨床能力試験(OSCE)で評価』するものである。そこで基本的な臨床能力としてOSCEで取り上げられる四肢脊柱の範囲の、頸椎の可動範囲の診察を対象として開発を行った。

[研究の内容, 成果]

診察中に医師から患者に対して行われる指示を分析し、どのようなセンシング方法によって指示を認識するかを検討し、システムの概要を決定した。その後、実際に、医師から患者に対しての指示を認識可能なモジュールの構築を行い、患者を診察可能な姿勢にしたかどうかで採点を行うことを目指した。本研究では、医師の指示の分析、システムの概要決定、認識モジュールの構築、それらの評価を行った。

1. 指示の分析実験

診察中に医師から患者に対して行われる指示の分析実験を行った。事前に、医師らへのヒアリングによって、診察において重要なことは、患者を診察可能な姿勢にすることだと判明した。そこで、学修評価項目を使用して、診察中に患者にとらせるべき姿勢についてリストアップを行った。学修評価項目とは学生が身につけるべき臨床能力の目標が記載されており、臨床技能訓練の際にも使用される教材である。

四肢脊柱の診察中に患者が取り得る5つの姿勢の写真(図1)を撮影(以下、 S_0 - S_4)した。被験者に S_0 の写真を見せ、 S_1 の姿勢になるように指示を思いつく限り列挙してもらった。同様に S_1 から S_2 への状態を遷移させる指示表現を列挙と、すべての姿勢遷移について取得した。なお、患者への指示のため、万人にわかりやす



図1 診察中に患者が取り得る姿勢

い表現としたが、指示方法に制限は設けなかった。

この実験によって、患者の現在の姿勢に応じて医師は指示を変えるということが判明した。さらに、音声指示については大きく2種類の表現があった。

- ・床を見てください
- ・もどしてください

の2種類であり、自立語が動詞1つの場合と自立語が動詞1つと名詞複数の組み合わせである。以上のことから、*Noun* を名詞、*Verb* を動詞として、以下の様に記載できる。

- ・[*Noun*を]*Verb* てください

[] 部分は省略可として、この文が認識できれば、指示が認識可能であると考えられる。

さらに、その際の *Noun* と *Verb* の組み合わせによって次の患者の姿勢が決定する。また、指示に使用される文章は限定された自立語からなっており、すべて記述しきることが可能な分量であると考えた。

さらに、補助的に動作情報を使うことがあると判明した。例えば、『こちらを向いてください』や『このようにしてください』といった、こそあど言葉を使用した際に、動作で表す指示である。動作を用いた指示表現は分析すると向きか手本を表す場合に限定されることもわかっ

た。

これらを整理すると次のようになる。

- ・医師は患者の姿勢に応じて指示を変える
- ・指示は主に音声指示、補助的に動作指示を使用
- ・音声指示に含まれる自立語は動詞1つと0以上の名詞
- ・使用される動詞・名詞の数は有限
- ・動作は指示代名詞に対応して、向きか手本を表す

以上のことから、本システムに必要な要素は、診察中の医師の指示を音声認識および動作認識によって認識し、動作による補助的な情報を含めてテキスト化する。テキスト化した情報から自立語を抜き出し、現在の姿勢に対して、実験によって取得した情報を照らし合わせることで、次の姿勢の遷移先について判断を行う。姿勢の情報管理については状態遷移モデルを用いることが適当であると考えた。そして、現在の患者の姿勢を医師に提示するための表示が必要である。最後に、状態遷移モデルに基づいて、各姿勢に達したかどうかを数え上げ、採点をフィードバックする必要がある。

2. 音声認識モジュールの構築

指示の分析実験によって、診察中の医師の音声は少ないバリエーションであることがわかっている。本システムではアドバンスト・メディア社の *AmiVoice SP2* を用いて音声認識を行った。*Amivoice* は学習済みの音響モデルが搭載されており、オフラインでの使用が可能である。さらに、認識可能な文型・単語のカスタマイズが可能という特徴を持っており、さらに、組み込み用のAPIが存在しているため、動作認識を含めた一連のシステムを一つのソフトとして開発が可能となる。

また、名詞や助詞や語尾については省略されることがあるため、下記の様な簡易な構文とし

て表した。

[〈名詞〉][を] 〈動詞〉 [ください]

[〈名詞〉][を] 〈動詞〉 [もらっていいですか]

[〈名詞〉][を] 〈動詞〉 [いただいいていいですか]

[] 内は省略可能であり 〈 〉 部分はカスタマイズした単語が認識可能となる。さらに、この方法によって記述した動詞と名詞は、AmiVoice の API を使用することによって、認識を行った際に、単語と品詞が抽出される。

4. 音声認識の認識率検証

構築した音声認識手法の精度検証実験を行った。対象は 20 代の男女 8 名とし、1) 認識可能な動詞、名詞の単語リストを用意。2) 任意の動詞もしくは動詞と名詞の組み合わせからなる文章を発話してもらう。3) 1 人につき 10 文章を発話してもらう。

以上の方法により、発話された 80 文章の認識結果と発話内容が一致するかについて比較した。

その結果、動詞は 80 語中 79 語一致し、名詞は 80 語中 80 語認識された。精度としては約 99% ととなった。また、認識が一致しなかったものは、発話は『向いて』だったが、認識は『見て』となっていた。

有効性について議論する。音声認識では、認識の際に使用される音響モデルの構築には膨大なデータが必要となり、昨今ではビッグデータを使用した認識も行われているが、その場合も膨大なデータが必要となる。つまり、その言語の話者の人口に影響を受けることになる。また、発話の状況が原稿を読むのか、自由発話なのかで大きく精度が変わる。日本語では、環境を整えた場合に Julius という音声認識ソフトでは認識率が 80-90% を示しており、国会審議の場（言い直しなどがあり認識しにくい点もあるが、使用される語数としては限定的）での認識率は 86% となっている。話者人口が多く、音声認識が進んでいる英語の場合で、アナウンサーの発話を認識させた場合は、90% 程度の高い認

識率を得ている。

以上のことから、本認識システムによる認識率 99% は、90% を大きく上回る物であることから、本音声認識方法による音声認識の精度は実用としても問題無い精度であると考えた。

また、認識ができなかった 1 語については、『向いて』を『見て』と認識しており、どちらもマ行下一段活用であり、音素分解した際もとても似ていることが原因と考えられる。

5. 動作認識モジュールの構築

次に、動作情報について認識方法を説明する。これまでの実験で動作情報は向きもしくは手本を表すことがわかっている。本研究で対象としている四肢脊柱の診察においては、手本を示す指示がほとんど見られないことから、向きの認識のみを行った。さらに、利き手でカルテを記載し、逆の手で患者への指示を行うことが一般的なため、右手を利き手と仮定して、左手が向きの指示を行う手と考えた。

診察中の動作指示は、手を使用することがわかっている。診察以外の場な日常的な場面では目線・顎などによる向きの指示も想定されるが、診察中は不適切と考えられる。そこで、親指の付け根から指の先端に向かうベクトルが、向きを表すと考えた（図 2）。

また、診察中に示される向きとしては上下左右以外に『私の方』という意味の『こちら』が存在した。『私の方』を示す場合、手は体の中心部分に向かい、手が胴体から出ない傾向であった。

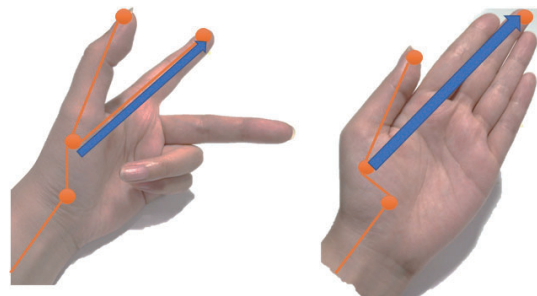


図 2 親指の付け根から先端へのベクトル

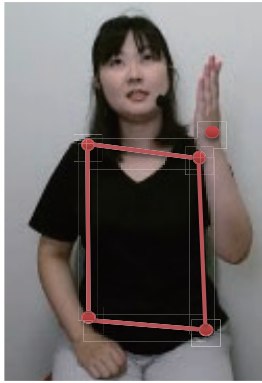


図3 手の位置と胴体の関係

以上のことを踏まえて、動作認識を以下の様に構築した。

Microsoft社の光学式3DモーションキャプチャKinect v2を用いて、学習者の骨格を認識する。両肩と左右の腰骨からなる四角形を仮定し、左手がその四角形に入っていた場合は『私の方』とした(図3)。そして、その四角形から左手が出ていた場合に、手首から指の先端へのベクトルを推定し、90度ずつ『上』『左』『下』『右』として認識した。

次に、音声認識モジュールと動作認識を統合する方法を記載する。まず、医師の振る舞いは1文章単位で認識を行う。

そして、その文中に指示代名詞が使用されている場合には、動作情報と置換する。しかし、『あちらを向いてください』と発話する間に約30~100フレーム程度の時間経過が起るため、発話中に複数の向きを示す可能性がある。さらに、発話の前後でも向きを示す可能性があるため、図4で示す6通りの状況について、検討した。『あちらをむいてください』という発話に対して、『あちら』に該当する向きを表す動作が行われたタイミングを示している。このうち発話以前に動作が終わる①と発話後に動作が始まる⑥に関しては不自然と考え、②~⑤が該当すると仮定した。

そして、発話中の最大フレーム数で示している向きが、『あちら』に該当する向きであると考えた。

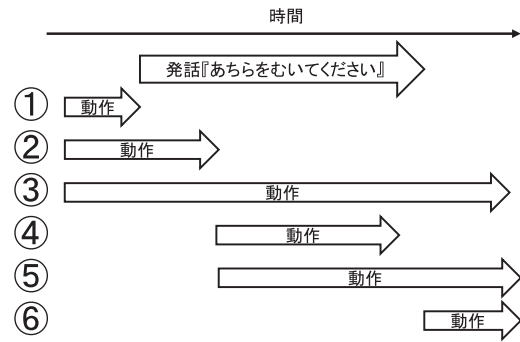


図4 発話と動作のタイミング

以上をまとめると、音声認識と動作認識の統合は次のように行う。まず、認識した文の向きを表す指示代名詞が含まれているか調べ、指示代名詞が含まれていた場合には、発話中の最大フレームで示していた向きについて名詞化して、指示代名詞に置換する。

6. 認識実験

構築した指示の認識手法について認識精度の検証を行った。被験者は2-30代の男性7名として、下記の方法で行った。

- 1) 医師役として患者に指示をしてもらう。
- 2) 指示書に従って、発話しながら向きを示してもらう。
- 3) 1人の被験者に対して上下左右自分の方の5つの向きを示してもらう。
- 4) 一連の動作を用いた指示をシステムによって記録する。

発話の順序による実験結果への影響を考慮し、指示書中の1~5の発話は被験者によってランダムに設定した。また、自分を示す際には違和感の軽減のために『こちらをむいてください』という文章にした。

その結果、35回の発話中31回で正しく、音声認識に対して向きの置換が行われた。精度としては89%となった。この実験では図4で検討した⑥の方法で向きを示した被験者がいた。⑥は発話が完全に終了してから向きを示すパターンであり、不自然であるとして、認識から

除外した。⑥の発話パターンについて不適切であるか、医師や患者らにヒアリングを行うなどして、さらなる検証を行いたい。

や人と対面をしない形で行えるようになることが、貴重な対面訓練の時間を有効に使用する一助となると期待する。

[今後の研究の方向, 課題]

本研究では、医師の指示の分析を行い、医師の指示が音声と動作の情報からなることを明らかにした。必要なシステムの概要決定を行った。そして、音声認識モジュールの構築・動作認識モジュールの構築、それらの統合および評価を行った。その結果、十分な認識率を得ていることを確認した。現在、医師に患者の状態をフィードバックするためのCG患者の開発を進めており、本システムを用いて認識および採点を行った結果が、医師らによる採点の結果と一致するか検証を進めている。

本研究は、医学生に臨床技能訓練の機会を提供することによって、学習効率を高めることを目的としているが、研究を進める中で新型コロナウイルスの蔓延が起これ、臨床技能訓練の実施が非常に困難になる事態が起きている。臨床技能訓練は指導医、患者、学生が行うことが最善であるが、対面・対話・接触が難しく、これまでよりも技能訓練の方法が制限されている。本研究を進めていき、初学者の臨床訓練が在宅

[成果の発表, 論文等]

■ 査読付き国際学会

- [1] Yurina Sugamiya, Takuya Otani, Ryu Nakadate, Atsuo Takanishi, "Construction of Automatic Scoring System to Support Objective Evaluation of Clinical Skills in Medical Education", 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4177-4181, 2019.

■ 査読無し国内会議

- [2] 菅宮友莉奈, 大谷拓也, 中橋龍, 大久保由美子, 山内かづ代, 高西淳夫, "状態遷移モデルと患者CGを用いた臨床技能訓練システム" 第37回日本ロボット学会学術講演会, 2019年.

■ シンポジウム, 交流会

- [3] TWIns10周年記念 東京女子医科大学・早稲田大学ジョイントシンポジウム, 音声認識を用いた臨床手技の自動採点システム, 2018年6月
- [4] 第6回研究交流セミナー (東京女子医科大学, 早稲田大学先端生命医科学センター), 音声認識を用いた臨床手技の自動採点システム, 2019年2月
- [5] Symposium for Women Researchers 戸山高校研究交流会, 音声認識を用いた臨床手技の自動採点システム, 2018年11月