

音声認識を活用した浄化槽維持管理システムの 取組みについて

株式会社 HHC ○東 晃一
齋藤 英俊

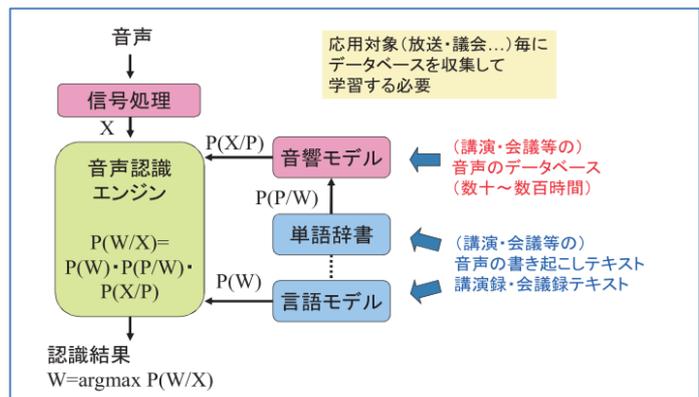
1. はじめに

近年、コンピュータによる情報処理システムの活用が広く浸透してきており、浄化槽の業界でも維持管理の業務を現場で入力できる情報処理システムが、数多く開発され運用されるようになってきた。

しかしながら、電子機器による入力作業は、フィールド作業における運用を考慮しているとは言い難く、だれにでも簡単に入力出来るインターフェイスにはなっていない。情報機器などの操作に不慣れな初心者は、システム操作以前にパソコンやタブレットの機器操作の習得に一定の期間が掛かるのが現状である。また、浄化槽維持管理の作業現場では汚水・洗浄水などの水分が多く使われ、濡れた手などでの操作は機器の故障を早める結果にもつながってくる。このように、フィールド作業に置いては、誰もが容易に操作できるインターフェイスの活用が今後さらに求められると考えられ、これまで主として用いられてきたキーボード、マウス、タッチパネルなどの入出力作業の負担を取り除くディクレーション（口述筆記）ソフトや装置の開発が期待される¹⁾。

音声は人間がコミュニケーションを行うためのもっとも基本で有効な手段の一つであることから、音声をコンピュータで処理する音声認識/合成などの技術が、近年飛躍的に発展してきた。これまでの入出力装置に置き換わる手段として音声認識技術が、浄化槽維持管理の現場作業で活用できないかさまざまな条件下で比較検証してみた。

図-1 音声認識のしくみ²⁾



2. 実験方法

(1) 実験の概要

本実験は、現在実用化され広く認知されている音声認識アプリケーションが、実際の浄化槽維持管理現場でどの程度利用可能かを検証するものであり、一般に市販されている情報機器を使用して検証した。また、音声認識技術を浄化槽設置先で用いるためには、騒音レベルが90dBを超えるような高雑音下での運用も視野にいれなければならない。音声を入力するマイクロフォン、音声認識を行う情報端末、さらに作業場所での環境雑音のレベルを想定し実験を実施した。

(2) 実験装置

実験装置の構成は、入力装置として内臓マイクロフォンと有線ヘッドセット、Bluetooth ワイヤレスマイクを用いた。(表-1) また、情報装置として携帯電話の iPhone と Android 端末、Windows タブレットパソコンを用いた。(表-2)

表-1 入力装置

有線マイク	ヘッドセット E 社製 HS-HP22TBK	周波数帯域:100~10,000Hz/感度:-58±3dB/インピーダンス 2.2kΩ/無指向性
無線マイク	Bluetooth ヘッドセット i 社製 BSHSBE33	周波数帯域:100~10,000Hz/感度:-58±3dB/ノイズキャンセル機能搭載(CVC)/無指向性

表-2 情報端末

名称	モデル	OS	音声認識システム
iPhone 携帯	iPhone 6 Plus	IOS Ver10.3.3	IOS 標準
Android 携帯	Motorola Moto G4 Plus	Android Ver7.0	Google アシスタント
パソコン	Lenovo Miix 3	Windows10	Cortana(コルタナ)

(3) 実験環境と方法

測定箇所として、以下の 3 か所を比較する。

- ①騒音の影響をほぼ受けない室内 (20~30dB 程度)
- ②一般家庭設置先を想定した低雑音下環境 (40~50dB 程度)
- ③大型ブローヤやポンプ設備などが設置されている大型浄化槽設備を有した高雑音下環境 (70~100 dB 程度) を想定した。

雑音のスペクトルが均一でなく、且つ時間的に変化する非定常雑音の環境下では計測データにばらつきが発生する。データを平均化する為、各々の計測地で 2 分間の環境雑音を録音した雑音データを実験のベースとし、他の環境雑音の影響を受けない状態の室内で各計測地を想定しながら測定を行った。(図-2)

写真-1 静寂地(40dB 程度)



写真-2 喧噪地(70dB 以上)



(4)
音声

データの構成

実験する音声データとして、浄化槽点検時に用いられる点検項目を想定した単語認識用データと、点検後の社内向け連絡メモを想定した文章認識用データを用いた。(表-3)

表-3 音声テストデータ

単語	pH:7.2/水温:28℃/透視度:20 cm/残留塩素:1.2/DO:4.5/亜硝酸濃度:+/塩素剤:450g 汚泥厚:20 cm/スカム厚:15 cm/タイマー時刻:2 時 10 分/逆洗装置 1 回 5 分
文章	8 月 25 日 11 時 25 分浄化槽コード 1024 山田太郎様にて浄化槽点検時に槽内の漏水を調査し亀裂・破損個所を確認しました。お客様とは後日見積り提示の約束をしました。次回清掃時に修理予定です。

(5) 実験設定

今回実験に使用した音声入力部分のシステムモジュールは、実際の浄化槽現場で運用されている弊社の維持管理システム ECOPRO をベースに試作したものである。入力された音声データがそれぞれの上表端末における音声認識システムを經由して ECOPRO システムに取り込まれるようにプログラム修正を行った。実験用に音声認識データとして設定した部分は、点検項目の数値項目部分（保守点検項目 12 項目）と社内向けメモ用コメント項目（清掃依頼処理の社内コメント項目）に限定し、他の点検項目「良否」判定部分や、設置先への告知文・備考に関しては、実験対象から除いている。各々の雑音環境下において上記の音声テストデータを入力装置から取り込み、音声認識システムが文字変換処理したデータの認識率を評価した。

図-2 実験環境



3. 実験結果と考察

実験は単語、文章の 2 種類の音声データを入力装置で取り込み、音声認識システムで文字に変換する。その際、環境雑音のレベルを 3 段階で想定し（無、低、高）、音声データに雑音データを加算して処理を行った。実験の要素は音声データ 2 個、入力装置 3 台、音声認識システム 3 個、加算する環境雑音データ 3 個とし、それぞれ 10 回ずつ実行し認識率の平均値を算出した。

表-4 認識率

	機種	音声認識	騒音低 (30~40dB)			騒音中 (50~60dB)			騒音高 (90~100dB)		
			Bl	有線	内臓	Bl	有線	内臓	Bl	有線	内臓
単語	Android	G.Asst.	79.1	81.7	79.1	78.3	76.4	78.8	69.3	74.4	68.1
	iPhone	IOS 標準	24.3	80.4	74.6	21.8	84.5	73.5	19.3	76.8	78.3
	Windows	Cortana	NG	27.8	9.6	NG	21.8	NG	NG	18.3	NG
文章	Android	G.Asst.	92.5	97.5	97.0	94.5	95.5	75.8	90.5	92.5	79.1
	iPhone	IOS 標準	26.3	99.5	98.7	19.3	96.1	94.5	16.1	90.8	80.0
	Windows	Cortana	NG	41.0	1.5	NG	35.0	NG	NG	23.5	2.1

※G.Asst.: Google アシスタント、Bl: Bluetooth (無線通信規格)、NG: 計測不可能の略

認識率の結果を表-4 認識率に示す。まず単語データの認識率は、予想していた値より低い結果となった。特に windows における変換率は低く入力装置によっては全く違う変換結果となり、実運用対象に程遠い低い認識率となった。

環境別では、雑音環境を考慮しない室内：72%、低雑音下環境：68%、高雑音下環境：62%となり、認識率は騒音の大きさに左右される。但し、同環境下でも入力装置が有線マイクの場合は、マイクと口の距離が近いためか、高雑音下環境でも認識率は維持された。機種別では Android：83%、iPhone：64%、Windows：10%となった。音声データ別では単語認識：59%、文章認識：69%となった。音声データの場合、個別比較しても認識率90%以上の値は全て文章側に存在し、単語側は誤認識が大きい。これは点検項目名+数値の発音の音素が音声認識の音響モデルや発音辞書にない識別的言語パターンを形成しているためであり³⁾、文章認識のように連続して入力される単語の前後を比較判断しながら変換作業を繰り返すことが無く誤認識が大きくなったと思われる⁴⁾。また今回、特徴的だったのは iPhone と無線マイクの組合せで認識率が低かったことである。有線マイクや内蔵マイクの認識率が高いため、機器固有に帰する事象かどうか、今後複数の無線マイクと組合せ相性を確認する必要がある。

4. まとめと今後の課題

音声認識を現行の維持管理システムに取り込むことで入力作業の一部分を処理することが確認できた。特に文章認識は高い認識率を示し、現場における入力作業を大きく改善させることが可能と考えられる。単語認識は認識率が低い結果に終わってしまったが数値音声部分だけ抽出すれば95%以上の認識率を上げており、十分にシステム活用が検討できる。項目名称をコード化し「1番 28℃」のように数値のみで水質入力を行うことが可能である。または音声合成と組み合わせて点検項目を音声合成でシステム側が発声し、それに応答する形で数値部分を音声認識で処理するなどの処理が検討できる。

今後本格的に音声認識をシステム化していくには次のような課題が考えられる。

- (1) 誤変換された音声データの修正作業をどのように行うか
- (2) 数値項目以外の点検項目「良否」判定部分をどのように処理するか
- (3) 現場作業に支障のない入力装置の形状は何か
- (4) 情報端末のコマンド操作は音声処理可能か

全ての作業を音声で処理するにはまだ技術的な課題が多い。現行のシステムで文字変換処理を行うなどある程度のシステム操作習得が必要な部分を音声認識処理に変えて運用するなど部分的な活用を検討すべきと考える。

参考文献

- 1) 独立行政法人工業所有権情報・研修館：音声認識技術(2005年)
- 2) 河原達也：話し言葉の音声認識の進展-議会の会議録作成から講演・講義の字幕付与へ-(2012年)
- 3) 鈴木雅之：背景雑音と話者の違いに頑健な音声認識(2013年)
- 4) 西崎博光：音声文章を対象とした音声入力型情報検索システムに関する研究(2003年)