

音声認識の応用事例からみた入力確認方式

3 J - 6

小川 博 (園田計器工業) 船越 洋二 (園田計器工業)
 和田 秀夫 (関西電力) 上西 勝也 (園田計器工業)

1. はじめに

特定話者による音声認識の応用として電力量計指示値の読み取り作業への適用システム [1] を開発した。電力の取引に使用される計器を検定作業等で一度に多数の指示値を目視と発声により読み取るが音声認識では読み取り誤りは避けられないため入力の確認と修正を行う方式を検討し従来読み合わせでの確認に替り表示装置による目視と音声応答の併用の方式を採用した。

2. 応用事例「電力量計指示値読み取り」

今回の開発は電力量計 (Watthour Meter) の検定作業等においてデータを音声で入力し認識結果を画面表示や音声応答で確認して蓄積し既存機器に入力データを転送するシステムであり作業フローを図1に示す。20~40台の電力量計の固有番号、前計量値を読み取り、試験終了後に後計量値を読み取り前後差と標準値の比較により電力量計の良否を判定する。この装置により従来は1人が読み上げ他の1人がキーボードで入力し2人で行っていた作業が1人で可能になる。音声認識装置は(1)誤認識が少ない(2)利用者が限定されている(3)コストの点からV社 [2] の既製品である特定話者方式の装置を採用した。

3. 音声の認識率

テストは5人の作業者が3または4桁の数字データと「計器番号」「前計量」「後計量」などの単語5種類につき合計14276語の入力を行った。テストに先立ち個人の音声登録を各人30分程度かけ

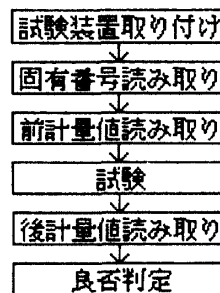


図1 作業フロー

行った。各作業者の認識率は平均95.8%標準偏差1.34%であった。認識率Reは式1のとおり全入力語数Twから誤認識語数Ewおよび無認識語数Nwを除いた語数の割合と定義する。

$$Re = \{ Tw - (Ew + Nw) \} / Tw \quad \text{式1}$$

数字データの入力について考察すると、誤認識 (例えば1を2、3を8、0を8に認識する) について90通りの組合せが考えられるが各人について3通りの組合せの誤り合計件数が全誤り件数の61.6%であった。誤認識の上位3通りの合計の全体に占める割合を図2に示す。また、最も誤認識の組合せ

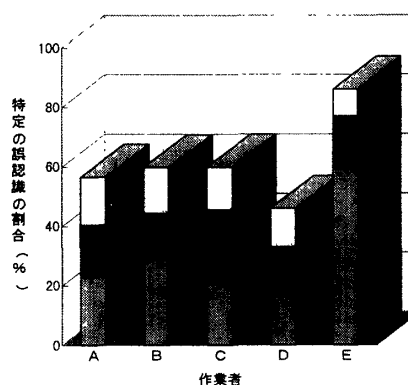


図2 誤認識上位3通り合計の割合

の多い作業者で25通りの組合せであり、全90通り中の30%に満たない種類の誤認識である。そこで標準パターンの追加学習を各作業者にを行い改善し

A Verification System for Applied Speech Recognition Processing
 Hiroshi Ogawa(1)*, Yoji Funakoshi(1)
 Hideo Wada(2), Katsuya Uenishi(1)
 (1)Sonoda Instruments (2)Kansai Electric Power
 * e-mail:ogawa@sonoda.co.jp

ようとしたが誤認識の結果は変わらなかった。このことから学習効果には過度に期待できないことが判明した。

4. 確認修正システム

100%の認識率を得られない音声入力では誤認識に対処するものが必要となるが、これは音声入力の都度、入力データが正しいことを確認し正しければ次の入力に移り、誤っていれば訂正、再入力を行うものである。このシステムに求められることは、入力の確認を素早く行い、訂正、再入力 that 確実に行うことである。[3] 今回の装置では、入力の確認のために認識結果を表示パネルに表示させた装置にスピーカを取り付けてヘッドフォンと共に音声応答させて入力の確認が素早く確実に行うようにした。図3に2つの入力確認方法について示す。

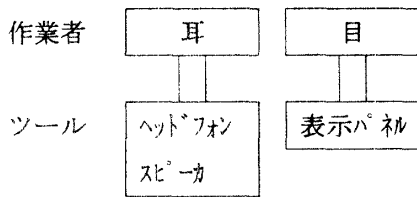


図3 入力確認方法

耳からの情報は人間の意識を向けなくても聞ける反面、聞き逃しや聞いても忘れてしまうことがあるが目からの情報は人間が意識して見なければならぬ反面、映像表示により確認できるという利点がある。このシステムでは入力の確認について音声応答と映像による表示を用いたがこれは音声応答のみの確認の時の聞き逃しや聞き間違いを起すのを回避するためである。また音声による再入力でも対応できない時にテンキーと数字以外のキーワードの入力として押しボタンスイッチによる入力できるようにした。図4に構成ブロック図を示す。

5. 音声入力による作業効率

最終的な入力結果についての誤りは皆無であった1語あたりの認識処理時間をRt、従来の1語あたりの入力時間をHt、リトライの合計回数をn、音

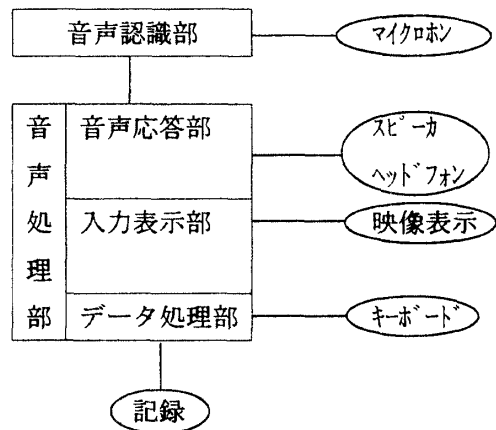


図4 音声認識システムブロック

声入力不能の語数をUwとすると、作業時間Wtは

$$Wt = \{TwRe+n(Ew+Nw)\}Rt+UwHt \quad \text{式2}$$

マンパワー比率Me (音声入力/従来方式)は

$$Me = Wt / (TwHt \times 2) \quad \text{式3}$$

となりMeは0.38程度となっている。

6. おわりに

今回試作した入力確認装置の音声応答は本装置採用以前の作業者2人での読み合わせと同雰囲気であるため現場では好評であった。音声認識装置は特定話者方式から不特定話者方式へ主流が移りつつあるが、その状況においても当然入力の確認は必要であり広く音声入力を普及させるためにも応用されるシステムに適した入力確認方法を確立していかなければならない。本研究では日本電気計器検定所尼崎試験所殿にご協力頂き深謝する次第である。

参考文献

[1] 上西：音声認識システムの現状、京都市工業試験所酒研会報1996.8
 [2] Verbex Voice Systems : Speech Commander Portable User's Guide1993
 [3] 嵯峨山：なぜ音声認識は使われないかどうすれば使われるか、情報処理学会音声言語情報処理研究会1994.5