

# ICT技術を活用した企業業務補助システムの開発 (音声入力型在庫管理システムの検討)

新型コロナウイルス感染症の影響により、接触の機会をできる限り減らしたニューノーマル社会に対応したインフラや機器の需要が高まって来ています。そこでは最新の画像処理を援用した技術などをはじめとした様々なICT技術が有望視されていますが、スマートスピーカーに代表される音声によるICT技術も重要です。本研究では、ものづくりの現場で容易に導入できるよう、廉価で簡便な音声を使った入力システム開発の可能性を検証するべく、在庫管理を念頭に音声入力を擁するシステムの構築を試みました。

## システム構築について

音声の入出システムを検討する上では、どのような音声処理システム(エンジン)を使うかが肝要です。本研究では名古屋工業大学のチームにより現在、開発・管理が進められている汎用大語彙連続音声認識エンジン「Julius」を使用することとしました。こちらのシステムは、①Linux / Windows / MacOSなど、対応するプラットフォームの幅が広い。②エンジン起動後は音声データを入力するだけで解析結果が出力される。など、汎用性が高く簡便に解析することが可能、という特徴があります。また、商用利用においても制限がなく、利用に際しての障壁が非常に低いエンジンです。

## 「Julius」の性能評価

この「Julius」の音声認識の精度を確かめるべく、簡易な調査を行いました。1つは「音声認識パッケージ ディクテーション(自動口述筆記)キット」という任意発話を日本語に変換するキットを、もう1つは「記述文法音声認識実行キット」という事前に設定した単語・文法に従い、日本語を認識するキットを用いました。

表1と2が各フレーズを100回試行した結果です。なお2回目(言い直し)の結果は、1度認識に失敗した後に、明瞭に発話するように心がけた場合を抽出した結果です。表1のとおり、この度は「ディクテーションキット」を用いた場合、言い直しを行ってもそれほど認識率の向上が認められず、また、十分な認識を得られませんでした。表2のとおり「記述文法音声認識実行キット」を用いた場合は、十分な認識結果を得ることができ、音声入出力に利用できると考えられました。

表1 ディクテーションキットの結果

| フレーズ         | 1回目   | 2回目<br>(言い直し) |
|--------------|-------|---------------|
| 「モード変更状態へ遷移」 | 35.0% | 33.8%         |
| 「担当 小山」      | 11.0% | 11.2%         |

表2 記述文法音声認識実行キットの結果

| フレーズ         | 1回目   | 2回目<br>(言い直し) |
|--------------|-------|---------------|
| 「モード変更状態へ遷移」 | 100%  | N/A           |
| 「担当 小山」      | 95.0% | 100%          |

## 外乱音への対策

前述のとおり、「Julius」の「記述文法音声認識実行キット」を用いると、システムの実現可能性がみえたところでしたが、他方で、意図しない外乱音にまで反応してしまう動作が認められました。このため「発話者が意図して発話していることを何らかの方法でトリガする」ことが必要となりました。

そこで本研究では①音声の到来方向が所定の方向であることをスマートスピーカー基板により認識する。②カメラを援用し、発話者がシステムに正対している(顔認識)をトリガとして音声認識するようにしました。

## 作成したシステム

この度作成したシステムの使用風景と外観を下図に示します。プログラミングには「Python」を、ユーザーインターフェースには「Webブラウザ」を、データベース管理には関係データベース管理システム(RDBMS)である「SQLite」を利用しました。

このように、システムに発話者が正対して利用することで、音声入力を実現するシステムを構築しました。



図 システムの使用風景と外観(インセット)

## まとめ

本研究では、音声認識エンジンの「Julius」を用いて、在庫管理を念頭に音声によって入出力を行うシステムを作成しました。この度のシステムは在庫管理の他、オフハンドで音声だけで入力することが望まれる現場に横展開することが可能です。興味のある方は電気通信係までお問い合わせください。

※本研究では「汎用大語彙連続音声認識エンジンJulius」を使用しています。(https://julius.osdn.jp/)