

音声モニタリングによる誤操作防止機能の検討

瀬川 修 † 高橋 誠 ‡ 水本 綾 ‡

†中部電力株式会社 電力技術研究所 E-mail: sega@rd.chuden.co.jp

‡株式会社東洋情報システム E-mail: {takahasi,mizumoto}@ssten01.tis.co.jp

概要

電力システムの監視制御システムのオペレーションにおける指差確認呼称とよばれる計算機に直面した発話行為に着目し、オペレータの発声をモニタすることによって音声認識を誤操作防止目的に積極的に利用するインタフェースの提案を行なう。本稿では、誤操作防止のためのインタフェースを自然言語を解する“仲介システム”によるチェック機能としてモデル化し、全体の機能の一部である音声認識モジュールの実装について初期検討を行なったので報告する。

A Study on False Operation Prevention System using Speech Monitoring

Osamu Segawa †, Makoto Takahashi ‡, Aya Mizumoto ‡

†Chubu Electric Power, Co., Inc. E-mail: sega@rd.chuden.co.jp

‡Toyo Information Systems, Co., Inc. E-mail: {takahasi,mizumoto}@ssten01.tis.co.jp

Abstract

In operation of power control and information system, there is an utterance — “pointing and calling check” by an operator in front of computer systems. We propose a human-computer interface using speech recognition technology positively for the purpose of false operation prevention. This paper proposes a model of false check system by intelligent “interface agent” and describes initial implementation of the speech recognition module which is a part of total function.

1 はじめに

電力系統の監視制御の分野では、システムのインタフェース機能の重要性は今後も高まる一方であり、系統状態の迅速かつ確かな把握や、運転操作時の人為的ミスに対する安全性の向上が期待される。これらの課題に対し、従来はメッセージ出力による警報など一部の限定された用途に採り入れられてきた音声メディアをインタフェースとして積極的に利用することが考えられる。

音声認識技術を応用したインタフェースは、人間にとって普遍的な自然言語を計算機との対話手段として用いることから、システムやアプリケーション環境に依存したコマンド体系の習得を要しない。また画面のメニュー選択において視覚や直接操作によらない迅速な画面・機能呼び出しが可能であるなど、音声入力はヒューマン・インタフェースの実現方式として有効であるといえる。

一方、誤操作防止への応用という観点から音声メディアの利用を想定した場合、オペレータの音声の意味内容をモニタし、何らかのチェック機能を設けることが考えられる。本研究では、システムのオペレーションで義務づけられている「指差確認呼称」という計算機に直面した発話行為に着目しているが、オペレータと制御計算機の間で自然言語を解する“仲介システム”によるチェック機能が介在することにより、音声メディアを誤操作防止への応用として積極的に利用できる可能性がある。

これまで発電所をはじめとする大規模プラントの運転支援のためのヒューマン・インタフェースの研究が数多くなされており、音声メディアの利用例としては、音声出力による警報や、さらに従来の直接操作による入力デバイスの機能を音声コマンドによって代替するなどのインタフェースが提案されている[6],[7]。しかし、音声メディアを誤操作防止目的でシステム側のモニタリング機能として積極的に採り入れたものは少ない。本稿では、音声認識技術の電力系統の監視制御の分野への応用として、誤操作防止のための新しい概念のインタフェースを提案する。誤操作防止のためのインタフェースを知的な仲介システムによるチェック機能としてモデル化し、音声モニタリングを実現するための要素技術について初

期検討を行なったので報告する。以下、2では音声メディアを利用した誤操作防止機能のコンセプトを述べ、3では仲介システムのモデルの概要を示す。4ではシステム全体の中の音声認識モジュールの実現方式について述べ、最後に5でまとめと今後の検討課題を提示する。

2 音声モニタリングによる誤操作防止機能

提案するインタフェースは、給電制御所システムとよばれる電力系統の監視制御のためのシステムの操作端末(制御卓CRT)における使用を想定している。対象となるユーザは電力会社の社員であり業務に関して専門知識を持ったオペレータである。給電制御所システムの通常業務の主なもの、(1)電力系統の状態の監視と(2)電力設備の保守・運用に伴う現地設備の遠隔操作である。制御卓CRTには入力デバイスとしてポインティングデバイス(スタイラスペン)とボタンが備わっている。オペレータは通常の画面オペレーションではスタイラスペンあるいはボタンを用いる¹。一方、遠隔で変電所などの開閉器操作を行なう場合は誤操作を防ぐ目的から、運転操作に際しては保安上必ず指差確認呼称が求められる。ここで言う指差確認呼称とは、ある操作手順を実行する時にオペレータが操作手順の内容を復唱してから実操作を行ない、また実操作を行なった後に実行した操作内容を復唱して確認する発話行為を含んだ一連の動作シーケンスのことである(例えば、スイッチ No.xxx を切るという操作手順において「スイッチ No.xxx を切ります」と発声してから、次に実際にスイッチ No.xxx を切るという操作を行なう。そして操作が完了した後に「スイッチ No.xxx を切りました」と操作内容を復唱する)。

指差確認呼称の目的はヒューマンエラーの防止であるが、その行為が持つ認知的要因として、(1)自己フィードバックによるセルフチェック、(2)周囲の人の注意を喚起し、これから実行しようとする行為を知らせるなどの重要な意味がある。

そこで、本研究では業務の中で見られる制御卓CRTに直面したこの発話行為²に着目し、遠隔操作

¹音声コマンドによる画面オペレーションについても現在検討中である。

²社内規定に従って発声することから、義務的発声タスクであるといえる。

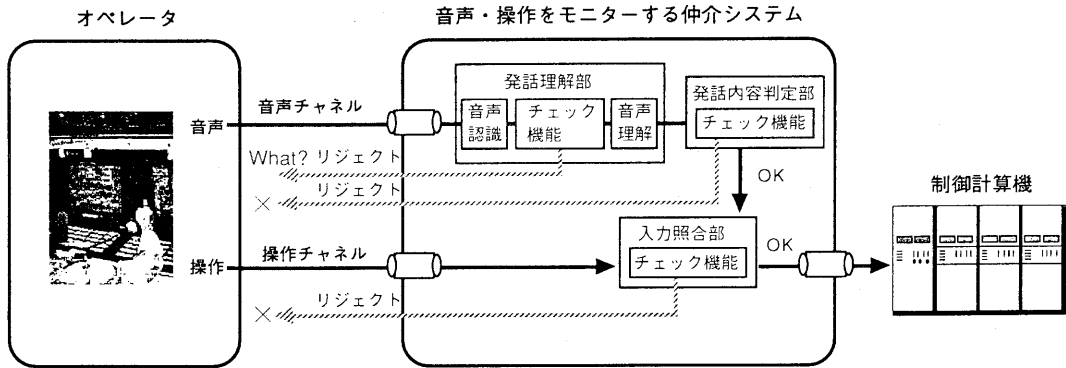


図 1: 仲介システムのモデル

のオペレーションが行なわれている間にシステムがオペレータの入力の挙動を逐次モニタすることによって発声の意味内容のチェックおよび音声と操作の照合を行い、誤操作を未然に防止する機能を設定した。次節では音声と操作をモニタする仲介システムのモデルの詳細を述べる。

3 仲介システムのモデル

図 1 に本研究で提案するモデルを示す。このモデルではオペレータと制御計算機の間知的な“仲介システム”が介在する。モデルではオペレータからの入力として音声と操作の 2 つのチャンネルを設けている。機器の遠隔操作など安全性が要求されるオペレーション時には、“仲介システム”がオペレータの入力（音声、操作）を逐次モニタすることによって誤操作防止のためのインタフェースとして機能する。

モデルは (1) 発話理解部、(2) 発話内容判定部、(3) 入力照合部の 3 つのモジュールより構成され、各モジュールごとにチェック機能を有している。誤操作防止機能では指差確認呼称の動作シーケンスに従い、この 2 つのチャンネルからの入力を個々に処理していく。

音声チャンネルにおいては、まず (1) 発話理解部において発話文の認識と理解が行なわれる。発話理解部はサブモジュールとして、さらに音声認識モジュール、音声認識の信頼性に基づくチェック機能、音声理解モジュールの 3 つの処理に分割される。音声認識モジュールは発話を認識する連続音声認識システム

である。フロントエンドとなる発話理解部では様々な環境要因のもとで発話が正しく認識・理解される必要があることから、簡易な対話制御方法として音声認識の信頼性に基づいた再入力促進 [1] のチェック機能を設ける。すなわち確率モデルによる音声認識のスコア（尤度）が、あらかじめ設定した閾値より低い場合は言い直しや言い淀みなど発声が不明瞭であったと判断し、入力音声はリジェクトされる。スコアが閾値より高い場合は認識結果は受理され、次の音声理解のサブモジュールに渡されて発話意図の抽出が行なわれる。

音声チャンネルではこの後に (2) 発話内容判定部として、発話の意味内容についてチェック機能を設ける。もしシステムが操作手順に関する知識を有していれば、発話内容が操作手順として適切であるかの判断をステップごとに行なうことが可能である。一方、発話内容のチェック機能として操作手順などの深い知識を用いるのではなく、オペレーションの状況下での浅い知識によって判定を行なう方式も考えられる。すなわち、操作の入力履歴やカレント画面の選択状態を知識として使い、発話理解部の結果と照合して矛盾がないかのチェックを行なうというものである。

最後に (3) 入力照合部では、音声チャンネルと操作チャンネルの入力照合によるチェック機能を設定した（発話内容と実際の操作内容が一致しているかを照合する）。これによって制御計算機への入力として適切であるかの最終的な判断を行なう。

以上がモデルの基本的な枠組であるが、全体のモジュールの中で(2)発話内容判定部の実現に関していくつかの問題点が残る。まず操作手順のような深い知識を前提とした場合、人間系が介在するマニュアル操作実行において、操作手順に関する完全な知識を登録しメンテナンスするのはあまり現実的ではない。また操作手順に関する完全な知識を前提とすると、操作チャンネルの方にチェック機能を設けるほうが実現が容易であり、音声チャンネルを用いる意味が薄れる。したがって、ここでは入力履歴やカレント画面の選択状態の知識などオペレーションの状況下での浅い知識を前提とした方式が実現性が高いと考えられる。

4 音声モニタリングに関する検討

図1のモデルでは音声チャンネルと操作チャンネルの2つのモードがあり、各モードごとに入力のモニタリングを行なう必要がある。以下、本稿では音声チャンネルの処理に相当する音声モニタリングに関する検討を行なう。音声モニタリングを実現するための中核となるのは音声認識モジュールである。ここでは、認識対象となる発話文の分析を行ない、音声認識処理としてのタスクを明確にする。

4.1 指差確認呼称の発話の特徴

指差確認呼称は実際の発話文の収集が可能であることから、業務分析のデータとして現場の給電制御所で発話コーパスを収集した(話者8名、218文)。以下に、発話文を書き起こしたコーパスの一部を示す。

えー、じゃあ、7万の名南寛政2号線停止します。 えー、名南寛政2号線、遮断器724、 えー、操作選択よし。 切ります。 えー、名南寛政2号線、遮断器724、切れました。
--

指差確認呼称の発話文は社内マニュアルで用語、発声内容が規定されていることから、もともと言語的な拘束性は強く、例えば「<線路名>を停止します」のように発話はほとんど定型パターンである。ただし、定型パターンの中に冗長語(間投詞など)、省略(格助詞、語尾の欠落)、言い淀み、言い直しなど音声言語に特有の曖昧性が多く見られる。表1に冗長語を対象としたコーパスの分析結果を示す。

表1: 冗長語を対象としたコーパスの分析結果

冗長語種類	出現箇所 (文頭)	出現箇所 (文節間)	小計
えー	114	68	182
じゃあ	17	0	17
はい	12	0	12
じゃ	10	0	10
それでは	6	0	6
では	4	0	4
えーっと	2	1	3
それじゃ	2	0	2
合計	167 (71%)	69 (29%)	236

自由発話文中に見られる言語現象と比較して、冗長語の種類と分布の傾向が異なる。自由発話文に対する分析[2]では上位にランクされる「まあ」、「あー」など人間同士の対話に頻出する冗長語が全く現れず、種類も限定されている。

冗長語の1文あたりの出現回数は平均1.1回、冗長語を含む文は75%、冗長語を含まない文は25%であった。また指差確認呼称では、もともと周囲に声を聞かせるのが目的の一つであることから協調的な発声が期待できることがわかった。

4.2 音声認識タスクの設定

コーパスの分析結果よりわかるように、指差確認呼称の発話は自由発話文と比較して曖昧性の度合は低いものの、音声認識タスクとしては以下のような言語的曖昧性に対するロバスト性を持たせる必要があると判断した。なお言語的曖昧性の中で言い淀みや言い直しについては今回は対象外とする。

格助詞の欠落 (例) 名南寛政2号線

<を>停止します。

語尾の欠落 (例) 名南寛政2号線を

止<します>。

間投詞 (例) [えー] 名南寛政2号線を

[えー] 停止します。

最後に話者の条件については、制御所のオペレータは特定であり、またシステムへの習熟期間を設けることも可能であることから、特定話者または話者適応方式による対応で十分であると考えられる。

5 音声認識モジュールの実現方式

言語モデルをネットワーク文法(有限状態オートマトン)によって記述し、オートマトン制御 One-Pass-DP 法による連続音声認識システムを作成した。

5.1 言語モデルと音声理解処理

図2に指差確認呼称タスクとして記述した言語モデルの例を示す。省略(格助詞、語尾の欠落)に対しては文法にヌル遷移を設けることによって対応している。

前述のとおり、指差確認呼称の発話は操作手順に従い一連の動作シーケンスとして進行することから、全ての動作シーケンスに対応する文法を用意しておく、[選択呼称] → [実操作呼称] → [確認呼称]のように動作シーケンスの進行によって文法の動的な切換を行なう。

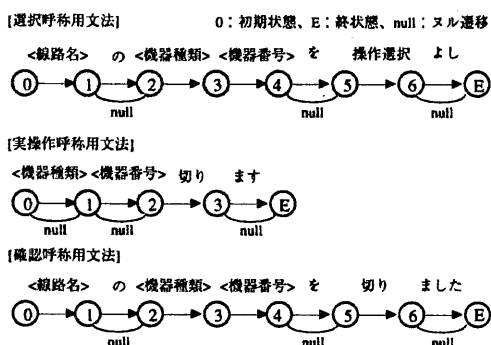


図2: 指差確認呼称タスクの文法

ここで述べる音声理解処理とは、認識処理結果から文の意味構造を抽出する処理である。本システムでは、図3に示す意味構造テンプレートによって発話意図の抽出を行なっている。

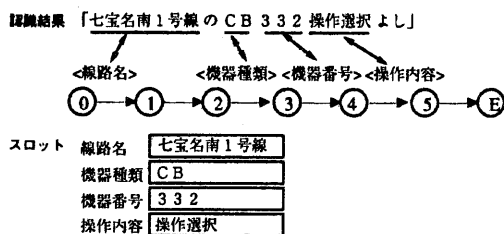


図3: 発話意図の抽出

5.2 冗長語の処理方式

コーパス分析より本タスクにおいては20種類程度の冗長語を辞書登録すれば、ある程度冗長語への対処を実現することができると考え、冗長語登録法により行なった。コーパス分析結果を参考にして出現頻度の高い間投詞を認識カテゴリとして登録し、文

頭と文節間に間投詞の挿入を受理できるオートマトン文法を記述した(図4)。

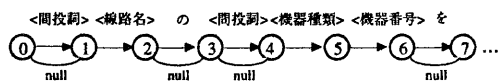


図4: 冗長語の処理方式

5.3 認識システムの概要

本研究で試作した認識システムの概要を述べる。音響分析条件は、サンプリング周波数12kHz、16bit量子化、20msハミング窓、フレーム周期10ms、特徴パラメータとして、16次ケプストラム、16次Δケプストラム、対数パワー、Δ対数パワーの合計34次元を用いた。音響モデルには音素環境依存モデルHMnet[3]を用いた。

5.4 基本性能の評価実験

音声モニタリング実現のための初期検討として、音声認識タスクの実現性を調べるために基本性能の評価実験を行なった。

5.4.1 音響モデルの学習

HMnetの学習には、ATR音声データベースのAセット5240単語(MHT)、Bセット115文(MHT)を用い初期モデル(400状態)を作成後に、Cセット5名分の話者データを用いて話者混合[4]により不特定話者モデルを生成した。実験では認識性能向上のために、この不特定話者モデルに対し男性話者1名の指差確認呼称の音声データ216文を用いて、移動ベクトル場平滑化方式(VFS)[5]により話者適応を行なったモデルを用いた。

5.4.2 実験条件

連続文認識と意図抽出の性能を評価した。実験では語彙数309、オートマトン平均分岐数12.58の規模の文法を使用し、ビーム幅は1000に設定した。冗長語は20種類を辞書登録した。評価データとして男性話者1名の指差確認呼称の音声データ(音響モデルの学習に用いていないオープンデータ、話者クローズ)を用意し、冗長語なし、冗長語ありの各々108文について連続文認識、意図抽出の実験を行なった。

5.4.3 実験結果

基本性能の評価実験の結果を表2に示す。

表2: 実験結果

	文認識率	意図抽出率
冗長語なし (108 文)	93.52(%)	97.22(%)
冗長語あり (108 文)	87.96(%)	97.22(%)

省略、冗長語 (間投詞) など言語的曖昧性を含んだ認識タスクについて、本手法による実現の見通しを得ることができた。基本性能としてはさらに意図抽出率を向上させる必要があるため、状況による辞書の局所化、すなわち使う場面によって使用語彙が局所化できることを利用して性能改善を図る予定である。また語彙数、文法を実際のタスクの規模に合わせて基本性能の検証を行なっていく必要がある。

6 おわりに

本稿では音声認識技術の電力分野への応用として、音声モニタリングによる誤操作防止機能、すなわちシステムが音声メディアを積極的に利用する新しい概念のインタフェースの提案を行なった。そして誤操作防止のためのインタフェースを自然言語を解する“仲介システム”によるチェック機能としてモデル化し、全体の機能の一部である音声認識モジュールの実装について初期検討を行なった。

今後の検討課題であるが、音声認識処理についてはロバスト性、リアルタイム性の向上を図ると共に、音声モニタリングを実現するためにはハンズフリーな高精度の音響入力系の実現が必要である。また本稿では述べなかった他のチェック機能のモジュールの実装を行ない、モデルの妥当性の検証を行なう予定である。

参考文献

1. 新美康永, 小林豊: “音声認識誤りを考慮した対話制御方式のモデル化”, 情報処理学会 音声言語情報処理研究会 95-SLP-5 pp.47-54 1995-2
2. 井ノ上直己, 武田一哉, 山本誠一: “ガーベジ HMM を用いた自由発話文中の不要語処理手法”, 電子情報通信学会論文誌, Vol.J77-A No.2 pp.215-222 1994-2
3. 鷹見淳一, 嵯峨山茂樹: “逐次状態分割法による隠れマルコフ網の自動生成”, 電子情報通信学会論文誌, Vol.J76-D-II, No.10 pp.2155-2164 1993-10
4. 小坂哲夫, 鷹見淳一, 嵯峨山茂樹: “話者混合逐次状態分割法による不特定話者音声認識と話者適応”, 電子情報通信学会論文誌, Vol.J77-A, No.2 pp.103-111 1994-2
5. 大倉計美, 杉山雅英, 嵯峨山茂樹: “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式”, 信学技報, SP92-16, 1992-6
6. 梶田征也 他: “原子力発電プラント マンマシンシステム (MMS)I-II”, 三菱電機技報, Vol.63, No.7, 1989
7. 三宅雅夫 他: “人にやさしい監視制御システムのヒューマンインタフェース”, 日立評論, Vol.78, No.10, 1996-10