

2T-2

音声操作による マルチメディア・タウン・マップ

加藤 誠巳 飯村 伊智郎 菊池 新 石田 裕三 能見 元英
(上智大学理工学部)

1 まえがき

不慣れた土地を訪れた場合、目的地にどのように行けばよいか迷うことがある¹⁾。目的地の情報を得るために地図を用いるのが普通であるが、目的地を見つけ出すのは必ずしも容易ではない。本稿では、パソコン PC9821 の Windows3.1 の下で動作する Windows Sound System の Voice Pilot および Text Reader を用い、音声認識および合成音声を用いた地理案内システムについて述べる。このように本システムの特徴は、キーボード、マウス、タッチセンサ等を使わず音声で操作する点にある。

2 案内の対象

2.1 対象地域

今回対象とした地域は、図1に示す四谷駅を中心とした2km四方である。

2.2 データベース

使用するデータベースは地図データベースと地理情報データベースより成る。

2.2.1 地図データベース

地図データベースとしては、国土地理院の数値地図10000(総合)の中から対象地域のデータを抽出し、使用している。地図情報項目は、以下にあげる8種類の項目を対象とした。

地図情報項目

一般道路(地上)、線路(地上)、線路(地下)、
駅(地上)、駅(地下)、丁目境界、区境界、
水涯線

2.2.2 地理情報データベース

地理データベースも、国土地理院の数値地図10000(総合)の中から対象地域のデータを抽出し、使用している。地理情報項目は、次に示す27種類に分類した。

地理情報項目

地域名、社寺、学校、大使館、ホテル、病院、
図書館、公園、会館、議会関係、皇室関係、
ビル、センター、警察、報道機関、事務所、
郵便局、名所、裁判所、生育地、劇場、区、
道路、インターチェンジ、鉄道、駅、堀



図1 案内の対象とする地域

A Multimedia Town Map System Operated through
Voice
Masami KATO, Ichiro IIMURA, Shin KIKUCHI, Hi-
romitsu ISHIDA, Motohide NOUMI
Sophia University

3 システムの概要

本システムの特徴はメニューの選択に音声認識を用いること、操作指定・確認に合成音声を用いることである。音声認識には Windows Sound system の Voice Pilot を使用し、音声合成には Text Reader を使用している。このような音声認識による指示は手による操作が不要となる長所がある反面、操作者の音声を誤認識してしまふ欠点がある。この欠点を補うために、認識された文字列を画面に表示すると共に、認識結果の当否を確認するようにした。正しい場合は次の指示に進み、間違っている場合は再度音声入力するものとする。

以下にシステムの操作画面および音声認識用単語辞書の構造について説明する。

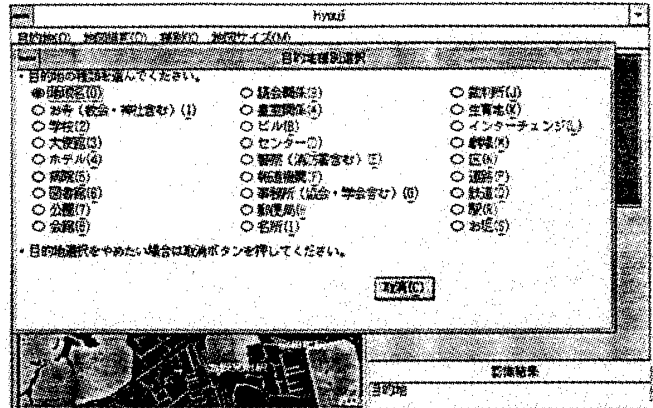


図 2 目的地の属性選択画面

3.1 目的地案内

前述の 27 種類の地理情報項目から目的地の該当する項目を選択する。直接目的地の名称を音声入力で選択しないで階層化しているが、これは種類が多いとそれだけ誤認識する確率が高くなり、システムの信頼性が低下するためである。更に前述の如く音声認識結果の当否を確認し、誤って深い階層レベルに入ることを回避するようにしている。

図 2 に目的地に属性選択画面の例を、図 3 に属性選択確認画面の例を示す。尚、案内後、駅と目的地の表示された地図を紙面などにコピーすることもできる。

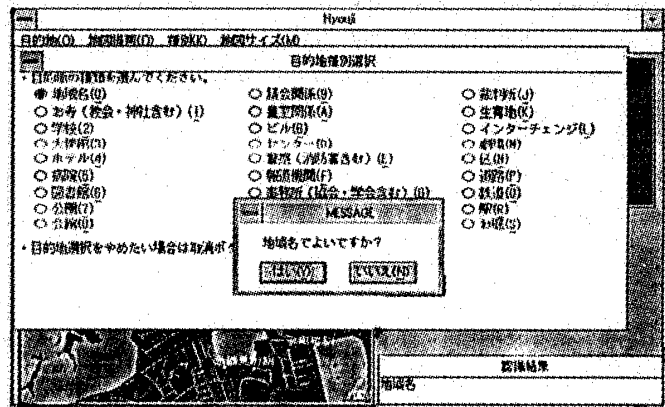


図 3 属性選択確認画面

3.2 音声認識用単語辞書の構成

複数の音声認識用単語辞書には、夫々に異なるシーン ID が付与されており、各辞書に属する単語の各々にはその単語が認識された後に移るべきシーン ID が指示されている。図 4 にこのように階層化した音声認識用単語辞書の構成の一部を示す。

4 むすび

音声によって操作するマルチメディア・タウン・マップについて述べた。今後ヒューマン・インターフェイスの更なる向上を図る予定である。

最後に、有益な御討論をいただいた本学マルチメディア・ラボの諸氏に謝意を表す。

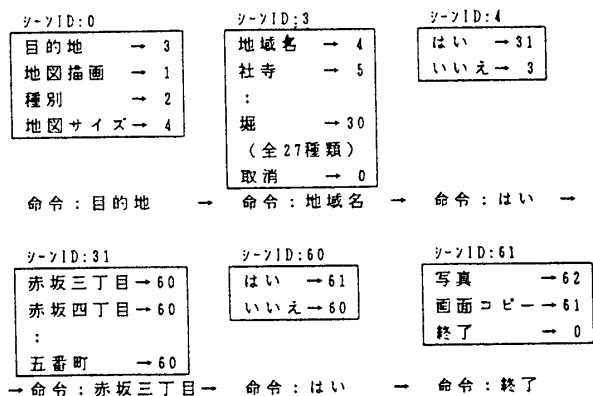


図 4 辞書の階層構造の一部

参考文献

[1] 加藤、本間: "歩行者用最適経路案内システム", 情報第 34 回全大, 5H-3(1987).