



物流ニュース

NO. 114

2014年4月

物流現場で注目される音声認識ソリューション

● はじめに

近年、スマートフォン（スマホ）が急速に普及している。総務省の平成24年通信利用動向調査（世帯編）によれば、わが国におけるスマホの保有率は49.5%であり、前年より20.2%の大幅な増加となった。スマホは指で押す物理キーが無いため、初心者が文字を入力するには、多少のコツがいる。そこで、音声入力を用いれば、繊細な手元操作が不要であるため、面倒な長文の打ち込みが回避できる。今日では、様々な無料の音声検索アプリが提供されているので、ネット検索などで日常的に音声入力を利用している人は、少なくないであろう。

音声認識技術が利用される分野に着目すると、スマホなどの個人向けアプリにとどまらず、物流分野における音声認識技術の活用が、特にこの数年において注目されている。

本稿では、音声認識技術について、定義、歴史、仕組み、企業における普及状況、物流現場への導入事例および物流分野で注目される理由を紹介して、最後にまとめを示す。

● 音声認識技術の定義、歴史および仕組み

はじめに音声認識の定義を確認する。本稿では、音声認識は、「人間が話す言葉をコンピュータが分析して、話者が発した内容を理解する技術」と定義する。

次に、音声認識の歴史を振り返る。音声認識技術は長い歴史があり、米国のベル研究所が研究を開始したのは、1950年代の初頭である。音声認識の歴史を整理すると、第一世代は単語の認識、第二世代は文章の認識、第三世代は不特定話者と自由発話文への対応、第四世代はクラウドサービスの登場と、4つのステップで進化してきた。ここ半世紀以上に及ぶ研究成果の積み重ねで応用研究が進み、商品の実用化が実現した。

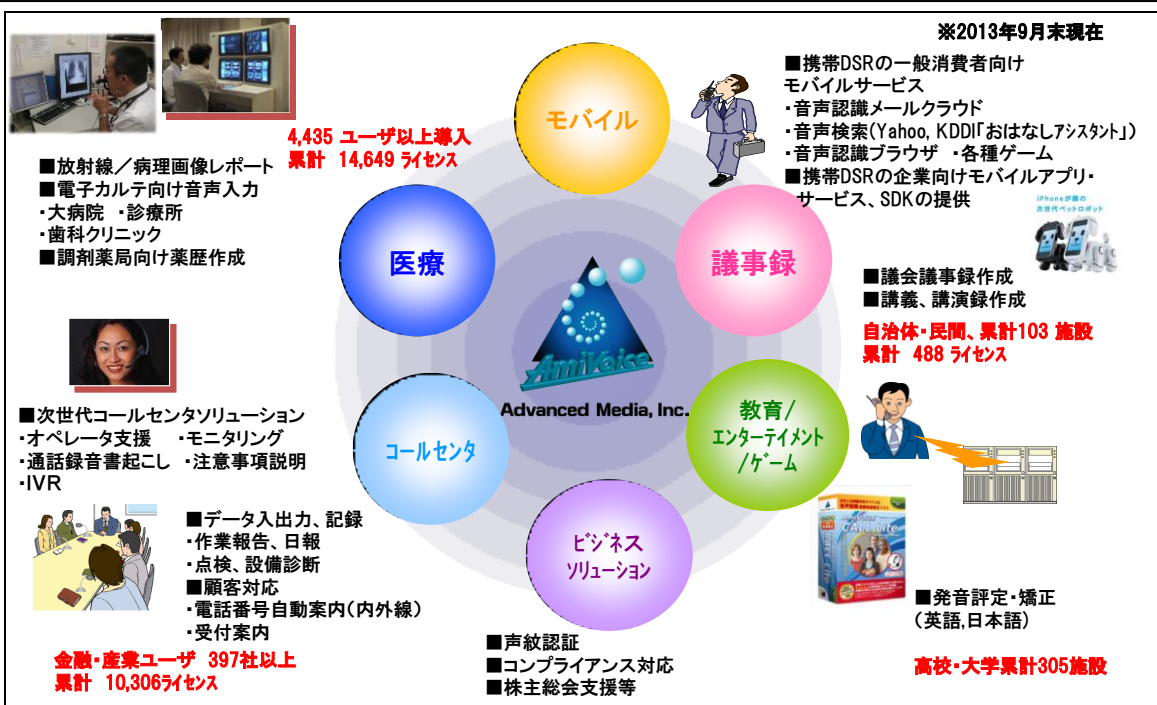
さて、人の声が言語として認識される仕組みについて、簡単に紹介する。人が発すると、音声は波形として認識される。音声認識の代表的な2つの特徴として、音響モデルと言語モデルがある。音響モデルとは音の波形データを集めたものであり、言語モデルとは単語とその並び方の情報を集めたデータである。この2つのモデルを用いて、声は何を言っているかを統計的に推定することで、文字に変換する。

● 企業における音声認識の普及状況

ここでは、業務用の分野において、音声認識技術の普及状況を確認する。

図表1に、音声認識の事業領域と導入数を示す。これは、音声認識ソリューションの企画・設計・開発等を行う㈱アドバンスト・メディアの導入実績数である。音声認識技術は、主に医療機関、民間企業、自治体などで既に実用化されている。

特に医師がカルテを作成する分野において、導入が加速している。例えば、実際の医療機関では、音声入力システムとして利用されている。患者を診察しながら医師の声を文字起こして電子カルテに記録することで、医師の入力業務が効率化される。



図表1 音声認識の事業領域と導入数 (出所：(株)アドバンスト・メディア提供)

● 物流分野への導入事例

音声認識技術を物流現場へ導入した事例として、銀座コージーコーナーの実例を紹介する。銀座コージーコーナーは全国におよそ 400 店舗を展開しており、人気のジャンボシュークリームなどが店舗の商品ディスプレイを華やかに彩っている。



図表2 ハンディ端末を用いた目視検品作業 (出所：(株)アドバンスト・メディアHPより)

まず、音声認識技術を用いた仕分け作業を行なう前の事前準備として、ハンディ端末画面を用いた目視検品作業の様子を図表 2 に示す。ハンディ端末画面には、ケーキの写真が表示されている。私たちが店舗で購入する生菓子は、見た目の美しさを追求するため、商

品名およびバーコードが記載されていないものが多い。そのため、商品知識が乏しい新人の担当者は、すべての商品を正確に目視確認することが困難である。その対策として、実物と画面を見比べることで、より正確に商品検品を行うことができる。



図表 3 音声認識による仕分け作業（出所：㈱アドバンスト・メディアHPより）

図表 3 に、音声認識による仕分け作業を示す。作業担当者は、商品を仕分けるロケーションおよび商品の出荷数量をヘッドセットマイクにより耳で聞いて、作業指示内容を理解する。作業終了後、担当者は、“次”というだけでよい。以下、これを繰り返す。この事例では、ある店舗の仕分け作業が完了した後、直ちに次の店舗の指示のタイミングを伝えるために、音声認識が利用されている。

銀座コージーコーナの事例において、仕分け作業に音声認識技術を採用した狙いを 3 つ挙げる。

第一に、作業生産性および作業品質の改善である。仕分け作業担当者は、両手を自由に使えるため、端末の出し入れや画面確認などは不要であり、効率的な作業が実現できる。また、やわらかいケーキなどは、両手を用いて丁寧に扱うことで、商品ダメージが低減される。

第二に、出荷作業精度の向上である。音声認識技術を用いた出荷精度を高める仕組みについて、以下に解説する。日々の事前準備として、現場管理者は、毎日変更される 4 桁のランダムな番号から構成される商品ラックコードを印刷して、店舗別のロールボックスに張り替える。商品ラックコードが日次で変更する理由は、作業担当者が番号を覚えてしまい、現物を目視確認せずに記憶している番号を読み上げることを防ぐためである。つまり、あえて毎日異なる番号を使うことで、作業担当者に常に緊張感を持たせて、正確な作業を行うことが狙いである。このように作業の正確性を保つために利用される文字列をチェックデジットと言う。本件では、チェックデジットを用いて、目視で確認し、番号を読み上げ、音声認識で正しい番号か確認する一連の工程により、仕分け作業精度の向上を実現している。

第三に、非接触による衛生面における配慮である。仕分け作業中は、ハンディ端末や物流機器に触れることがないため、極めて衛生的である。商品特性上、仕分け作業時には、

商品以外の物にできるだけ触れないことが望ましい。

● 音声認識技術が注目される理由

今日において、物流分野で音声認識が注目されてきた3つの理由をまとめる。

第一に、ソフトウェア・ハードウェアの性能の向上である。音声認識の研究は進み、音声認識エンジンは、産業用分野で導入事例が加速して、劇的に進化した。例えば、コンベアの駆動音やフォークリフトの走行音や衝撃音が発生する騒々しい物流現場においても、音声認識の精度は、実務で活用できるまでに高められた。

また、近年、端末が高性能化しただけでなく、通信環境が充実してきた。従来、物流センター内における無線通信はPHSなどが用いられたため、大容量のデータ送受信には、ある程度の時間が必要であった。現在は自宅でも無線LANが簡単に構築でき、大容量の光回線も手軽に利用できる。このように、国内における通信環境は高速化し、使い勝手がよい環境が整った。この良好な通信環境を用いることで、物流施設に設置されるコンピュータの演算量は、非常に少なくなった。そして、音声認識処理は遠隔地の高性能サーバーにて超高速で解析され、直ちにその結果を物流現場へ返すことが可能となった。

第二に、システムの低価格化が挙げられる。以前は、音声認識技術は、顧客別にシステムを構築するため、システムの開発費だけで数百万円は必要であった。現在では、製品のパッケージ化により低価格で導入でき、比較的小規模な物流現場でも投資対効果が期待できるようになった。

第三に、音声認識の魅力が認知されたことである。CMなどで音声認識の魅力が広まり、いわゆる“秘書アプリ”などの使い勝手が高まり、スマホを通じて、私たちが音声認識に触れる機会が増加した。そこで、音声認識が日常的なツールとして使われはじめ、業務向けの利用を考える企業が増えたことが挙げられる。

● まとめ

本稿では、音声認識技術に関して、定義、歴史、仕組み、企業における普及状況、導入事例および物流分野において注目される理由を確認した。大学や民間企業の研究機関を中心に半世紀以上に渡り研究されて、音声認識技術は飛躍的に進化した。スマホのアプリのような個人利用だけでなく、産業界においてその導入が加速しており、物流分野において注目されている。

将来的に、音声認識システムは、端末の軽量小型化などにおいて、これまで以上の技術革新が起きるかもしれない。それが作業担当者にとって使い勝手が良いものであれば、現場で利用する物流機器の選択肢が増えるため望ましいことであり、物流関係者は今後の動向を注視する必要がある。

KEY WORD

無線LANとは

無線通信を利用した特定地域内通信網 (local area network)。パソコンなどをケーブルなしでLANやインターネットに接続することが可能であるため、その手軽さから事務所内だけでなく、一般家庭においても普及が進んでいる。物流分野ではフォークリフトやハンディターミナルに無線LAN機能を持たせ、ピッキングデータの送受信などにも利用されている。

(出所：ロジスティクス用語辞典。日通総合研究所 [編])

—日通総合研究所 ロジスティクス コンサルティング部—