

一般口演

一般口演2

病院情報システム2（ハードウェア，インターフェース）

2018年11月23日(金) 10:15 ～ 11:45 F会場 (5F 502+503)

[2-F-1-4] 音声認識技術と自動操作技術を併用した電子カルテ音声入力操作システムの開発

○山ノ内 祥訓¹, 中村 太志², 宇宿 功市郎²（1.熊本大学医学部附属病院総合臨床研究部, 2.熊本大学医学部附属病院医療情報経営企画部）

【はじめに】現在電子カルテシステムは機能が多く操作が複雑化している。利用者の操作習熟に多くの時間が必要で、一つの業務終了に時間を要している。この課題の解決にはユーザインターフェースデザインの再設計が必要であるが、コスト面から困難である。本研究では、電子カルテ自体に手を入れることなく、Speech to Text技術と Robotic Process Automation技術を併用した電子カルテの音声入力と音声操作システムを試作し評価を行うことで解決を目指した。【方法】当院の IBM社製電子カルテシステムを操作対象とした。音声認識はクラウドサービスを用いてインターネット接続 PCで音声を文字に変換し、結果を Bluetooth通信で電子カルテ端末に送信した。音声操作は電子カルテ端末に Microsoftの UIAutomationを使用した自動操作プログラムを配置し、受信した結果から電子カルテを操作する仕組みとした。【結果】音声認識は1795字の文章を読み上げたところ94%の精度で行え、入力時間は16%の短縮が行なえた。音声操作はカルテ記載やオーダ入力等の操作シナリオを5本用意し音声による電子カルテの操作を行った。音声のみで完結できたのは2本、途中手入力による補正を行い完結できたのは3本だった。よって音声認識技術と自動操作技術を用いることで、必要最小限の操作で入力操作を完結できることを確認した。【考察】音声によるテキスト入力は、一部認識間違いがあったがキーボードによる補正が少なく実用可能と考える。音声による自動操作は、今回のシナリオ操作はほぼ行えたが、入力中発生する各種チェックや複雑な条件分岐などに対応した自動操作を音声のみで実現するのは困難なことが確認できた。今後は、音声入力化、音声操作化したときのメリット・デメリットを考慮した入力場面に本システムを適用することで業務軽減への貢献が期待できる。

音声認識技術と自動操作技術を併用した電子カルテ音声入力操作システムの開発

山ノ内祥訓^{*1}、中村太志^{*2}、宇宿功市郎^{*2}

^{*1} 熊本大学医学部附属病院 総合臨床研究部、^{*2} 熊本大学医学部附属病院 医療情報経営企画部

Development of voice input and operation system in the Electronic Medical Record using Speech to Text and Robotic Process Automation

Yamanouchi Yoshinori^{*1}, Nakamura Taishi^{*2}, Usuku Koichiro^{*2}

^{*1} Department of Clinical Investigation, Kumamoto University Hospital,

^{*2} Department of Medical Informatics and Administration Planning, Kumamoto University Hospital

Since the medical care service and the input of the electronic medical records (EMRs) have become complicated, the time to become familiar with the use of EMR system and the one to input articles of EMRs are increasing. In order to solve these problems, we have developed a voice input and voice operation system of EMR using Speech to Text technology and Robotic Process Automation technology, respectively. Speech recognition uses cloud service to convert voice into letters on Internet connected PC by transmitting the result to the EMR terminal on Bluetooth communication. In the voice operation, an automatic operation program using Microsoft's UI Automation was installed on the EMR terminal, and those were operated by the received result. Speech recognition showed 94% precision, and input time was shortened by 5.56%. Among five prepared operation scenarios such as chart description and order entry after selecting patients by voice, 2 scenarios were able to be completed only by voice. From these results, our system can be expected to contribute to work load reduction on applying to the character input scene by considering the merit and demerit of voice input and voice operation.

Keywords: Electronic Medical Record, Speech Recognition, Robotic Process Automation.

1. 緒論

電子カルテシステムをはじめとする病院情報システムの普及に伴い、医療情報を電子的に入力することが多くなっている。情報システムの入力は大半がキーボードとマウスによって行われ、特に医療情報の大半を占める文字情報はキーボードがほぼ唯一の入力手段となっている。しかし、キーボードによる文字入力は、以前に比べればある程度スムーズに入力できる医療者は増えているが、習熟しないと高速に入力することが困難である。また、キーボードやマウスは医療者が多く触れる個所でもあることから、院内感染の媒介になっているという報告もある。そのため、キーボード以外の文字入力方法として以前より音声テキスト変換(Speech to Text)技術を使用した音声認識システムの研究が進められている¹⁾。

医療分野における音声入力は、これまで放射線検査や超音波検査、ホルター心電図検査のレポート作成、看護記録の記載、歯科検査の記載、トリアージ記録の記載などの事例が報告されており、一定の成果が見られている²⁻⁸⁾。しかし、いずれの場合も、マイク付きヘッドセットを装着しなければならないこと、音声認識システムに対し事前に用語辞書の登録が必要であり運用後のメンテナンスも必要であること、音声認識システムの特性に合わせた発話が必要で習熟に時間がかかること、音声認識システム一式を準備すると高額であること、といった制限があり、一部部署のみの普及にとどまっている。その一方で、スマートホンやスマートスピーカーには高精度の音声認識エンジンが搭載され一般に普及している。しかし、これらの音声認識はクラウドサービスとして実装されていることから、

クローズドなネットワークで運用されることの多い病院情報システムでの活用が困難であった。

また、電子カルテシステムは、以前に比べて非常に多くの機能が実装されており、様々なことが電子化され医療の質と安全の向上に役立っている。しかし、その一方で操作が複雑化し、利用者の操作習熟に多くの時間が必要である。さらに、一つの業務を行う上で必要な操作が多く、業務終了に多くの時間を要している。この課題の解決にはユーザインターフェースの再設計が必要であるが、ユーザインターフェースの根本的な変更は、ほぼシステムを新しく作り変えるに等しいコストがかかると予想され、カスタマイズで行うには非現実的である。電子カルテ自体を極力カスタマイズせずに複雑な操作を簡略化する手段の一つとして、近年注目されている技術にRobotic Process Automation(RPA)がある⁹⁾。この技術はもともとソフトウェア開発のテスト工程で行われる GUI の再帰テストを自動化するために開発されたが、繰り返し業務が発生する事務作業においても有用であることから金融分野をはじめとして導入が始まっている。RPA は、あらかじめ決められたシナリオを元に PC 上に表示される GUI の解析を行い、得られたデータをパラメータとして認識し、人間が行っていた入力操作をロボットプログラムが代行する。そのため、最初に必要なデータを与えてシナリオを実行すれば複雑な情報システムの入力操作を自動的に行うことで業務の効率化が期待できる。

2. 開発目的

本研究では、電子カルテ自体にカスタマイズを行うことなく、Speech to Text 技術とRPA 技術を併用した電子カルテの音声入力と音声操作システムを開発し評価を行うことで、入力業務の効率化を行うことを目指した。

3. システム概要

当院の電子カルテシステムは IBM 社製であることから、この電子カルテシステムを音声入力及び音声操作の対象とした。音声入力は、電子カルテ端末に併設したインターネットに接続可能な小型 PC(Diginnos Stick DG-STK4)にフラット型マイク(SANWA MM-MCU06BK)を接続して、このマイクを通じて音声を認識するようにした。これをクラウドサービスに送信しテキストデータに変換した。使用するクラウドサービスは Google 社の Google Cloud Platform (GCP)で提供されている Speech API¹⁰⁾と Microsoft 社の Microsoft Azure (Azure)で提供されている Bing Speech API¹¹⁾の二つを使用して精度を比較した。2 つのクラウドサービスはどちらも医療情報システムに関する諸ガイドライン(3 省 4 ガイドライン)に適合するように構成した。次に、変換結果を電子カルテ端末に送信するが、クローズドなネットワークとなっていることから、小型 PC と電子カルテ端末を Bluetooth による通信で送信するようにした。ただし、Bluetooth は無線通信となるため Bluetooth 以外の通信方式として RS232C による有線通信も実装し、設置状況に合わせて選択できるようにした。電子カルテ端末では、受信したテキストデータを表示し、さらにキーボードによるテキスト修正が行えるアプリケーションを開発した。通常の音声入力はこのアプリケーションに表示されたテキストを適宜修正し、コピーしたうえで電子カルテシステムの入力欄に貼り付けて入力する操作とした。

一方、音声操作については、Speech to text で変換されたテキストデータを言語理解クラウドサービスに送信し操作コマンドとパラメータに分割した。このクラウドサービスは、Microsoft 社の Microsoft Azure で提供されている Language Understanding (LUIS)を使用した¹²⁾。電子カルテ端末では、受信した操作コマンドに対応した RPA の入力シナリオにパラメータを設定して実行し、電子カルテシステムの自動操作を行うようにした。自動操作ライブラリは、Microsoft 社の UIAutomationを使用した¹³⁾。全体のシステム構成図と実際に使用した小型 PC 及びマイクの写真を示す(図 1, 2)。

4. システム評価

4.1 音声入力

音声入力の評価に使用したテキストは、あらかじめ実際の電子カルテシステムに記録された SOAP 記録から 5 テキスト(合計 616 文字)、患者への説明記録から 3 テキスト(合計 452 文字)、入院診療計画書から 4 テキスト(合計 727 文字)である。これを 1 人が 2 回読み上げた音声データを 2 種類の音声認識クラウドサービスへの入力とし、元文書からのテキスト変換精度を比較した。なお、音声入力の習熟期間を与えていない。音声入力の環境は、専用マイクではなくノート PC に搭載されている内蔵マイクを使用した。変換精度の結果を示す(表 1)。SOAP 記録では、GCP の平均訂正文字数 71.5 字で認識精度 88.39%に対し、Azure の平均訂正文字数 180.5 字で認識精度 70.7%だった。説明記録では、GCP の平均訂正文字数 21.5 字で認識精度 95.24%に対し、Azure の平均訂正文字数 45.5 字で認識精度 89.93%だった。入院診療計画書では、GCP の平均訂正文字数 19.5 字で認識精度 97.32%に対し、Azure は平均訂正文字数 80 字で認識精度 89.0%だった。全

体文字数でみると GCP が認識精度 93.73%で Azure が認識精度 82.95%となり、GCP の精度が高かった。

表 1 文字認識精度の比較

	平均訂正文字数	平均認識精度
SOAP記録(616文字)		
GCP	27.5	88.39%
Azure	180.5	70.70%
説明記録(452文字)		
GCP	21.5	95.24%
Azure	45.5	89.93%
入院診療計画書(727文字)		
GCP	19.5	93.73%
Azure	80.0	89.00%

次に、音声入力速度に関する結果を示す(表 2)。音声入力速度の評価は、音声認識と同じテキストを用いて行った。こちらは GCP のみを使用し、すべてキーボードで入力したときの時間と音声入力とその後のキーボードによる定性時間の合計を比較した。入力操作は 7 人が 1 回入力した結果の平均を比較した。評価者は、キーボードの操作は習熟しているが音声入力の習熟期間を与えていない。その結果、SOAP 記録では、キーボード入力時間 746 秒に対し音声入力時間 827.86 秒、説明記録では、キーボード入力時間 425.86 秒に対し音声入力時間 370.57 秒、入院診療計画書では、キーボード入力時間 691.43 秒に対し音声入力時間 561.29 秒となり、全体ではキーボード入力時間が 1863.29 秒、音声入力時間が 1759.71 秒で、音声入力のほうが 5.56%速かった。

表 2 キーボードと音声入力速度の比較

	SOAP記録 (616文字)	説明記録 (452文字)	入院診療計画書 (727文字)
平均キーボード入力時間 [sec]	746	425.86	691.43
平均音声入力時間[sec]	267	127.29	210.71
平均入力訂正時間[sec]	560.86	243.29	350.57
平均音声+訂正時間[sec]	827.86	370.57	561.29
速度改善割合(訂正無[%])	36.95	31.78	31.84
速度改善割合(訂正有[%])	112.13	88.99	83.14

4.2 音声操作

音声操作に関する評価は、以下の 5 つの操作シナリオを準備し、音声で入力と操作ができるかどうかを確認した。

- A) 氏名で患者検索を行い該当の患者のカルテを開いた後、SOAP 記録を入力し、カルテを保存終了する。
- B) 氏名で患者検索を行い該当の患者のカルテを開いた後、説明記録を入力し、カルテを保存終了する。
- C) 氏名で患者検索を行い該当の患者のカルテを開いた後、入院診療計画書を登録し、印刷後にカルテを保存終了する。
- D) 氏名で患者検索を行い該当の患者のカルテを開いた後、スキャナオーダを登録し、カルテを保存終了する。
- E) 氏名で患者検索を行い該当の患者のカルテを開いた後、セットオーダを登録し、カルテを保存終了する。

それぞれのシナリオについて 1 人が 2 回試行したところ、シナリオ A 及び D は 2 回とも音声のみによる操作で完結することができた。シナリオ B,C では入力画面に複数の入力欄があるが、音声で入力したテキストを正しい入力欄に入力できな

かった。シナリオ D では、セッ一覧画面内でのツリービューの操作ができず、マウスによる操作が必要となった。また、音声で画面を操作する場合、画面の展開に合わせて発話による操作を行うことになるため、マウス操作を行うより時間がかかる傾向が見られた。

5. 考察

音声によるテキスト入力、一部認識間違いがあったが修正が少なく全体の入力速度もキーボードと比較して速いため実用可能と考える。これは先行研究の結果とも一致する。一方、音声による自動操作は、今回のシナリオでは音声のみで完結できたシナリオより、キーボードやマウス操作が必要になったシナリオが多かった。入力中発生するGUI部品の操作や各種チェック、複雑な条件分岐などに対応した自動操作を音声のみで実現するのは困難なことが確認できた。

音声入力の対象とした SOAP 記録、説明記録、入院診療計画書のうち、SOAP 記録が最も認識精度が低かった。その理由として、他の 2 種類のテキストに比べて数値や医療用語、略語が多く、医療記録特有の言い回しの存在が挙げられる。一方、説明記録や入院診療計画書は患者向けの内容になることが多いため、平易な表現で記載することが多いことから認識精度が高くなったと考えられる。本研究では、音声認識システムとしてクラウドサービスを使用した。先行研究ではオンプレミスの音声認識システムを使用しているが、本システムを開発するにあたり、メーカーから事前にデモ環境を借りて予備的評価を行ったところ、認識精度が低く診療現場からも不評であったことから比較対象として採用しなかった。メーカー側からの運用環境説明では、ヘッドセットが必須ということであったため、マイクの精度や雑音が影響していると考えられるが、同じハードウェア環境でクラウドサービスのほうが高い認識精度が得られたことから今後もクラウドサービスを中心として採用することとした。入力速度の比較については、評価者がキーボードに習熟していることから 57.8 字/分という十分速い入力速度(日本情報処理検定協会日本語ワープロ検定 1 級が 70 文字/分)に対し、音声入力では 61.2 字/分であり音声入力の方が速い結果となったが、大きな差は認められなかった。特に、SOAP 記録はキーボードの入力速度の方が勝っていた。これは、SOAP 記録が 3 種類のテキストの認識精度で最も低い値であったことが訂正作業に影響していると考えている。実際に音声でテキスト化された内容と訂正後のテキストを比較したところ、漢数字とアラビア数字の訂正やアルファベットが混じった分の訂正や、句読点や改行を追加する、という訂正が多かった。また、実際に訂正作業を行った評価者にヒアリングしたところ、テキストの一部のみを修正するために単漢字変換を多く使用したことや、見直しの時間をとったことで訂正作業に時間がかかった、という回答が得られた。これらの結果から、キーボード入力に対する音声入力の優位性は大きなものではないという考え方もできる。しかしながら、診療現場の使い方として、これまで患者との対話内容や説明内容をその場で速記することは困難であり、多くの場合終了後に思い出しながら文章としての構成を整えて記載していた。そのときに必要なのは、誤変換はあっても何を話したかが一目でわかる機能であり、単純なディクテーション速度の比較で業務改善効果の評価できないと考える。

音声操作については、単純な操作シナリオであれば音声のみで操作を行えることが示された。しかし、電子カルテシステムの入力画面はテキスト入力欄以外にもラジオボタンやコンボボックス、リストボックス、タブ、スクロールなど様々な GUI

部品で構成されている。これらを全て音声で一つ一つ操作していくことは技術的には可能であるが、非常に時間のかかる操作となってしまう、実運用における業務効率化に寄与しない。また、入力画面内の GUI 部品名の特定方法を一般化することができないため、画面毎に違うコマンドを実装する必要があることから、すべての電子カルテ操作を音声で行うことは現実的ではないと考える。しかし、RPA の特徴である同じ操作の繰り返しとの組み合わせでは、行いたい業務を発話することで電子カルテの該当の画面まで開くことや、複数患者に対して同じ文書を印刷する、など定型的な業務シナリオに限定すれば有用であると考えられる。

今後は、音声入力については 9 月より複数の診療科の患者説明室や外来診察室の電子カルテ端末に環境を構築し、運用評価を行う。その後、順次設置場所を拡大していく予定である。音声操作については、対象業務と操作シナリオの見直しを行い、音声入力なしの RPA のみを先行して開発し、運用評価を行いながら音声入力まで実装するかを検討する予定である。

6. 結論

本研究では、音声入力及び音声操作システムの開発評価を行った。音声入力は診療現場においても十分実用的であり今後院内への展開を行う予定である。音声操作は、実用的な業務シナリオを再検討し、音声操作化したほうが良いものだけを今後の開発対象とする。音声入力化、音声操作化したときのメリット・デメリットを考慮した入力場面に本システムを適用することで業務軽減への貢献が期待できる。

参考文献

- 1) Tobias H, Enrico C. Risks and benefits of speech recognition for clinical documentation: a systematic review. *Journal of the American Medical Informatics Association*. 2016;23:169-179.
- 2) 市川珠紀, 小泉淳, 高原太郎, ら. 放射線科レポート作成における音声認識システムの有用性: ディクテーション経験による比較. *日本医会誌*. 2005;65(4):384-386.
- 3) 小宮山恭弘, 水野成人, 花阪智, ら. 音声認識ツールを用いた超音波領域でのレポート作成の有用性について. *医療情報学*. 2005;25(1):29-35.
- 4) 杉山裕章, 今井靖, 鈴木健樹, 永井良三. ホルター心電図判読レポート作成における音声認識システムの有用性—音声認識ソフトウェアAmiVoiceを用いた検討—. *心電図*. 2011;31(2):158-164.
- 5) 丸上 輝剛, 谷 昇子, 松田 淳子, ら. 音声認識技術の電子看護記録への応用に関する基礎的研究. 第 28 回医療情報学会連合大会論文集. 2008;231-238.
- 6) 漆原譲治, 村上恵子, 伊藤正満, 鈴木丈一郎. 歯科検査における音声認識システムの有効性について. 第 28 回医療情報学会連合大会論文集. 2009;194-195.
- 7) 芦田廣, 脇坂仁, 鎌田志保, ら. 音声認識トリアージデータ入力システムの試作と評価. 第 27 回医療情報学会連合大会論文集. 2007;953-955.
- 8) 平井優司, 長井美和, 横井英人, 澤田秀之. オントロジーに対応した電子カルテ入力支援システムと医療現場での音声入力実験. *医療情報学*. 2012;32(2):73-81.
- 9) 島田優子. インサイト RPA ツールの全貌: 「開発者不要」で導入可能. *日経 systems*. 2017;2017-11(297):46-51.
- 10) Cloud Speech-to-Text - 音声認識 | Google Cloud. Google, 2018. [https://cloud.google.com/speech-to-text/ (cited 2018-Aug-23)]
- 11) Bing Speech API - 音声認識ソフトウェア | Microsoft Azure. Microsoft, 2018.

- [<https://azure.microsoft.com/ja-jp/services/cognitive-services/speech/> (cited 2018-Aug-23)]
- 12) Language Understanding (LUIS) | Microsoft Azure. Microsoft, 2018.
[<https://azure.microsoft.com/ja-jp/services/cognitive-services/language-understanding-intelligent-service/> (cited 2018-Aug-23)]
- 13) ユーザー補助 | Microsoft Docs. Microsoft, 2017.
[<https://docs.microsoft.com/ja-jp/dotnet/framework/ui-automation/index> (cited 2018-Aug-23)]

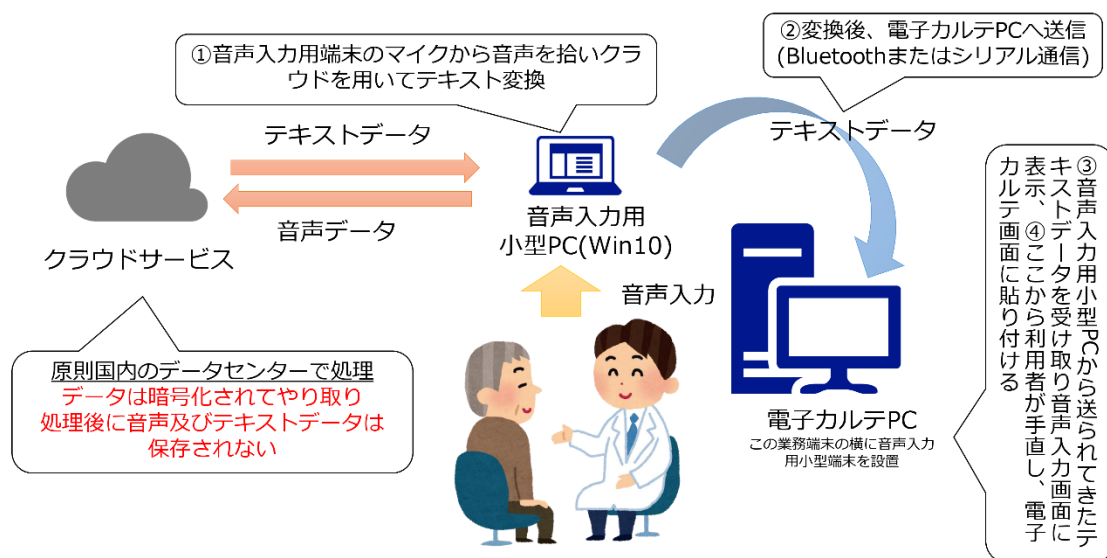


図1 音声入力システムの全体構成



図2 音声入力システムの使用機器