

今回の覚えておきたい
このまちのサイエンス NOW !

「音声認識学」とは？

「音」から見える化
つづきに、楽に、
豊かに

私の研究は「音中心のセンサー情報による行動・心身状態理解」です。学問領域としては人工知能系の音声認識分野に位置しています。音は画像等に比べると小さな一次元のデータですが、言語情報はもちろん、文字だけでは伝わらない感情に加え、話し方の癖、性別、年齢といった多様な情報を含んでいます。このような音情報を集積、分析することで個人情報を選定することもできます。さらに、最近では心的状態に関する情報も抽出できることがわかってきており、それを応用した認知症

サイエンス全史

明日から見えてる世界がハイレゾ化しちゃう！
静岡県でいま研究されている先端サイエンスの紙上講義

文・成田真也（静岡大学人文社会科学部三年）
写真・林優介（静岡大学人文社会科学部三年）



音は思っている以上に
物を言います

西村雅史先生

静岡県立大学情報科学部教授

喉嚨マイクの実物を用意していただきました。
皆が首にマイクを巻いて歩いてくれる日が
普通になることを願っているそう。

の診断やストレス測定の研究も行っています。

私が音という分野で研究し始めたのは大学四年生からで、卒業後も同じ分野で研究を続けたと思います。当時音声認識の分野で多くの成果を挙げていた研究室の研究所に入りました。

そこでは三十年に渡って大規模な音声認識の基礎研究を行うことができました。大学に來たのは今から七年前と最近のことで、最初は予算や環境の違いに驚きました。ただ、ビジネス的理想が付き纏う研究所時代と比較すると自由度はかなり高く、音声認識の基礎研究からそれらを応用した技術の開発という研究にシフトしました。

私が現在取り組んでいる主なプロジェクトは、高齢者の健康維持・管理を目的として「食べる」「話す」という行為に着目したものです。その中でもまず「食べる」行為に重きをおいた研究が「心身状態の認識・見守りシステム」です。この研究は高齢者にマイクをつけて過こしてもらい、その行動や音声をモニタリングするというもので

す。ここでは「咽喉マイク」と

いう皮膚に直接押し当てる特殊なマイクを使用します。このマイクは声を発した時の皮膚の振動を拾い、音に変換する原理なのでノイズに強く、装着負荷も小さいデバイスです。このマイクを使用して「発話」だけでなく、今は左で囁んでいる、今飲み込んだというような咀嚼から嚥下（飲み込む）の過程、「むせ」などの音情報を収集します。

マイクで集めた音情報からは、咀嚼の回数や左右のバランス、嚥下の頻度や持続時間というような装着者が無意識で行う食行動を抽出することができます。このような食行動の認識は、日常の健康維持・管理だけでなく、表記の早期発見や診断の手助けにもなります。嚥下を例にすると、高齢者は筋肉の衰えから、若者と比較して嚥下（飲み込む）の動作が遅くなります。そうすると、誤って気管に水や食べ物が入り込む「誤嚥」が起りやすくなります。「誤嚥」は重症化すると誤嚥性肺炎を引き起こし、ひどい時には死に至ることも多い病気です。

「誤嚥」の兆候を探るためには喉の動きを見る必要がありますが、現在ではX線を当てるかマイクロスコープを鼻から入れなければなりません。このような診断方法は、費用や労力という面で患者にかなりの負担をかけます。しかし、この「誤嚥」を音情報で解析できれば、患者への負担を軽減すると共に、容易に正確な診断が可能になります。音情報による解析は、衰え切ってしまう前に兆候をつかみ、後々の予防・診断に役立てることができるのです。

他にも、「話すこと」で重きを置いてるのが「回想法音声対話システム」です。こちらは、介護施設に設置することを想定して、ロボットから話題を提示して、話し相手となる高齢者にたくさん話をしてもらおうというものです。発話は脳を活性化し、認知機能の維持に大変役立ちます。

つことがわかっています。

また「話す」という行為には、当人の心理状態や認知能力が音になって現れます。それには言い淀み、声の裏返り、相槌の頻度など様々な特徴があります。

例えば、認知症患者には会話のタイミングが測れない特徴があります。普通の人は相手が喋り終わるタイミングが態度や息遣いから予測できますが、認知症患者はそれが難しくなります。

そのため、相手が喋り終わる前に話出してしまう「かぶり」が発生します。認知能力が低下すると、対話においての先読み能力が衰えてしまうのです。この他にも認知機能が低下すると声自体の特徴の変化に加え、使用する語彙が少なくなったり、代名詞が多くなったりといった変化も現れます。健常者にはみられないこのような様々な変化

高精度な音声認識で、グ レードアップした生活を



たくさん話すことは認知症
の予防につながるんです



を捉えることで認知症やその前段階の早期発見につなげられることもわかってきています。

ただ、まだまだこれらの研究にも大きな課題があります。二つのプロジェクトに共通しているのが、「いかに違和感をなくすか」というものです。現段階では高い精度を誇る食行動の音認識ですが、間違いを犯すこともあります。例えば、実際は左で囁んでいたのに、システムは右で囁んでいると誤った判断を下すことが多々あります。また、対話システムにおいては「頷き」の問題があります。人間は会話を続けるためにタイミングを図って「頷き」を行うのですが、ロボットはそれができません。なので、的外れな場所でも頷いたりしてしまいます。日常的に扱うことを目的としたシステムと認識のズレが起こると、使用者に不快感を与えてしまい、続かなくなってしまうのです。この認識のズレを失くすためには、日々研究を進めています。

音声認識は目標の達成が難しい分野です。そのため、目標設定が重要になってきます。集め

音情報から人間の全てを
理解するのが野望です



「よく囁まないで飲み込むくせ、直そうと思います」
文・成田



「最近独り言が減りました。老いかもしれません」
写真・林