

음성을 이용한 수치지도정보 검색시스템의 구현

°김 태 수* 정 현 열*
*영남대학교 정보통신공학과

Implementation of Digital Map Information Retrieval System Using Speech Recognition

°Tae-Soo Kim*, Hyun-Yeol Chung*
*Dept. of Information and Communication Eng., Yeungnam University

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

요 약

본 논문에서는 음성인식기능을 이용하여 지리정보 검색시스템을 구현하고 그 성능 평가 결과에 관하여 보고한다.

성능평가실험을 위하여 사용한 수치지도 자료는 1:5,000의 대구광역시 수성구의 국가 기본도를 사용하였으며, 이 수치지도에서 속성단어와 제어단어를 포함한 68단어를 음성인식과 지도정보 검색의 대상으로 한다.

또한, 음성인식 결과와 지도자료와의 인터페이스는 OLE Automation 프로그래밍 기법을 이용하였으며, Pentium 150MHz, 64MB 메모리를 가진 개인용 컴퓨터 환경하에서 개발하였다.

성능평가 실험결과 남성 3인이 지도검색용 68단어를 실제 사무실 환경하에서 마이크를 통하여 on-line 테스트 하여 평균 98.02%의 인식을 얻었다.

이상의 결과로부터 음성인식기능을 이용한 수치지도 정보검색 시스템의 실용화 가능성을 확인할 수 있었다.

I. 서 론

최근 디지털신호처리 기술의 발전, 각종 미디어 기술의 발달, 초고속 정보통신망의 구축 등과 더불어 개인용 컴퓨터의 대량 보급으로 인하여 거의 모든 분야의 각종 서비스 업무를 컴퓨터를 이용하여 처리하는 것이 일반화되고 있다. 또한 일반인들의 서비스 품질에 대한 요구

가 점점 다양화되고 있다.

이와 더불어 지리정보 분야에서도 컴퓨터의 보급으로 제도기술, 공간분석방법 등의 발전으로 인하여 현대사회에 적합한 지표공간분석 도구로서 지리정보시스템(Geographic Information System; GIS)이 개발되어, 지표에 대한 다양한 분야, 즉, 토지, 자원, 도시, 환경, 교통, 농업 및 군사 등에 이용되어 왔다. 또한, 선진 각국에서는 멀티미디어를 GIS에 채용하여 보다 신속한 공간정보를 각종 산업은 물론이고 재난시에 활용하고 있다[4].

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

이와 같은 단순 수작업에 의한 지리정보의 검색을 보다 신속하고 효율적으로 처리하기 위하여 미국, 일본 등 선진 각국에서는 음성인식기능을 이용한 연구가 활발히 진행되고 있으나, 국내에서는 그 연구의 사례를 찾아볼 수가 없다.

따라서 본 연구에서는 각종 지도정보의 검색에 있어서 효율적인 처리를 위해, 음성인식기능을 이용하여 컴퓨터를 모르는 일반 이용자도 지도의 위치식별과 그에 따른 속성정보를 신속하고 쉽게 획득할 수 있는 수치지도정보 검색시스템을 개발하고자 한다.

이를 위하여, 음성자료는 한국전자통신연구원(ETRI)에서 작성한 445단어 중 14인 1회 발성과 3인의 화자가 사무실 환경에서 발성한 68단어를 이용하고, 수치지도자

료는 대구광역시 수성구의 1:5,000 국가 기본도를 사용하며, 이중 특정 지역명과 제어단어를 포함한 68단어를 검색의 대상으로 한다.

이때, 인식은 구문제어를 이용한 OPDP (One Pass Dynamic Programming)법을 이용하고, 인식결과와 수치지도와의 인터페이스는 OLE (Object Linking Embedding) Automation 기법을 이용하였다.

II. 수치지도정보 검색시스템

2.1 음성 자료 및 수치지도 자료

2.1.1 음성 자료 및 특징추출

인식을 위해서는 ETRI에서 작성한 PBW(Phoneme Balanced Words) 445단어 음성 데이터베이스 중 14인 1회 발성을 이용하여 화자독립 초기모델 (Speaker Independent HMM; SI-HMM)을 구성한다.

또한, 발성환경이 다른 3인의 화자가 사무실 환경하에서 발성한 68개의 단어를 이용하여 화자적응화를 수행하였다. 표 1에 HMM모델의 학습과 평가를 위한 음성 자료를 나타내었다.

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

2.2.2 수치지도 자료

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

또한, 지도 전체를 하나의 수치지도로 만드는 경우, 화면에 나타내는 시간이 많이 소요되므로 여러 개의 지도로 분할하여 작성하였으며, 필요시 하나의 지도로 통합시키는 것도 가능하도록 하였다. 그림 1에 데이터베이스 및 수치지도 구축과정을 나타냈으며, 표 2는 이를 이용해서 작성한 수치지도 지명과 좌표 목록을 나타냈다.

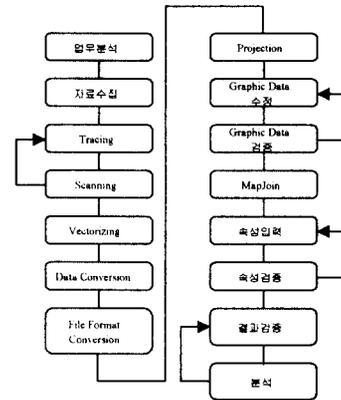


그림 1. 수치지도의 작성과정.

표 2. 수치지도의 지명과 좌표

번호	지명-지불명	X-좌표	Y-좌표
1	대구상공회의소	166,073.4	263,492.06
2	귀반예식장	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

2.2 HMM을 이용한 인식

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 확률(MAP)을 이용한 적응화 기법[2]으로 초기 HMM 모델(SI-HMM)을 적응화한다.

이때 사용한 HMM은 4상태 3출력분포의 연속출력분포 HMM(CHMM)을 이용한다.

또한 인식단계에서는, 적응화된 화자종속 HMM (Speaker Dependent; SD-HMM) 모델과 평가용 자료로부터 추출한 특징파라미터를 사용하여, 유한상태 오토마타에 의한 구문제어를 이용한 OPDP 알고리즘으로 인식을 수행한다[1].

2.3 음성인식과 수치지도정보의 인터페이스

본 시스템에 사용된 수치지도정보와 음성인식과의 인터페이스는 음성인식부분은 DLL(Dynamic Link Library)로서 작성을 하고, Visual C++ Programming 기법 중의 하나인 OLE(Object Linking Embedding; 객체삽입) Automation을 이용하여 GIS용 소프트웨어 MapInfo와의 인터페이스를 구성하였다.

OLE 기법은 엑셀과 같은 프로그램에서 작성한 도표

를 다른 응용프로그램에서 쉽게 사용가능토록 하는 프로그래밍 기법이며, 새로운 프로그램을 작성할 경우 이 기법을 이용하면 적은 시간과 노력으로 윈도우즈가 제공하는 여러 가지 기능을 이용할 수 있다는 장점이 있다.

이와 같은 OLE기법 적용하여 "대구문화방송"을 발생했을 경우의 실제 표현된 화면을 그림 2에 나타낸다.

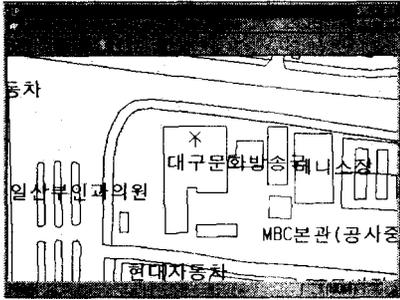


그림 2. 수치지도정보 검색시스템의 화면구성.

2.4 시스템 개요

그림 3에 이상의 방법으로 구성된 시스템의 전체 구성도를 나타냈다. 먼저 ETRI 445단어에서 특징을 추출하여 초기 SI-HMM모델을 구성한 다음, 발생환경 변동, 마이크의 채널왜곡 등을 고려해서 최대사후확률추정법을 이용하여 적용화용 자료로부터 특징을 추출하여 SD-HMM을 작성한다. 또한, 평가용 자료로부터 특징을 추출하여 음성인식부에서 OPDP법을 이용하여 인식한 후[3], 인식결과에 해당하는 수치지도를 윈도우 화면에 출력한다.

이때, 수치지도 검색은 음성뿐만 아니라 키보드, 마우스를 통해서도 가능하며, Pentium 150MHz, 64MB의 메모리를 가진 개인용 컴퓨터환경에서 개발되었다.

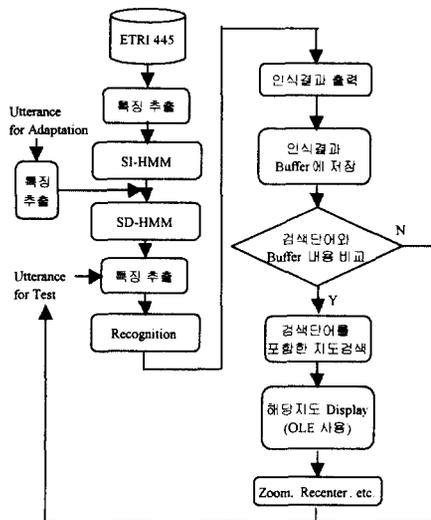


그림 3. 수치지도정보 검색시스템의 전체 구성도.

III. 인식실험 및 결과

본 시스템에 사용된 성능평가용 음성 자료는 사무실 환경하에서 68개의 검색용 단어를 남성화자 3명이 콘텐서 데스크탑 마이크를 사용하여 발생한 자료를 이용하였다.

실제 음성인식 시스템에 사용되는 여러 환경에 대하여 고정도 인식률을 얻기 위하여 최대사후확률추정법에 의한 적응화 기법을 도입하여 각 화자별 1회 발성을 이용하여 초기 화자독립 HMM을 적응화[2]하였다.

또한, 인식은 OPDP법을 이용하여 각 화자별 off-line 테스트에서는 적응화 전·후 그리고 on-line 테스트에서는 적응화 후에 대한 테스트를 수행하였으며, 그 결과를 그림 4, 5, 6에 나타내었다.

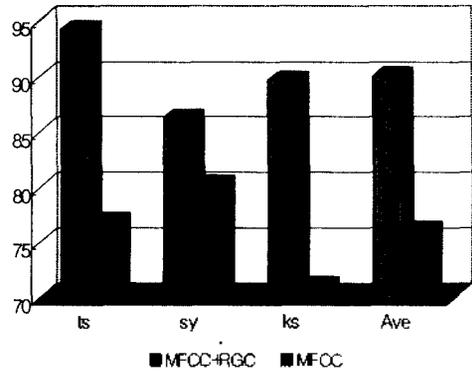


그림 4. 수치지도정보 검색용 단어에 대한 인식성능. (off-line Test : 적응화 전)

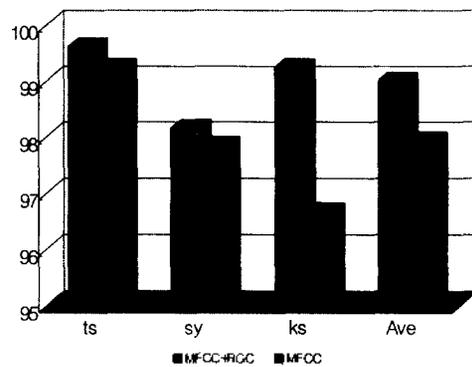


그림 3. 수치지도정보 검색용 단어에 대한 인식성능. (off-line Test : 적응화 후)

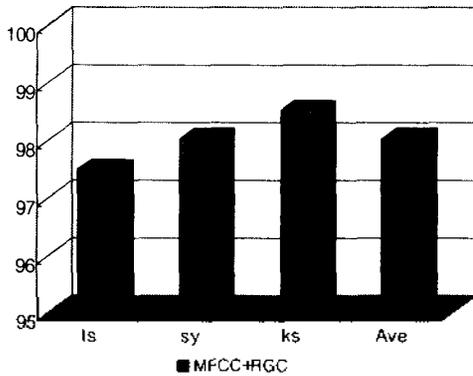


그림 6. 수치지도정보 검색용 단어에 대한 인식성능.
(on-line Test)

이상의 결과로부터 off-line 테스트에서는 MFCC와 RGC를 특징 파라미터로 사용했을 경우, 적응화 전 3인의 화자 평균 90.08%, 적응화 후 99.01%의 높은 인식률을 얻었고, on-line 테스트의 경우 3인의 화자 평균 98.02%의 비교적 높은 인식률을 얻어 실용화 가능성을 확인하였다.

IV. 결론

본 논문에서는 음성인식을 이용한 수치지도정보 검색 시스템을 구현하여 대구광역시 수성구의 속성단어와 제어단어를 포함한 68단어를 대상으로 사무실 환경하에서 성능을 평가한 결과 유효성을 확인하였다.

성능 평가를 위한 off-line 테스트에서는 3명의 화자 평균 적응화 전 90.08%, 적응화 후 99.01%의 인식률을 얻었고, on-line 테스트에서는 평균 98.02%의 비교적 높은 인식률을 얻었다.

이상의 결과로부터 음성인식기능을 이용한 수치지도정보 검색시스템의 실용화 가능성을 확인할 수 있었다.

또한, 현재의 인터페이스 기법 (OLE Automation 기법)을 적용한 경우는 프로그램 개발에는 이점이 있지만, 상용화라는 측면에서 경제적이지 못한 점이 있다.

따라서, 향후에는 검색대상 단어를 확장하고, OLE Automation 기법을 적용하지 않고, 독립적으로 구동되는 시스템 개발에 관한 연구를 진행할 예정이다.

참고문헌

1. 이시욱, "주소입력시스템을 위한 음성인식의 고속화에 관한 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.