

## 音声対話と全文検索を利用した電子ガイドシステム (1)

2E-4

## -システム概要-

藤田 稔 伊藤史朗 酒井桂一 山田雅章 上田隆也 池田裕治

キヤノン(株) 情報システム研究所

## 1 はじめに

最近、音声による対話システムの研究が盛んになってきている [1, 2]. 我々は、音声対話を用いることができることに加え、全文検索により文書データから取り出す情報を利用してガイダンスを行なう対話システムの研究開発を進めている。

今回、ガイダンスのドメインを「旅行」とし、その情報を収録した CD-ROM の全文検索結果からガイドするシステム: TARSAN (Travel Assistant system composed of information-Retrieval, Speech conversation Access and Natural-language techniques) を試作した。

本稿では、TARSAN の概要について述べる。

## 2 システム構成

本システムの全体構成を図 1 に示す。

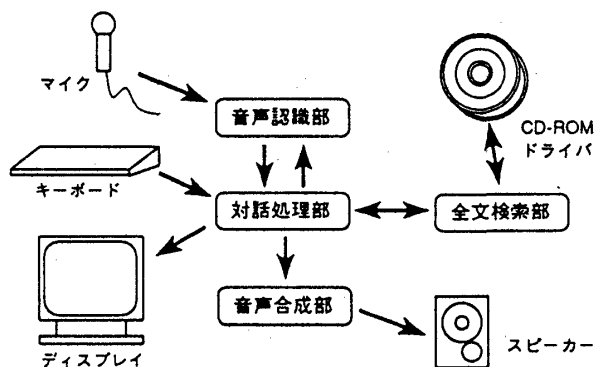


図 1: 全体構成

本システムは、4つのモジュール: 1) 音声認識部, 2) 対話処理部, 3) 全文検索部, 4) 音声合成部, から構成される。

全文検索の対象として用いている 2 枚の CD-ROM には、日本全国にある観光地 [8] と宿泊施設 [9] に関する

情報、約 8 万件が収録されている。全文検索部では 24 種類のジャンル (ホテル, ゴルフ場, 温泉, 神社, 美術館, ... など) および各ジャンルに依存する項目 (場所, 行き方, 説明, 電話番号, 料金, 立地条件, ホール数, ... など) を検索条件として、検索条件を満たすデータの名称および項目の内容を検索することができる。

入力はマイクによる音声入力とキーボード入力を切り替えることができる。キーボード入力では全 8 万件の文書データの検索に対処しているが、音声入力では現実的な時間で認識できる語彙数の制約から、場所を関東地方に限定し、16 個のジャンル名と 16 個の項目名の文書データに制限している。

対話処理部は、我々が以前に作成した、文書整形システム L<sup>A</sup>T<sub>E</sub>X のガイダンスシステム [3] の対話処理部に改良を加えたものを用いている。

出力としては、検索結果の詳細を出力するディスプレイ出力と、対話処理部で生成した適量の読み上げ文から音声合成部で規則合成した音声をスピーカから出力する音声出力とがある。

本システムは全て NeXT 上に実現した。対話処理部, 全文検索部, 音声合成部に 1 台の NeXTstation Turbo Color を用い、音声認識部として NeXTcube 1 台, NeXTstation Turbo Color 2 台の 3 台を用いた。

音声入力後、4つのモジュールで各処理を行ない、音声出力し始めるまでの時間は、検索条件として都道府県あるいは市町村名を指定した場合 [6], 10 ~ 20 秒である。

## 3 各モジュールの特徴

以下に各モジュールの特徴を簡単に述べる。

## 3.1 音声認識部 [4]

- 不特定男性話者, 一文単位発声の連続音声入力。
- 音素単位の連続分布型 HMM。
- 正規文法表現による有限状態オートマトン。
- 検索や対話の状況に応じた単語と文法の動的変更。
- 複数台の NeXT によるパラレル認識処理による高速化。

\*An Electronic Guidance System with Speech Conversation and Full Text Retrieval (1) - Overview -

Minoru FUJITA, Fumiaki ITOH, Keiichi SAKAI, Masayuki YAMADA, Takaya UEDA, Yuji IKEDA (Information Systems Research Center, Canon Inc.)

湯ノ花沢温泉
所在地：神奈川県箱根町
名称ヨミ：ユノハナザワオンセン
交通：小田原駅バス50分
宿泊施設(軒数)：1
宿泊施設(人数)：240
概要：芦ノ湯の上にあり展望絶好。硫化水素泉55～85度。標高940m。
効能：皮膚病
利用者数(年間)：61年(1月～12月)21,509人

図2: 観光地のデータ例 [8]

### 3.2 対話処理部 [5]

- ムード解析により、要求、提示、挨拶、条件変更などの発話意図を抽出。
- パラフレーズ機能による自然な文の入力が可能。
- 検索結果が多い場合、市町村名や項目値でグループ化して提示。
- 項目の一覧出力など、メタ質問が可能。
- 条件、検索項目の追加・変更が容易。

### 3.3 全文検索部 [6]

- 複数のCD-ROM文書から直接検索が可能。
- 任意の単語で検索が可能。
- 定型文書に対して項目単位での検索が可能。
- 単語検索だけでなく数値検索・地名検索などが可能。
- インデックスを併用することにより高速な検索が可能。

### 3.4 音声合成部 [7]

- NeXTのDSPによりソフトウェアだけで実現。
- 任意の漢字混じり文の正確な読み上げ。
- 自然で明瞭性の高い音声合成。

## 4 対話例

観光地に関するデータ例を図2に示し、そのデータに至る対話例を図3に示す。

## 5 おわりに

本稿では音声対話と全文検索を利用した電子ガイドシステムTARSANの概要について述べた。今後は、

- 対話処理にユーザモデル、推論機能などを導入し、対話管理機能を充実させる
- 音声認識の速度および精度を向上し、認識語彙数を増やす

ことにより、高度なガイダンス機能の実現を目指す。また、画像など、音声・自然言語以外のインタフェースの導入も検討していく。

U <sub>sr</sub> : おはよう。
S <sub>ys</sub> : おはようございます。TARSANです。
U <sub>sr</sub> : 神奈川県にある温泉について教えて。
S <sub>ys</sub> : 箱根町に16件、厚木市に4件、横須賀市に2件、湯河原町に2件など、合計で31件あります。
U <sub>sr</sub> : 箱根町では。
S <sub>ys</sub> : 芦ノ湖温泉、木賀温泉、底倉温泉、宮城野温泉など、合計で16件あります。
U <sub>sr</sub> : その効能を示せ。
S <sub>ys</sub> : 病後回復・ストレス解消が1件、皮膚病が1件あります。
U <sub>sr</sub> : 効能が皮膚病では。
S <sub>ys</sub> : 他の条件、所在地が箱根町、を残しますか。
U <sub>sr</sub> : はい。
S <sub>ys</sub> : 湯ノ花沢温泉です。
U <sub>sr</sub> : 湯ノ花沢温泉の行き方は。
S <sub>ys</sub> : 湯ノ花沢温泉の行き方は小田原駅バス50分です。
U <sub>sr</sub> : ありがとう。
S <sub>ys</sub> : またどうぞ。

図3: 対話例

## 謝辞

本研究を行なう機会を与えて下さいました知能工学研究部 田村秀行部長、UIを作成していただいた金子和恵氏に感謝いたします。

## 参考文献

- [1] 天野他：音声対話システムの試作，日本音響学会講演論文集，1-1-20，1992。
- [2] 竹林他：音声対話システムにおける Spontaneous Interaction の実現，日本音響学会講演論文集，1-1-22，1992。
- [3] 酒井他：文書整形システムを対象とする自然言語インタフェース，「自然言語処理の新しい応用」シンポジウム論文集，pp.22-31，1992。
- [4] 山田他：音声対話と全文検索を利用した電子ガイドシステム(2) -音声認識部-，情報処理学会46回全国大会予稿。
- [5] 酒井他：音声対話と全文検索を利用した電子ガイドシステム(3) -対話処理部-，情報処理学会46回全国大会予稿。
- [6] 伊藤他：音声対話と全文検索を利用した電子ガイドシステム(4) -全文検索部-，情報処理学会46回全国大会予稿。
- [7] 日本語の文字データを音声合成して出力 ネクストで動作するソフトウェア，EYE-COM 第4巻第15号，p.11，(株)アスキー，1992。
- [8] 「旅蔵」電子ブック，企画 製作(株)廣済堂，発行(社)日本観光協会，1990。
- [9] JTBの「宿泊情報」(電子ブック)，編集人 中嶋隆一，発行人 岩田光正，発行所 JTB 日本交通公社出版事業局，データ編集凸版印刷株式会社，1992。