

# 給電指令電話の音声検索システム

音声認識を用いた自動書き起こしとキーワード検索

## An Information Retrieval System for Telephone Dialogue in Load Dispatch Center

Automatic Transcription and Keyword Search Using Speech Recognition

(電力技術研究所 お客さまネットワークG 情報通信T)

系統運用業務の重要な連絡・打ち合せを行う給電指令電話では、打ち合せ内容の記録を目的として社員同士の通話の録音が行われており、蓄積された音声内容の検索に対するニーズが高い。また、音声テキストに変換することによって蓄積された情報の付加価値が高まる。そこで、本研究では音声認識を用いた自動書き起こしを行い、任意キーワードによる音声内容の検索を実現するシステムの開発を行った。

(Information and Communication Team, Customer Supply Network Group, Electric Power Research and Development Center)

In load dispatching operations, a telephone dialogue is recorded for the purpose of logging of operation jobs. Recently, the needs for information retrieval of a telephone dialogue are high. Moreover, the value of the accumulated information increases by transforming a speech into a text. Therefore, we have developed an automatic transcription system using speech recognition, and realized the keyword search function for telephone dialogue.

### 1 研究の背景と目的

近年の音声認識技術の進展に伴い、自然な話し言葉の自動書き起こしが実現されつつある。日常聞き流している音声情報を文字情報に変換して記録保存することができれば、膨大な情報の蓄積によって業務分析や教育訓練など事例活用の可能性が飛躍的に拡大するものと考えられる。

本研究で検討対象とする給電指令電話とは、系統運用業務で用いる専用の業務電話のことであり、電力を安全に供給するために電力系統の利用者に対し必要な指示、連絡を行うことを目的としている。業務の重要な連絡・打ち合せを行う給電指令電話では、打ち合せ内容の記録を目的として社員同士の通話の録音が行われており、蓄積された発話内容の検索に対するニーズが高い。また、音声テキストに変換することによって蓄積された情報の付加価値が高まる。しかしながら、音声データは言語情報であるにもかかわらず、キーワードを手掛かりとした内容検索が困難であり、人手による音声聴取と書き起こしは大変なコストを要する。そこで、本システムでは音声自動書き起こしによって通話内容のテキスト化を行い、キーワードによる音声内容の検索を実現する。

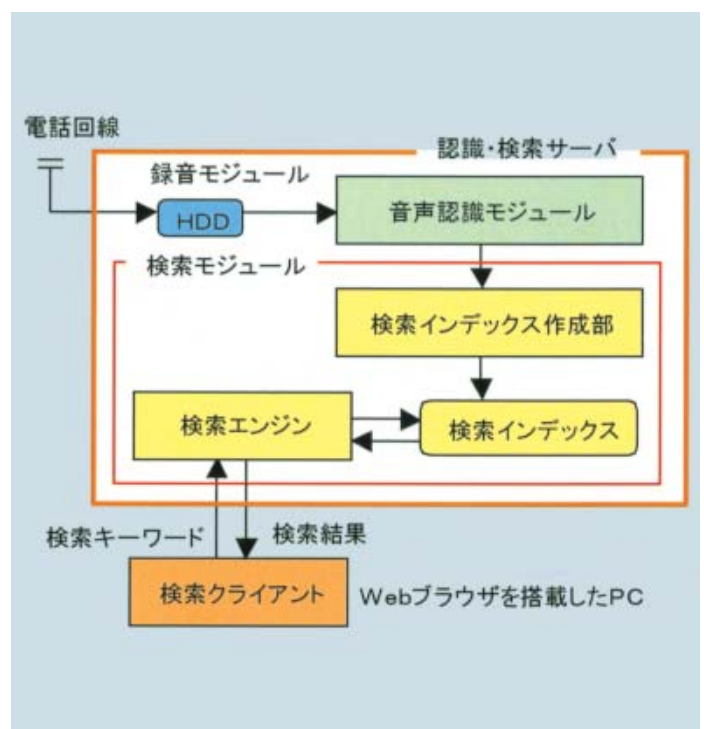
通話全体の音声聴取を実現する。

本システムの概要を第1図に示す。本システムは、認識・検索サーバ(Linux PC)と検索クライアント(Windows PC)より構成され、認識・検索サーバは録音、音声認識、検索の3つのモジュールより構成される。音声認識モジュールは自然な話し言葉に対応した大語彙連続音声認識システムであり、系統運用業務で用いられる約1万語を辞書登録している。ハードディスクに録音された音声ファイルは、音声認識モジュールによってテキストに変換され、検索インデックスに自動的に追加されていく。ここで、検索インデックスとは、音声認識結果に基づき通話中に現れる単語とその発話時刻を記録したデータベースのことである。検索クライアントのインターフェースはWebブラウザで実現しており、検索のための専用ソフトが不要な方式である。

### 2 システムの概要

本システムではユーザのニーズの分析を行った結果、以下の機能を実現することにした。

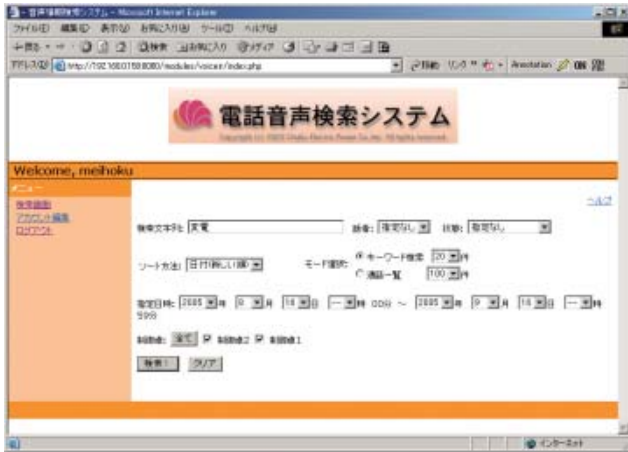
- (1) キーワード検索機能：任意キーワードによる音声内容の検索を実現する。
- (2) 時間範囲指定による検索機能：時間範囲指定による検索を実現する(キーワード検索と併用可)。
- (3) 書き起こし全文表示機能：通話ごとの書き起こし全文の表示を実現する。
- (4) 音声聴取機能：検索にヒットした音声区間、または



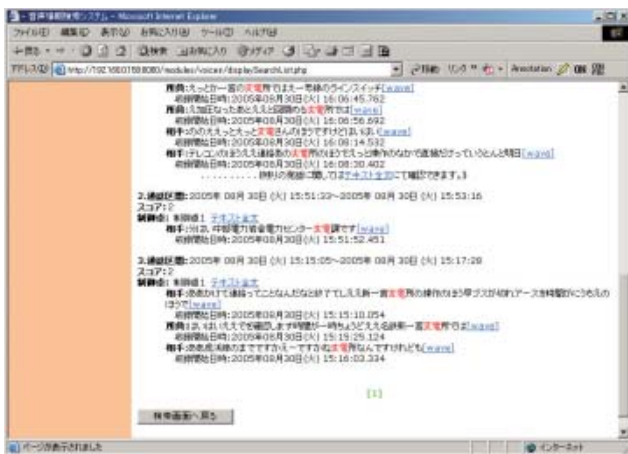
第1図 システム概要

### 3 システムの実装

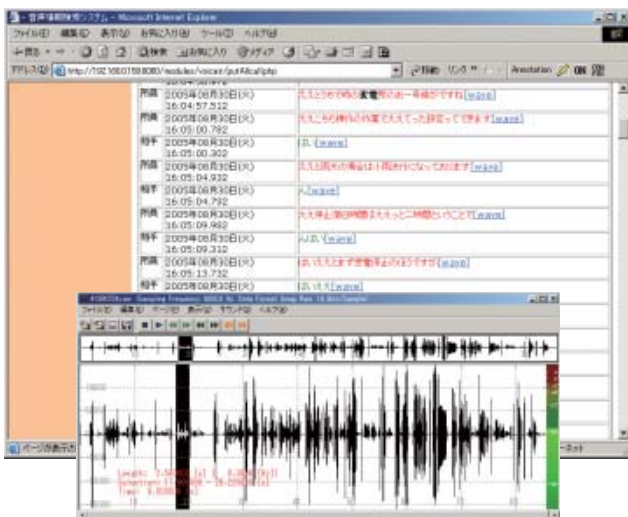
本研究で開発したシステムの検索画面例を第2図に示す。検索では複数キーワードの指定も可能である。また検索結果の画面例を第3図に、全文書き起こし結果の例を第4図に示す。第3図、第4図では画面の「wave」というリンクをクリックすることによって通話全体の音声波形を表示し、該当区間の音声を聴取することができる。



第2図 検索画面



第3図 検索結果の例



第4図 全文書き起こし結果例と音声聴取機能

### 4 システムの有効性評価

システムの有効性を評価するために実際の録音データを用いた評価実験を行った。実験では、給電制御所で録音した所員同士の打ち合せの30通話(総時間数 約35分、総単語数 5134語)を用いた。

まず認識性能の評価では、30通話の平均で単語認識率59.5%が得られた。ここで単語認識率の定義は、認識結果の単語数をN、置換誤りの単語数をS、脱落誤りの単語数をDとした時に次式のとおりである。

$$\text{単語認識率(\%)} = \frac{N - S - D}{N} * 100$$

次に、検索性能の評価では、検索キーワードとして人名、地名、一般名詞(変電所、電力センターなど)、話題語(打ち合せ、件名、工事など)の合計20単語から成るセットを用いた。実験では各セットの検索キーワードを一つずつ検索質問として合計20回の発話文の検索を行った。評価指標としては次式で定義される適合率(検索結果の精度)と再現率(検索結果の網羅性)を用いた。

$$\text{適合率(\%)} = \frac{\text{検索された発話文中の適合文の数}}{\text{検索された発話文の数}} * 100$$

$$\text{再現率(\%)} = \frac{\text{検索された発話文中の適合文の数}}{\text{全通話の発話文中の適合文の数}} * 100$$

その結果、適合率と再現率の平均でそれぞれ87.3%、67.2%という結果が得られ、重要な一般名詞や固有名詞を手掛かりとした検索は見かけの認識性能以上にうまく機能していると言える。現状の認識性能では完全な適合率と再現率を得ることは困難であるが、検索条件として複数キーワードの指定や時間範囲による絞り込みによってある程度カバーすることが可能である。実際の音声検索では、キーワードと共に時間範囲の条件を併用して候補の絞りこみを行った後に、詳細な通話内容の確認というプロセスを経ることが多い。本システムは、膨大な通話記録の中から検索対象の高速な絞りこみを行う上では実用的な性能が得られていると考えられる。

### 5 今後の展開

本システムはH17年3月より当社給電制御所の一つに試行導入され、実業務での評価を行っている。今後は他部門への技術展開を検討し、応用システムの開発を行っていく予定である。



執筆者 / 瀬川 修  
Segawa.Osamu@chuden.co.jp