

음성인식기능을 이용한 지도정보 검색 시스템을 위한 검색엔진의 성능평가

°김 태 수* 조 명 회** 정 현 열*

*영남대학교 정보통신공학과

**경일대학교 측지공학과

Performance Evaluation of Search Engine for Speech Recognition Based Map Information Retrieval System

°Tae-Soo Kim*, Myung-Hee Jo**, Hyun-Yeol Chung*

*Dept. of Information and Communication Eng., Yeungnam University

**Dept. of Geodetic Eng., Kyungil University

E-mail : { kts, chy }@speech.yeungnam.ac.kr

mhjo@bear.kyungil.ac.kr

요 약

본 논문에서는 음성인식기능을 이용한 지도정보 검색 시스템의 실용화를 위하여 독자적인 지도검색 알고리즘을 구현하여 기존의 GIS(Geographic Information System)용 검색툴(MapInfo)을 이용함으로써 소요되는 비용을 최소화하면서도 어느 정도의 검색속도를 유지할 수 있는 음성구동지도검색시스템 구현을 위해 개발한 검색엔진의 성능 평가 실험을 통하여 그 유효성을 확인하고자 한다.

지도정보 검색시스템은 크게 음성인식부, 지도검색부로 나눌 수 있으며, 음성인식부에서는 유한상태오토마타(Finite State Automata)에 의한 구문제어를 통하여 OPDP(One Pass Dynamic Programming) 법으로 대상 단어의 인식을 수행하고, 지도검색부에서는 기존의 시스템에 사용된 OLE(Object Linking Embedding) [7]기법에 기저한 Mapinfo 툴을 이용하지 않고, Visual C++를 이용한 독자적인 알고리즘을 구성하여 지도자료를 읽어 들이도록 구성하였다.

평가결과, 사무실 환경하에서 지도검색용 68단어를 대상으로 실시한 on-line test에서 검색 대상 단어인식률은 98.02%를 얻었으며, 이때 해당지도 화면에 나타내는데 걸리는 시간은 평균 18.2초가 소요되었다.

I. 서 론

최근 지리정보 분야에서는 컴퓨터를 이용한 계도기술, 공간분석방법 등의 발전으로 인하여 현대사회에 적합한 지표공간분석 도구로서 지리정보시스템(GIS)이 개발되어, 지표에 대한 다양한 분야, 즉, 토지, 자원, 도시, 환경, 교통, 농업 및 군사 등에 이용되고 있다[4].

그러나, 기존의 GIS는 위치 좌표계로 지도를 검색하여 공간을 분석하는데 있어서, 키보드와 마우스를 사용함으로써 많은 시간과 노력이 소요된다.

이와 같은 단순 수작업에 의한 지리정보의 검색을 보다 신속하고 효율적으로 처리하기 위해서는 음성인식기능을 이용하면 가능하나 아직까지 이에 관한 응용에는 찾아보기 힘들다.

본 연구실에서는 각종 지도정보의 검색에 있어서 효율적인 처리를 위해, 음성인식기능을 이용함으로써 컴퓨터를 모르는 일반 이용자도 지도의 위치식별과 그에 따른 속성정보를 신속하고 쉽게 획득할 수 있는 지도정보 검색시스템[7]을 개발하였다. 그러나, 이 시스템은 인터페이스부에 OLE 기법에 기저한 고가의 Mapinfo 툴을 이용하고 있기 때문에 상용화를 위해서는 경제적이지 못한 단점이 있다.

본 논문에서는 이러한 점을 고려하여 개발한 지도검색 알고리즘을 소개하고 지도정보 검색을 고속화하기 위하여 음성인식 고속화기법을 도입한 on-line 테스트 결

과에 관하여 보고한다.

II. 지도정보 검색시스템

2.1 시스템 개요

그림 1에는 지도정보 검색시스템의 전체 구성도를 나타내며, 그림 2에는 실제 윈도우즈 화면을 나타낸다. 시스템은 크게 음성인식부와 지도검색부로 나눌 수 있다.

먼저, 음성인식부에서는 입력된 지도검색 명령어를 인식하여 그 결과를 지도검색부로 넘겨준다. 지도검색부에서는 음성인식부에서 넘겨받은 결과를 지도좌표로 변환한 후 이를 이용하여 검색어에 해당하는 지도를 화면에 나타내게 된다.

이때, 지도정보검색은 음성뿐만 아니라 키보드, 마우스를 통해서도 가능하며, 시스템의 개발환경은 Pentium 150MHz(Cyrix), 64MB의 메모리를 가진 개인용 컴퓨터환경이다

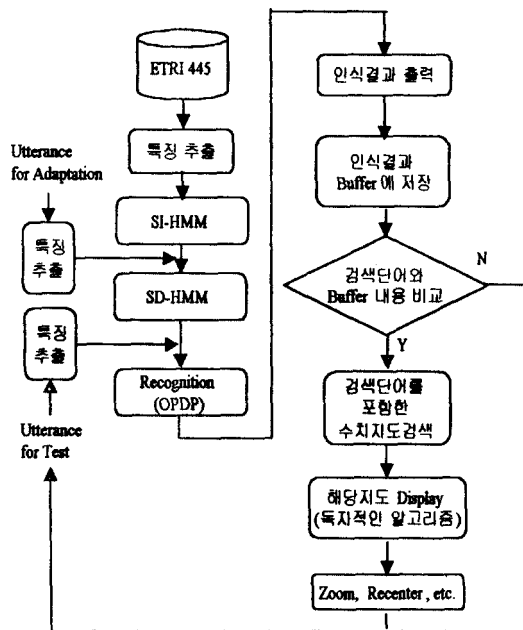


그림 1. 지도정보 검색시스템의 전체 구성도.

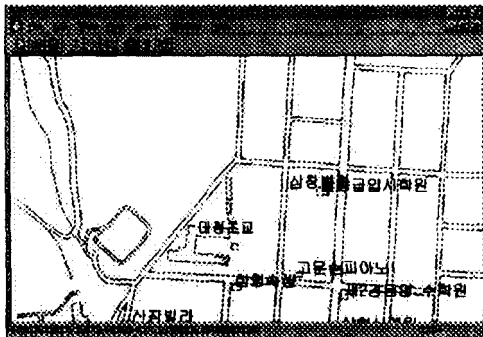


그림 2. 지도정보 검색시스템의 윈도우즈 화면.

2.2 음성인식부

2.2.1 음성 자료 및 특징추출

학습을 위해서는 ETRI에서 작성한 PBW(Phoneme Balanced Words) 445단어 음성 데이터베이스 중 14인 1회 발성을, 평가에 있어서는 발성환경이 다른 3인의 화자가 사무실 환경하에서 발성한 68개의 단어 4회 발성을 이용하였다. 표 1은 HMM모델의 학습과 평가를 위한 음성 자료를 나타낸다.

표 1. 학습용 음성 자료.

화자	단어수	발성	사용 단계	발성환경
14인	445	1회	SI-HMM 학습	방음부스
3인	68	1회	화자적응화	사무실
3인	68	3회	인식단계	사무실

음성의 특징추출은 먼저 발생된 음성 자료로부터 음성구간을 검출한 후, 14차의 LPC(Linear Predictive Coefficients) 분석을 통하여 10차의 멜렙스트림 계수(Mel Frequency Cepstral Coefficient, MFCC)를 구하고, 이 멜렙스트림 계수와 그 회귀계수(Regressive Coefficient, RGC)[5]를 추출하여 음성의 특징 파라미터로 사용한다.

2.2.2 HMM을 이용한 인식 및 고속화기법

HMM을 이용한 인식은 크게 학습단계와 인식단계로 나눌 수 있으며, 학습단계에서는 지도검색용 단어인식을 위한 인식의 기본단위는 48개의 유사음소를 이용하여 초기모델을 작성한다.

일반적으로 마이크의 채널왜곡, 발성환경의 변동 등은 인식률에 직접적인 영향을 미친다. 따라서, 최대사후 확률추정법(MAP)을 이용하여 SI-HMM을 적용화시킨다.[2].

또한 인식단계에서는, 적용화된 SD-HMM과 평가용 자료로부터 추출한 특징파라미터를 사용하여, 유한상태 오토마타에 의한 구문제어를 이용한 OPDP 알고리즘으로 인식을 수행한다[1].

음성인식에 있어서 가장 간단한 방법은 예측되어진 전체의 후보와 입력음성을 정합시키는 방법이다. 그러나 이 방법은 대상 어휘수가 증가하고 인식 알고리즘이 복잡해짐에 따른 대규모 탐색공간이 필요한 경우에 대해서는 많은 처리시간을 요구한다. 실시간 음성인식을 위해서는 전체의 후보와 정합을 행하지 않고도 고정도의 인식성능을 얻을 수 있는 효율적인 탐색수법이 필요하게 된다.

따라서, 본 시스템의 실용성을 고려하여 최근 많이 연구되어지고 있는 고속화 기법[1]으로 빔 탐색법(Beam Search)과 목구조사전(Tree-structured lexicon)정보를 도입한다.

2.3 지도검색부

2.3.1 지도 자료

수치지도를 구축하기 위해서는 대구광역시 수성구의 1:5,000 국가 기본도를 사용하며, 이중 특정지역명과 제어단어를 포함한 68단어를 검색대상단어로 한다. 개인용 컴퓨터 및 Workstation ARC/INFO software와 도서계획도 등을 이용하여, 행정 지번별로 표기되어 있는 지적도, 토양도, 지질도 등을 수치입력하여 데이터베이스화한다.[6]

또한, 지도 전체를 하나의 수치지도로 만드는 경우, 화면에 나타내는 시간이 많이 소요되므로 여러 개의 지도로 분할하여 작성하였으며, 필요시 하나의 지도로 통합시키는 것도 가능하도록 한다.

이와 같이 작성한 지도자료를 효율적으로 액세스하기 위해 MIF(MapInfo Interchange File)라는 파일형식으로 변환하여 사용한다.

표 2는 본 시스템에 사용된 수치지도 지명과 좌표 목록의 예를 나타냈다.

표 2. 수치지도의 지명과 좌표

번호	지명-지물명	X-좌표	Y-좌표
1	대구상공회의소	166,073.4	263,492.06
2	귀변에식장	166,244.04	263,503.05
3	국제오피스텔	166,247.11	263,253.01
:	:	:	:
67	황금2동사무소	166,079.235	260,223.29
68	황금초교	165,470.34	260,495.88

2.3.2 지도검색 알고리즘

본 연구실에서 개발한 기존의 지도정보 검색시스템 [7]은 OLE 기법에 근거한 GIS툴인 MapInfo를 이용하기 때문에 검색속도면에서는 우수하지만, 고가이기 때문에 상용화에 문제가 따른다.

따라서, 본 시스템에서는 이를 고려하여 지도를 읽어 들이는 부분과 검색 부분을 Visual C++ Programming 기법으로 독자적으로 구성하여 이런 문제를 해결하도록 하였다.

그림 3은 본 연구실에서 독자적으로 구성한 지도검색 알고리즘의 순서도를 나타낸다. 먼저 인식된 결과와 매칭하는 검색단어를 찾아낸 후, 그 검색단어가 포함된 지도를 찾아서 읽어들인다. 다음으로 X좌표와 Y좌표를 주어서 검색단어를 화면의 중앙에 위치시키는 순서로 구성되어 있다.

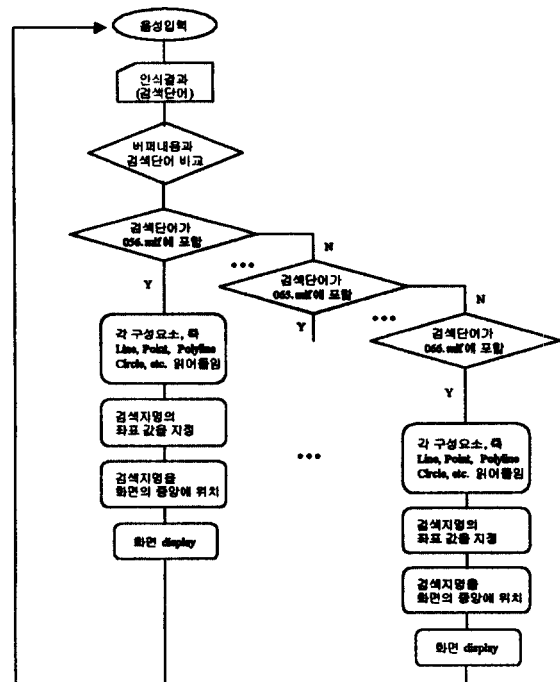


그림 3. 지도검색 알고리즘의 순서도

III. 인식실험 및 결과

본 시스템에 사용된 성능평가용 음성 자료는 사무실 환경 하에서 68개의 검색용 단어를 남성화자 3명이 큰 목소리로 데스크탑 마이크를 사용하여 발성한 자료를 이용하였다. 실제 음성인식 시스템에 사용되는 여러 환경에 대하여 고정도 인식률을 얻기 위하여 최대사후확률추정법에 의한 적응화기법을 도입하여 각 화자별 1회 발성을 이용하여 초기 화자독립 HMM을 적응화[2]하였다.

또한, 인식은 OPDP법을 이용하여 각 화자별 on-line 테스트 결과와 그때의 화면 display 시간을 그림 4, 5에 나타내었다.

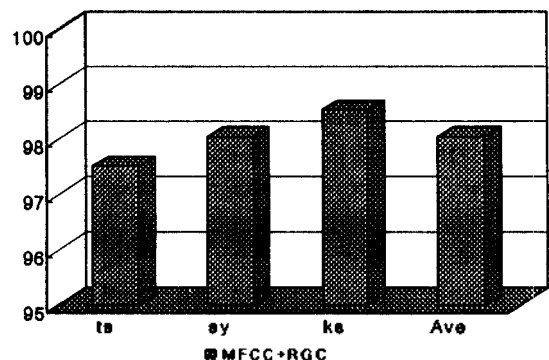


그림 4. 지도정보 검색용 단어에 대한 인식성능.

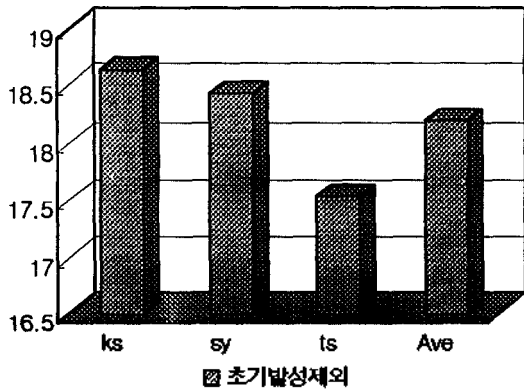


그림 5. 지도를 화면에 나타내는데 걸리는 시간.

이상의 결과로부터 MFCC와 RGC를 특징 파라미터로 사용했을 경우, 3인의 화자 평균 98.02%의 비교적 높은 인식률을 얻었으며, 이때 화면 display 시간은 평균 18.2초로서 기존의 OLE기법[7]을 사용했을 경우보다 비교적 긴 시간이 소요되었다.

전체적으로 볼 때 기존의 Mapinfo Tool[7]을 사용했을 경우는 지도를 화면에 display 하는데 평균 약 5초 정도로 비교적 빠르게 동작하지만, 경제적이지 못해 상용화 측면에서는 단점이 있다. 이를 고려할 때 현재의 시스템의 지도 display 속도를 개선한다면 음성인식기능을 이용한 지도정보 검색시스템의 실용화 가능성이 크다. 향후 본 시스템의 지도 display 속도개선을 위해 집중적으로 검토할 예정이다.

IV. 결론

본 논문에서는 음성인식을 이용한 지도정보 검색시스템에 있어서 기존의 OLE기법 기반의 검색률을 이용하지 않는 독자적인 지도 display 알고리즘을 구현하여 이를 평가한 결과 상용화 가능성을 확인하였다.

평가를 위하여 음성인식대상을 대구광역시 수성구의 속성단어와 제어단어를 포함한 68단어로 하고 이를 사무실 환경하에서 발생하여 시스템의 성능을 평가하였다.

성능평가 결과, on-line 테스트에서 음성인식에 있어서는 평균 98.02%의 비교적 높은 인식률을 얻었으나 지도 display에는 평균 18.2초의 비교적 시간이 많이 소요되었다.

그러나, 기존의 시스템의 문제점인 OLE 기반의 도구를 경제성을 고려하여 독자적으로 구성하여 음성인식기능을 이용한 지도정보 검색시스템의 실용화 가능성을 확인할 수 있었다.

향후에는 검색대상 단어를 확장하고, display 속도 등의 문제점 수정 보완에 관한 연구를 진행할 예정이다.

참고문헌

1. 이사육, "주소입력시스템을 위한 음성인식의 고속화에

관련 연구," 석사학위논문, 1997
 2. 越川忠, "連続音聲認識システムにおけるHMMの話者適應化に関する研究," 修士學位論文, 1993.
 3. Shi-Wook Lee, Deuk-Sou Kim, Hyun-Yeol Chung, "A Korean Address Input System Employing Automatic Speech Recognition," Proc. ICSP '97, Vol. 1, No. 1, pp. 373-377, Aug. 1997.
 4. Cartwright, W.E., and G.J.Hunter, "Beyond Maps : Using Multimedia to Enrich the use of Geographic Information," 1995 Annual Conference Proceedings, URISA, Vol. II, pp.420-433, 1995.
 5. Sadaoki Furui, "Digital Speech Processing, Synthesis, and Recognition" MARCEL DEKKER, INC, 1992.
 6. 조명희, Landsat TM영상에 의한 토지피복분류, 경북대학교 박사학위논문, 1989, 154p.
 7. 김태수, 정현열 "음성을 이용한 수치지도정보 검색 시스템의 구현" 한국음향학회 학술발표대회, pp.55~58, 1998년 7월