

음성구동에 의한 수치지도 정보 검색시스템*

조명희** · 정현열***

A Voice Activated Digital Map Information Retrieval System*

Myung-Hee Jo** · Hyun-Yeol Chung***

요약 : 지리정보시스템(GIS)을 이용하여 특정지역을 수치지도화하고 이를 기초로 지역계획이나 자원 및 환경 관리에 적용한 연구 예는 국내외적으로 많이 찾아 볼 수 있다. 이들 GIS는 위치좌표체계를 이용하여 지도를 검색하고 공간분석하는 데 있어서 마우스의 조작으로 화면의 크기를 조절함과 동시에 대상지역의 범위를 지정하기 때문에 수행과정에서 동일한 동작을 수없이 반복하게 되어 검색대상지역을 화면에 구현하는 데 정시간이 소요된다. 그러나, 이와 같은 단순 수작업에 의한 정보의 검색은 음성인식을 이용하면 신속하고 효율적인 처리가 가능하나 음성을 이용하여 정보를 검색하는 것은 GIS 분야에서 아직 찾아볼 수 없다.

본 연구에서는 각종 지도정보의 검색에 있어서, 번거로운 수작업을 필요로 하는 검색용 명령어를 음성입력으로 처리하여 컴퓨터를 모르는 일반 사용자도 정확한 도시공간정보를 보다 신속하게 입수할 수 있도록 하는 음성구동 수치지도정보 검색 시스템을 개인용 컴퓨터를 이용하여 개발하였다.

지도정보 검색의 범위는 대구시의 1개 區로 하고, 수치주제도인 목록으로부터 Top-down방법으로 공공기관, 특정건물, 도로의 순으로 인식, 검색하게 하였다.

기본Window 화면상에 圖내의 수치주제 후보 단어를 빌정한 음성이 마이크를 통해 입력된 되면 A/D변환기(또는 사운드 카드)를 통해 16 kHz/16bit 정도의 디지털신호로 변환되어 음성구간 검출 알고리즘을 통해 음성부분만 검출된다. 인식기의 성능은 단어인식률 98% 이상으로 하고 이를 위한 인식방법은 48HMM 음향모델을 기본으로하는 OPDP법을 이용하였다. 아울러 사후확률추정법을 개발하여 환경적응 혼란을 추가한다.

주요어 : 음성구동, GIS, 수치주제도, 음성구간알고리즘, 사후확률추정법

Abstract : There are a lot of case studies on applying the digital mapping of a specific area by GIS (Geographic Information System) to the regional plan or the mangement of the resource and environment. In retrieving and analyzing the map spatially using the coordinates, it takes a long time for these GIS to display the retrieval area on the screen because they repeat the same operation in the execution process by regulating the screen size and at the same time, designating the range of the study area with a mouse. In fact, the information retrieval by a mere manual labor can be treated more promptly and effectively if voice recognition is used. However, there have been no precedents for information retrieval using voice in the field of GIS so far.

In this study, voice activated digital map information retrieval system was developed on the personal computer. It turned a command language for retrieval involving a cumbersome manual labor into a voice input and thus enabled the general users to obtain exact spatial information on city more rapidly.

As the map information retrieval area, an urban district, Gu, in Taegu was designated and Top down method was employed to recognize and retrieve the governmental institutions, landmark buildings and streets by order from the list of digital subject items. The voice pronouncing the candidate key of digital subject items within Gu was input through a microphone and turned into 16kHz or 16bit digital code by A/D Converter (Sound Card). As a result, the voiced portion alone was detected through end-point detection algorithm.

The capacity of the recognition equipment was set to recognize more than 98% of the words and the method used for it was One Pass Dynamic Programming based on the 48HMM acoustic model. Besides, Environment Adaptation Training was added with the aid of Posterior Probability Estmation.

Key Words : voice activation, Geographic Information System, digital thematic map, end-point algorithm, posterior probability estimation.

* 본 연구는 1996년 학술진흥재단 회계간 연구지원사업 학술연구비의 지원에 의해 수행되었음

** 경인대학교 측지공학과 조교수(Assistant Professor, Department of Geodetic Engineering, Kyungil Unversty)

*** 영남대학교 전자통신공학과 부교수(Associate Professor, Department of Electronic Engineering, Yeungnam Unversty)

1. 서 언

최근 인구의 증가로 인하여 토지가 점차 세분화, 다양화 되어감에 따라 지표에 대한 과학적인 정보는 매 5년마다 2배씩 증가하고 있으며 도시계획과 재개발, 환경자원의 균형있는 이용을 위하여 지표정보의 효율적인 관리에 대한 요구가 점점증하고 있으나 지표정보의 관리를 위해 현재까지 주로 사용되어온 기존의 종이지도나 항공사진으로서는 복잡하고 집약적인 지표양상과 물체를 표현하는 데는 극히 제한적이었다.

한편, 디지털 신호처리 기술의 발전, 각종 미디어 기술의 발달, 초고속 정보통신망의 구축 등과 더불어 개인용 컴퓨터의 대량 보급과 각종 통신수단의 발전으로 인해 여러분야에서 각종 서비스 업무를 컴퓨터 시스템을 이용하여 처리하는 경우가 상식화되어가고 있고 이와 더불어 일반인들의 서비스의 질에 대한 요구도 점점 고급화 되고 있다.

지리학 분야에서도 컴퓨터시스템에 의한 제도기술의 발달, 공간분석 방법의 발전의 결과로 현대사회에 적합한 지표공간 분석도구로서 지리정보시스템(GIS:Geographic Information System)이 개발되어 지표(地表)에 대한 다양한 분야, 즉, 토지, 자원, 도시, 환경, 교통, 농업, 해양 및 군사 등에 활용되어 왔다. 아울러 수치지도제작(Digital Mapping)을 비롯하여 토지정보시스템(Land Information System:LIS), 도시정보시스템(Urban Information System:UIS), 도면자동화(Automated Mapping:AM)그리고 시설물관리(Facility Management:FM) 등이 활발하게 이용되고 있다.

그러나, 기존의 GIS는 위치좌표체계를 이용하여 지도를 검색하고 공간분석하는 데 있어서 키보드, 또는 마우스를 이용하여 화면의 크기를 조절함과 동시에 대상지역의 범위를 지정하는 조작을 수회 반복하여 수행함으로써 검색대상 지역을 화면에 구현할 수 있는 데 이를 위해서는 많은 시간과 노력이 소요된다.

이와 같은 단순 수작업에 의한 정보의 검색은 음성인식을 이용하면 신속하고 효율적인 처리가 가능하다. 음성에 의해 구동되는 지도정보 검색시스템을 개발하는 데 있어 가장 핵심이 되는 음성인식 기술은 지난 30여년간의 기술축적에 힘입어

팔복할 만한 발전을 가져와 연속음성인식에 있어서도 단어인식률 97%를 상회하고 있다. 이 기술을 이용하여 외국에서는 자동통역시스템, 여행정보안내 시스템, 관광안내 시스템, 증권정보 안내시스템을 개발하여 상품화하고 있으며 국내에서도 음성구동 퍼스날 컴퓨터, 증권정보안내 시스템이 개발되어 소개되고 있고, 미국, 일본등과 같이 자동통역시스템 개발사업에도 참여하고 있다. 또 음성다이얼링 휴대폰을 개발하여 필드테스터등에 있다. 그러나, 이러한 연구 결과들이 개인용 컴퓨터에서 구현되어 일반 사용자들이 쉽게 이용할 수 있는 응용제품을 개발한 사례는 많지 않은 실정이다.

특히 GIS분야에서 구미 각국에서는 멀티미디어를 GIS에 채용하여 보다 신속한 공간정보를 각종 산업을 비롯한 제난시에 활용(Fung:1995 Hirschfield etal:1995 Lao:1995 etc)하고 있으나 국내에서는 아직 연구 개발되지 못하고 있는 실정이다.

이와같은 점을 고려하여 본 연구에서는 각종 지도정보의 검색에 있어서, 번거로운 수작업을 필요로 하는 검색용 명령어를 음성입력으로 처리하는 시스템을 개인용 컴퓨터를 이용하여 개발하고, 특정지역의 수치지도를 바탕으로 시각적인 인터페이스를 뒷받침하는 좌표체계를 음성에 의한 위치식별체계로 전환함으로써 이용자가 지도의 위치식별과 그에 따른 속성정보를 신속히, 쉽게 획득함으로써 필요한 지역의 공간정보를 입수할 수 있게 하고자 하는 데 그 목적을 두고 있다. 즉, 음성인식 기법과 GIS기법을 통합한 다기능의 공간정보체계를 개발함으로써, 컴퓨터를 모르는 일반 이용자도 정확한 도시공간정보를 보다 신속하게 입수할 수 있도록 하고자 한다.

2. 시스템의 개요

본 연구에서는 불특정 화자의 발생에 의해 마이크로를 통해 입력된 지도정보 검색 요구에 대해 수치주제별에 해당하는 단어에 대해 인식률 98%이상의 정확도로 동작하여 필요한 정보를 출력하는 시스템을 개발함을 연구의 최종목표로 한다. 본 연구에서의 지도정보 검색의 범위는 1개 區로 하고 이 구역내의 21종의 수치주제 도면 목록으로부터

Top-down방법으로 공공기관, 특정건물, 도로의 순으로 인식, 검색하게 한다. 이때 시스템은 개인용 컴퓨터(Pentium)의 윈도우95 환경에서 동작할 수 있게 한다.

수치지도정보 검색 시스템(이하 시스템이라함)의 전체구성을 그림1 에 나타내었다. 시스템 동작의 전체 흐름을 간략하면, 먼저 마이크를 통해 입력된 음성은 A/D 변환기(또는 사운드 카드)를 통해 16 kHz/16bit 정도의 디지털신호로 변환되어 음성구간 검출 알고리즘을 통해 음성부분만 검출된다. 검출된 수치지도정보에 관한 단어구간은 인식부에서 인식하여 인식결과에 해당하는 지역의 지도와 정보를 window 화면에 출력한다. 수치지도 정보 음성인식에 있어서는 광역 단위로부터 Top-down방법으로 인식, 검색된다. 입력은 음성과 컴퓨터의 기본 입력 장치인 키보드, 마우스 모두를 통해서 가능하며, 음성에 의한 입력도중에도 시스템의 기본 명령어에 대한 입력이 가능하도록 하였다.

인식결과, 바르게 인식하였을 경우에는 사용자가 요구한 지도 정보를 필요한 형태로 출력할 수

있게하고 인식오류가 발생할 경우에는 재발성, 또는 마우스, 키보드 입력으로 수정할 수 있게한다.

지도정보 입력에 있어서는 연구지역의 여러종류의 수치주제 후보 단어를 window 화면상에 나타내고 인식 결과에 따른 하위 단위의 후보만을 보여줌으로써 사용자가 지역정보를 선택 하는 데 있어서의 오류를 방지하고 지역의 지도와 정보 출력에 있어서는 지역단위, 또는 공공기관 등의 대상을 발생하여 데이터베이스에 정리, 저장된 필요정보를 출력하게 한다. 이때 출력형태는 전자의 경우와 동일하며 사용하는 마이크는 Head Mount 또는 Desk-top 형태를 사용하였다.

음성구간 검출 알고리즘을 통해 입력된 수치지도 단어는 그림 2와 같은 과정을 통해 인식된다. 즉 먼저 48PLU(Phoneme Like Unit)를 미리 훈련한 후 발생된 지도정보 음성을 이용하여 재학습하고, 학습된 PLU를 이용하여 인식할 경우는 인식률에 있어 OPDP(One Pass Dynamic Programming)알고리즘을 이용하여 인식한 후 그 결과를 출력하게 하였다.

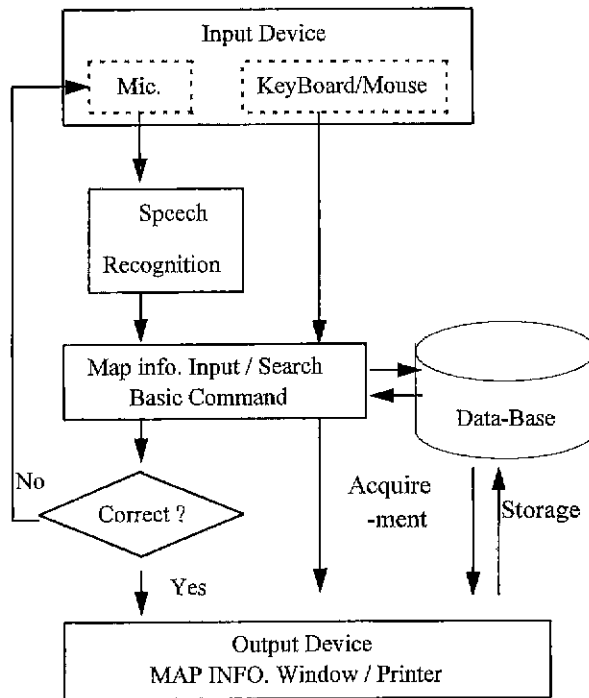


그림 1. 전체 시스템의 개략도

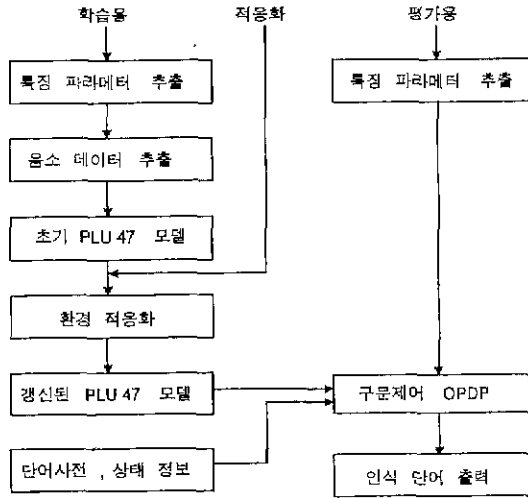


그림 2. 인식 시스템의 구성도

3. 수치주제도 작성 및 속성자료 데이터베이스 구축

1) 수치주제도의 작성

(1) 연구지역 선정

본 연구에서는 대구광역시의 구역에서 수성구를 연구대상지역으로 선정하였다. 이는 대구광역시의 7개의 구(區)중에서 토지이용이 다양하게 나타나고 지형지물의 구성도 다른 구에 비하여 각각 그 특성이 독특하기 때문에 본 연구의 주제가 되는 음성신호에 의한 식별가능성도 높을 것으로 사료되기 때문이다. 또한 동구나 북구에 비하여 대상면적도 적당하여 연구수행시의 작업용량이나 완성후의 실행시에 시스템을 효과적으로 운용할 수 있는 적정범위로 판단되었다. 아울러 음성에 의한 지도검색과정에서 용도지역의 구성밀도가 조밀하게 되면 효율적인 결과를 기대할 수 없고 인터페이스상의 문제도 감소되기 때문에 일반적으로 용도지역의 구성패턴이 균일하지 않은 지역의 선정에 역점을 두게되어 본 연구지역을 선정하였다.

(2) 지도작성자료 및 방법

- ① 수치지도 구축을 위해서 1:5,000의 지형도를 기본도로 사용
- 지역의 정보검색에서 상위체계에서 하위체계로

이동하게되므로 보다 상세한 정보를 획득하기 위해서는 더욱 축척이 큰 기본도면을 사용할 수 있으나, 특정 범위내의 지역에서 대축척 도면단위 일수록 그래픽 데이터 및 속성자료의 량이 과다하여 검색과정에서 하나의 도면을 모니터상에 구현하기 위한 시간이 극히 지연될 위험이 있으므로 연구지역의 규모와 정보량의 균형을 유지함으로써 효율적인 지도검색시스템을 구축을 위해 1:5,000의 지형도를 기본도로 사용하였다.

② 주제도별 정보검색을 위한 Map/Info software 사용

기존지형도에서는 지표의 복합적인 정보가 한정된 지면상에 포함되어 있어 지표환경(예: 등고선, 수계망, 도로망, 토지이용 등)에 대한 주제별 분석이 불가능하였으며, 또한 지표물체가 변화되어 갱신을 요할 때도 일정한 기간이 경과하지 않으면 수정인쇄가 어려운 현실이다. 그러나 layer개념이 강조되어 설계된 Map/Info software를 이용하여 주제별 데이터베이스구축이 일단 완료되면 자료갱신이 용이해 질 뿐만 아니라 중첩분석(overlay analysis)을 이용한 주제별 공간분석이 용이하게 된다.

(3) 수치 주제도 작성

- ① 수치주제도면 목록
- 수치주제 도면의 목록과 각 주제도면의 자료원은 표 1과 같다.

표 1. 수치주제도면의 목록

No	주제도	자료원(data source)	비 고
1	등고선도	1 : 25,000 지형도	제곡선
2	수계망도	1 : 25,000 지형도	
3	행정경계도	1 : 28,000 대구광역시도	동경계
4	도로망도	1 : 5,000 지형도	아스팔트, 포장도로
5	토지이용도	Landsat 인공위성영상	
6	철도망도	1 : 28,000 대구광역시도	
7	도시계획도	1 : 25,000 도시계획도	
8	동사무소분포도	1 : 28,000 대구광역시도	도청, 시청, 구청
9	우체국분포도	1 : 28,000 대구광역시도	
10	경찰서	1 : 28,000 대구광역시도	
11	병원	1 : 28,000 대구광역시도	대학병원
12	학교	1 : 28,000 대구광역시도	초등, 중, 고등, 대학교
13	시장	1 : 28,000 대구광역시도	재래시장, 백화점
14	도서관	1 : 28,000 대구광역시도	공공도서관
15	방송국	1 : 28,000 대구광역시도	
16	아파트	1 : 28,000 대구광역시도	아파트단지
17	은행	1 : 28,000 대구광역시도	
18	지가분포도	1 : 3000 대구광역시 지적도	각 그리드 중앙의 지가
19	기타	1 : 28,000 대구광역시도	인지도가 높은 지형지물

② 주제도면 제작과정

수치주제도 작성을 위한 데이터베이스 구축과정은 그림 3과 같다.

③ 수치주제도의 작성결과

위 과정을 통하여 작성한 수치주제도는 21종이며 이의 일례를 그림 4에 나타내었다. 이 주제도는 다음과 같은 특징을 가진다. 본연구에서는 주요한 지형지물의 지명을 중첩되지 않도록 도면상에 적절히 배정하여 도면상에 표시하였다.

가. 종래의 지도에서는 지도 데이터를 일괄적으로 도면에 나타내었기 때문에 특정한 항목만을 이용하기가 곤란했으나 수치 지도에서는 항목별로 구분되어 저장되므로 지도데이터를 선택적으로 이용할 수 있다.

나. 도로 경로의 거리 계산 등 수치 지도 데이터를 처리하여 각종의 표나 도면으로 출력할 수 있다.

다. 수치 지도 데이터에 속성 정보를 추가하면 지형 정보시스템으로써 이용이 가능하다.

라. 수치화 되어 있기 때문에 신속한 출력이 가능하며 수정 작업이 용이하다.

또한 지도 작성의 생략 기능이 우수하여 편집도

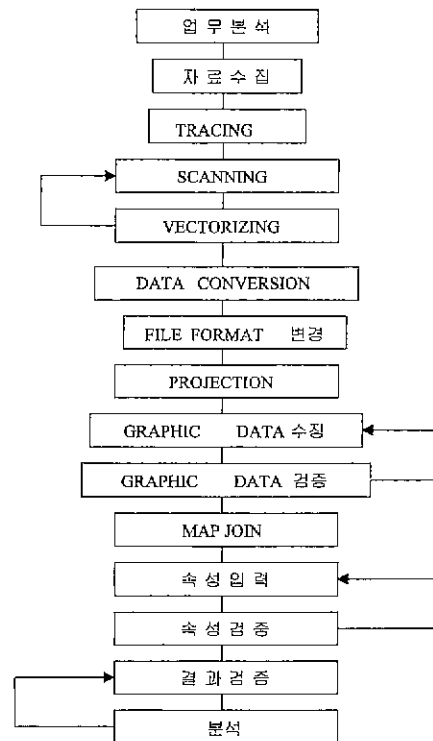


그림 3. 수치주제도면의 제작과정

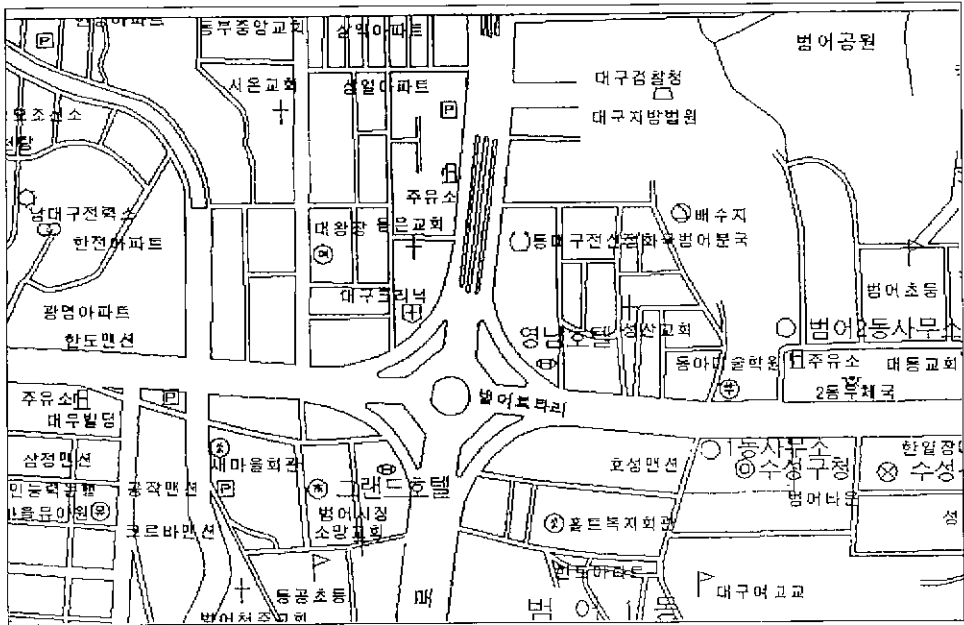


그림 4. 도로주제도를 나타낸 지도에 (대구시 수성구 범어로터리 주변)

제작이 편리하다.

마. 지도 투영법의 변환이 가능하다.

4. 음성구동을 위한 음성인식 알고리즘 작성

1) 입출력 인터페이스 설계

(1) 음성 구간 검출

음성 구간 검출이란 지도검색단어 음성이 발생되었을 때 입력된 음성신호로부터 원하는 음성의 구간을 찾아내는 과정이다. 이것은 일반적인 음성인식 시스템 구현에 있어 인식기의 인식 성능에 큰 영향을 끼치는 데, 특히 본 시스템과 같은 단어를 기본으로 하는 음성인식에서는 매우 중요하다. 즉, 단어 인식의 경우 음성신호의 첫부분은 대부분 자음으로 시작되는 데 이때 잠음구간이 크면 그 구간이 자음으로 오인식되어 인식을 저하를 가져오게 된다. 따라서 끝점 검출은 인식기의 성능을 좌우하는 중요한 문제로 정밀한 끝점검출 알고리즘이 요구된다.

이를 고려하여 본 연구에서는 다음과 같은 2단

계의 검출과정을 거쳐 끝점검출을 실시하여 정확도를 높였다.

① Input Magnitude에 의한 방법

이 방법은 현재 마이크가 켜진 상태에서 입력되는 잠음과 음성을 입력 level에 의해 구분하여 음성구간을 판정한다. 이를 위해서는 먼저 buffer에 약 10초간(16bits/16kHz시 320kBytes)의 음성을 저장할 수 있는 공간을 확보하고 분석구간 160 samples 마다의 입력레벨을 측정한 후, 문턱치(threshold)값을 넘는 입력신호가 입력될 경우 음성구간이라 판정하고 buffer에 저장된 데이터를 입력으로 받아 들인다. 이때 CallBack 함수를 사용하여 전후의 묵음구간을 조정(보통30msec) 할 수 있도록 설계하였다.

② ZCR(영교차율)과 Power에 의한 방법

2단계 음성구간 검출은 식(1)에 의해 결정된 일차 음성구간을 보다 정확하게 검출하여 인식성능 향상, 인식시간 단축을 도모함을 목적으로 한다. 이를 위해 현재까지 음성 구간 검출에 많이 이용되고 있는 ZCR과 Power정보를 이용한다.

진체적인 끝점검출 알고리즘은 다음과 같이 수행된다. 먼저16kHz로 샘플링률로 A/D 변환된 데이터는 10msec 단위의 블록으로 나눈다. 여기서,

처음 입력된 10 블록(100msec)의 데이터를 묵음구간의 잠음으로 가정하고 그 구간에서 에너지와 영교차율의 통계적인 특징을 구하여 기준값을 정한다. 이때 에너지의 기준값 ThEL과 ThEU은 식(1) - (4)와 같이 각 블록 에너지 값의 평균에 비례하도록 결정한다. 블록 i의 에너지 E(i)는 계산량을 고려하여 각 샘플의 절대값의 합으로 하였다.

$$E(i) = \sum_j^{80} abs[data(j)] \dots\dots\dots (1)$$

$$ThEL = \min(1.6 * [\sum_i^{10} E(i)] / 10 + offset, THEL) \dots\dots\dots (2)$$

$$ThEU = 2 * THEL \dots\dots\dots (3)$$

$$ThBig = 4 * ThEL \dots\dots\dots (4)$$

영교차율(Zero Crossing Rate)은 원래 신호값이 0축 양쪽으로 부호가 바뀌는 횟수로 정의하고 있지만 여기서는 잠음의 영향을 줄이기 위하여 다음과 같이 정의한다. 먼저 식(5), (6)과 같이 ThEL에 비례하는 SilenceU, SilenceL을 정하여 신호값을 4개의 영역으로 나눈다.

$$SilenceU = ThEL / 35 + 1 \dots\dots\dots (5)$$

$$SilenceL = -SilenceU \dots\dots\dots (6)$$

이를 이용하여 그림 7과 같이 신호값의 변화가 2개 영역이상인 횟수를 계산하여 영교차율로 한다. 즉 신호값이 1→3, 1→4, 2→4, 3→1, 4→1, 4→2로 변화하는 횟수만을 계산한다. 이렇게 하여 각 블록의 영교차율을 구한 후 (7)을 이용하여 영교차율의 기준값 Th_ZCR을 계산하였다. 이 기준값 Th_ZCR은 예비실험 결과 25로 하였다.

$$ThZCR = \min[mean(zcr) + 2 * \sigma, ThZCR] \dots\dots\dots (7)$$

음성이 아닌 묵음이 계속되는 경우, 0.5초 간격

으로 10블록의 데이터를 이용하여 기준값을 다시 계산한다. 새로 계산된 값(Th2)과 이전 값(Th1)을 비교하여 식(7)과 같이 새로운 기준값(ThNew)을 설정하였다. 이렇게 하여 주위 잠음의 변화에 적응하도록 하였다.

(2) 윈도우즈용 기본화면설계

음성인식 지도정보검색 시스템의 기본화면은 다음과 같은 수순에 의해 설계하였다. 먼저 시작 단추에 의해 시스템이 동작하기 시작하고 입력되는 음성의 레벨이 표시되며 음성이 입력되어 처리중에 있을 때는 처리중임을 나타내게 하였다. 이때 음성구간 검출알고리즘을 통하여 검출된 입력음성 파형도 동시에 표시된다. 음성구간추출이 완료되면 표시화면이 바뀌어 인식 진행과정, 인식결과, 인식 단어후보등이 표시된다. 인식결과 오인식이라고 판단될 경우 단어 후보를 선택할 수 있으며, 자판 입력, 인쇄 등을 선택할 수도 있게 된다.

윈도우 환경에서 동작하는 음성인식 시스템 구현을 위해 윈도우즈용 음성인식 시스템의 사용자 인터페이스(Graphic User Interface)와 음성의 입력 및 추출을 위한 사운드카드의 인터페이스를 중심으로 고찰한 후 이를 이용하여 본 시스템의 인터페이스를 설계하기로한다. 그림 7에 본 시스템 설계를 위한 시스템의 구성요소를 나타내었다.

설계할 윈도우 인터페이스 부분은 그림의 블록 다이어그램 중 음성입력장치를 통한 음성입력 부분과 이를 파일로 기록하는 인식결과 부분이다.

① 음성인식 시스템의 인터페이스 설계

PC를 이용한 음성인식 시스템은 기존의 내장된 사운드 카드를 사용할 수 있으므로 별도의 음성입력 장치를 없앨 수 있다. 따라서 여기서는 우선 내장된 사운드 카드를 이용하는 것으로 하고 인식속

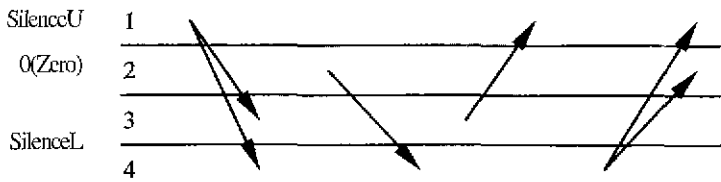


그림 5. 영교차율을 구하는 방법

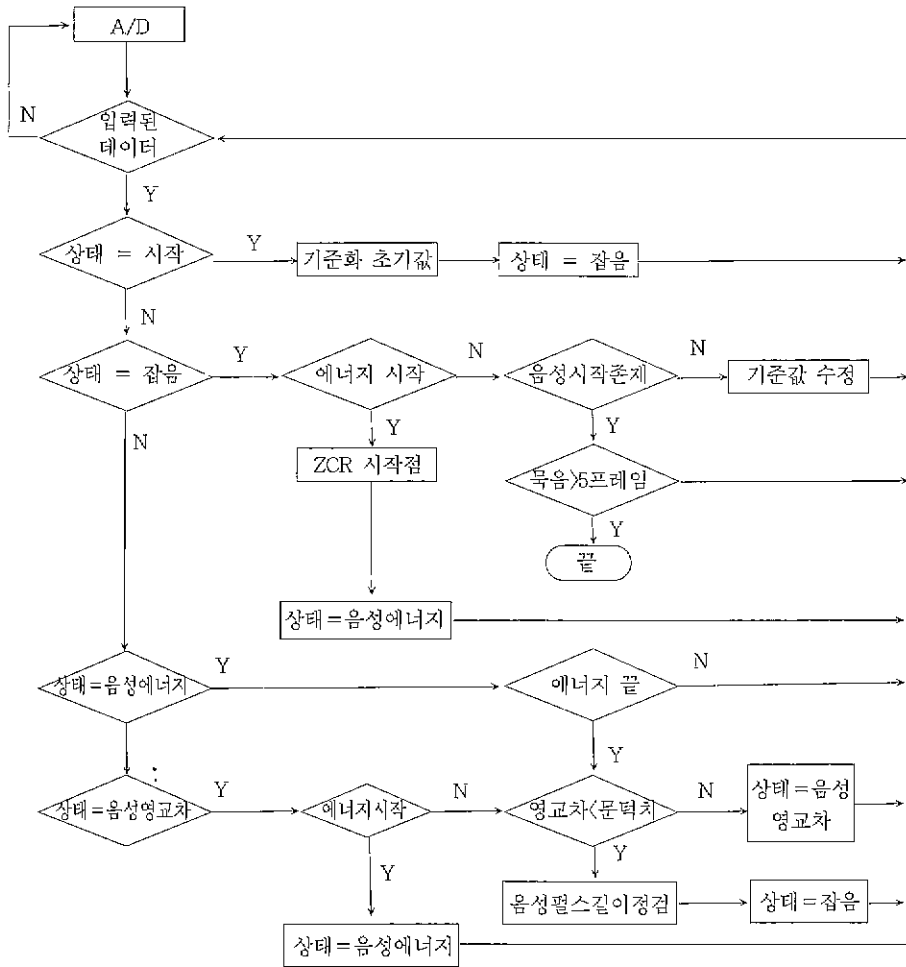


그림 6. 음성구간 검출

도에 문제가 발생할 경우 별도로 제작한 A/D 보드를 사용하기로 한다. 인터페이스 제작 대상 환경은 아래와 같다.

〈시스템 실행환경〉

- 윈도우즈95 기반의 32bit 어플리케이션 환경
- 윈도우즈 호환의 16bit 사운드 카드
- 16M의 램 및 Pentium(권장)

설계될 인터페이스 프로그램은 윈도우즈95 기반의 32Bit로 컴파일 가능하며 윈도우즈 내의 MDK(Multimedia Developing Kit)의 사운드 API를 사용하여 높은 호환성을 유지하게 되도록 한다. 이 부분은 주소음성 메시지가 입력되면 실시간으로

음성구간을 추출하여 실제 음성인식이 이루어지는 내부 블록으로 넘겨주기 위해 사운드 파일(WAV)로 저장하는 것이다. 인터페이스 화면설계를 위해서는 스탠포드 대학의 Dragon Dictate System을 참고하였다.

② 인터페이스 구조 설계

Dragon Dictate 시스템의 분석에서 얻어진 결과를 참고하여 인터페이스 화면을 설계하였다. 설계된 인터페이스 화면들은 시스템 매뉴얼(그림 8), 메뉴버튼을 눌렀을 경우의 창(그림 9), 입력된 음성중 음성구간 추출과정을 완료한 음성파형을 표시하는 창(그림 10)으로 이루어져 있다.

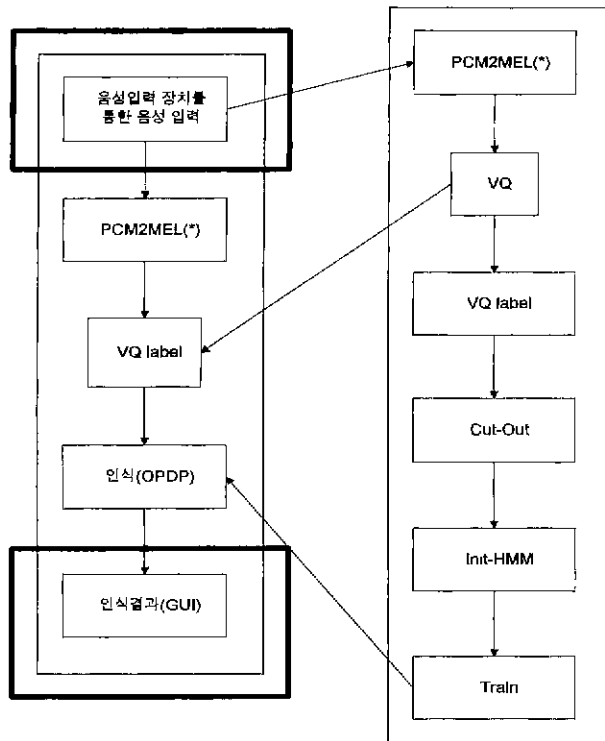


그림 7. 시스템 구성요소

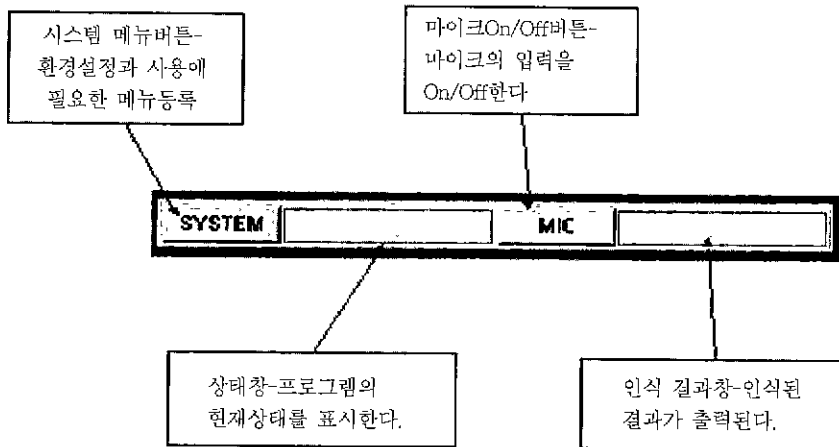


그림 8. 설계한 시스템의 Tool-bar

실제 인터페이스는 단순하면서 동작상태를 잘 알 수 있도록 하기위해 마이크 버튼을 누르면 음성의 입력이 시작되는 것으로 하였다. 음성파형은

음성구간 자동추출 알고리즘을 이용하여 입력된 음성을 계속 검사하여 음성으로 판단되면 추출하여 16bit, 16kHz의 wave 파일형태로 저장한 후 원

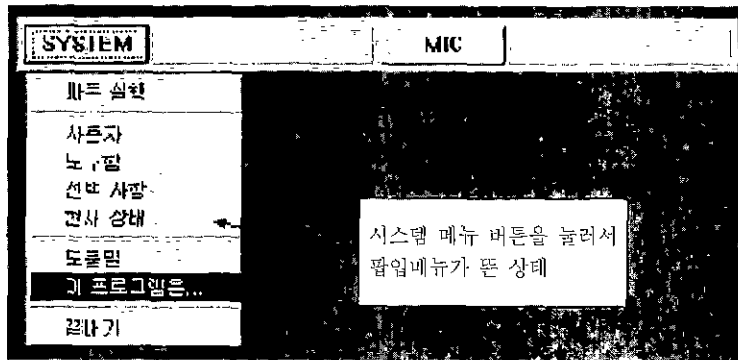


그림 9. 설계한 시스템 화면(메인메뉴 버튼이 눌러진 상태)

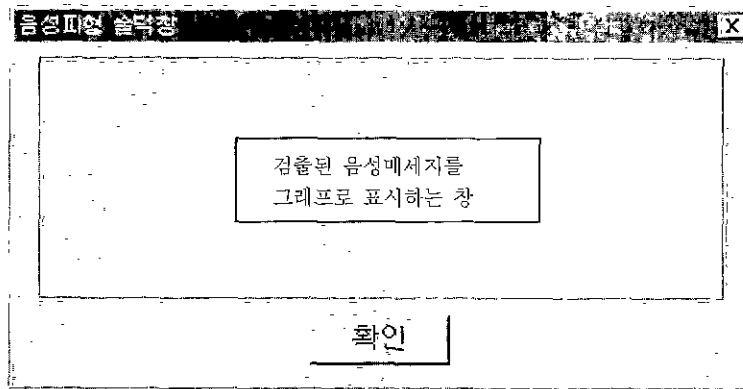


그림 10. 설계한 시스템 화면(음성파형 표시창)

도우95의 바탕화면에 저장되어서 바로 결과를 알 수 있게 하고 그 음성을 그래프로 음성파형창에 표시하여 볼 수 있도록 하였다.

③ 음성구간추출 방법

가. 음성구간의 시작부분 검출

입력된 음성신호로부터 음성메시지 부분만을 추출하는 데에는 여러 가지 방법이 사용될 수 있다. 예를 들어 한정된 구간의 Zero Crossing Rate를 측정하거나, 구간내의 전체 Power를 구하여서 이것이 메시지 부분인지 아니면 잡음인지를 판별하는데 이용할 수 있다. 그러나 인터페이스 구현에서는 단순히 음성메시지를 약간의 여유를 두고 비교적 정확하게 추출한 후 다음 단계의 정밀 검출루틴에 넘겨줄 수 있는 것을 목표로 하기로 한다.

여기서 사용한 알고리즘은 마이크의 기본상태

(입력이 없는 상태)로부터 얻은 적당한 스투스 홀드값을 이용하여 일정한 검색구간 내의 최대값과 최소값의 차이를 구간의 대표값으로 하여 이와 비교하는 방법을 사용하였다. 그림 11에 이과정을 나타내었다.

이 방법은 단순하지만 검색구간이 짧은 경우 빠르고 비교적 정확하였다.

나. 음성구간의 끝부분 검출

음성 메시지의 끝부분도 처음과 마찬가지로 스투스 홀드값과 비교하는 형식으로 구할 수 있지만, 그렇게하면 음성 메시지가 여러 음절로 이루어진 경우, 음절사이를 메시지의 종료점으로 판단할 수 있기 때문이다. 그림 12에 이러한 예를 나타내었다.

이 예는 “close window”를 발음한 경우인데, “clo-”와 “se”, 그리고 “window” 사이에 명령이 발생중인 데도 묵음구간이 나타나는 것을 볼 수 있

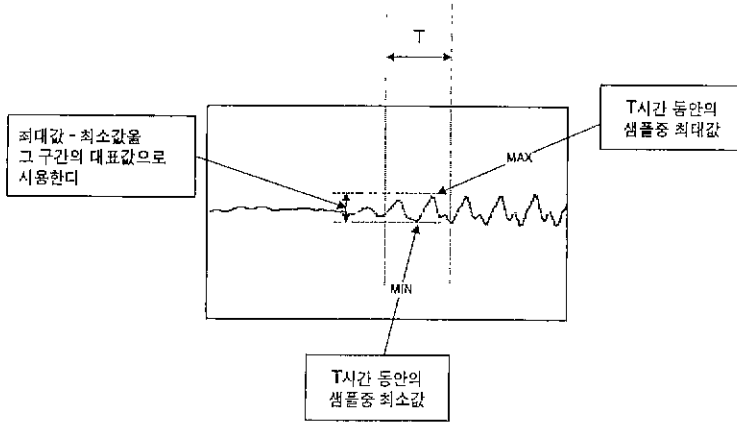


그림 11. 음성구간의 시작부분 검출

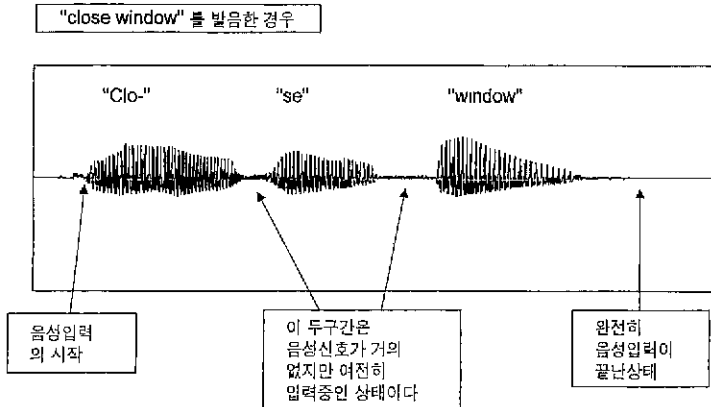


그림 12. 음절음성인 경우의 끝점 검출

다. 단순히 스투스홀드 검사만 시행한다면 이것은 음성 메시지의 종료로 인식되어 뒤쪽의 메시지는 잃어버리게 될 것이다. 이러한 현상을 막기 위해서 종료점 판별에는 일정시간 동안의 음성신호 변화를 체크하여 스투스홀드 검사에 의해 종료점으로 판별이 되더라도 그것이 일정한 횟수 이상 반복되지 않으면 종료시키지 않는 방법을 사용했다. 그러나 이 경우 여분의 부분이 음성 메시지로 포함되므로 마지막 종료점이 아닌 종료체크 카운터가 처음 시작한 부분을 기억했다가 전체가 종료되는 것이 확실해지면 첫 점을 기준으로 종료시키는 방법이 사용되었다. 이렇게 해서 입력단어의 끝점을 정확히 검출할 수 있게 되었다.

다. 여분의 버퍼 첨가 후 저장

1과 같은 스투스 홀드방법에 의한 음성 메시지 검출방법에 의하면 초성이나 종성이 파열음, 파찰음 등으로 소리나는 경우 이를 정확하게 체크해내지 못해서 음성 구간이 유실되어 버리는 일이 생기곤 한다.

그러나 이런 파열, 파찰음은 음성인식 시스템에서 아주 중요한 부분으로 이 부분이 생략되면 전혀 다른 음성으로 오인식되는 결과가 생기곤 한다. 이것을 방지하기 위해서는 Zero Crossing Rate법을 사용하는 것이 적절한 데 이를 위해 전체 음성 메시지의 앞뒤에 충분한 여분을 두어 음성추출을 하여 메시지가 삭제되지 않도록 하였다. 파열음, 파

찰음은 대부분 아주 짧은 지속시간을 가지므로, 조급의 여분으로도 좋은 결과를 얻을 수 있었다.

그림 13에 설계한 윈도우 인터페이스의 실행과정을 나타내었다.

5. 수치지도 검색과정 및 결과

1) 수치지도 단어 음성데이터베이스 구축

대구시 수성구의 수치지도 속성단어는 다음과

같이 지형지물을 중심으로 100단어를 선정하였다. 이를 10인이 5회 발성한 단어를 이용하여 음성데이터베이스를 구축하였으며 이를 표 2에 나타내었다.

2) 인식기의 구성

음성구간 검출 알고리즘을 통해 입력된 수치지도 단어를 인식하기 위해 인식시스템은 앞에서 언급한 그림 2와 같이 구성하였다. 이때 구성한 48PLU(Phoneme Like Unit)는 다음과 같다. 48PLU는 미리 훈련한 후 발생된 지도정보 음성을 이용하여 재학습한다. 이 학습된 PLU를 이용하여

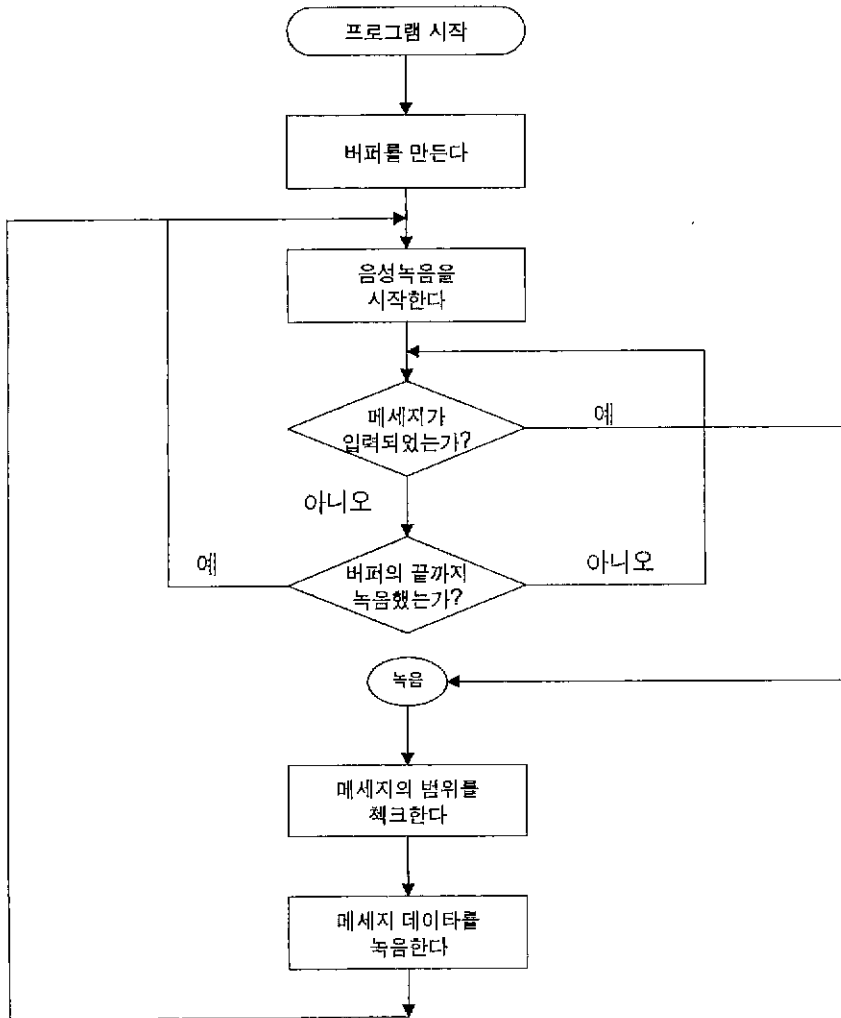


그림 13. 윈도우즈 인터페이스 동작의 흐름

표 2. 속성정보 음성데이터베이스(수성구)

No	지명	비고	No	지명	비고	No	지명	비고
1	경동초등학교		35	동일초등학교		69	수성경찰서	
2	경북고등학교		36	동천초등학교		70	수성관광호텔	
3	경산대한방병원		37	두산동사무소		71	수성구청	
4	경신중고등학교		38	두산초등학교		72	수성여자중학교	
5	골든엔선		39	만촌1동사무소		73	수성중학교	
6	그랜드호텔		40	만촌2동사무소		74	수성초등학교	
7	남대구수성전화국		41	만촌3동사무소		75	시립도서관	
8	남부정류장		42	만촌동우체국		76	신명여자중학교	
9	남산여자고등학교		43	만촌초등학교		77	신일전문대학	
10	농인중고등학교		44	망우공원		78	아리아나호텔	
11	대구과학고등학교		45	목화엔선		79	어린이회관	
12	대구과학연수원		46	MBC문화방송		80	영남공업고등학교	
13	대구광역시교육청		47	범물여자중학교		81	영남호텔	
14	대구남양학교		48	범물중학교		82	오성중고등학교	
15	대구등중학교		49	범물초등학교		83	용지초등학교	
16	대구스포츠헤터		50	범어1동사무소		84	유성프라자	
17	대구여자고등학교		51	범어2동사무소		85	지봉초등학교	
18	대구은행본점		52	범어2동파출소		86	정화여자중학교	
19	대구지방법원		53	범어3동사무소		87	중동동사무소	
20	대구환경청		54	범어4동사무소		88	중동우체국	
21	대등은행본점		55	범어공원		89	중앙상업고등학교	
22	대륜중고등학교		56	범어동우체국		90	지산동사무소	
23	대청초등학교		57	범이초등학교		91	지산중학교	
24	덕원중고등학교		58	범일초등학교		92	지산초등학교	
25	덕화여자중학교		59	보건환경연구원		93	파동사무소	
26	등공초등학교		60	삼육초등학교		94	파동초등학교	
27	동대구전신전화국 범어분국		61	상동사무소		95	파크호텔	
28	등도여자중학교		62	상동우체국		96	해화여고	
29	동문초등학교		63	소선여자중학교		97	황금관광호텔	
30	동부여자중학교		64	수성1동사무소		98	황금동사무소	
31	동산초등학교		65	수성1동파출소		99	황금동우체국	
32	동성초등학교		66	수성2동사무소		100	황금초등학교	
33	동원중학교		67	수성3동사무소		101	효목아파트	
34	동원초등학교		68	수성4동사무소				

인식할 경우는 인식물에 있어 OPDP알고리즘을 이용하여 인식한 후 그 결과를 출력하게 된다.

3) 검색결과

윈도우용 인터페이스를 Dragon Dictate 시스템을 참고하여 윈도우 컨트롤 컴포넌트를 이용하여 설계하고 이를 시험하였다. 설계한 인터페이스는 단순하지만 비교적 정확하고 안정적으로 동작하였다.

윈도우용 프로그램은 일단 구현하기가 까다롭지만, 제작 후에는 같은 윈도우환경의 시스템이라면, 어느 컴퓨터에서나 사용할 수 있기 때문에 여러 분야에서 이용 가능한 장점이 있다. 연구수행결과를 지면상에서 음성구동에 의한 동작으로 표현하는데는 제약이 수반되므로, 음성구동에 의한 수치지도의 정보검색 결과만을 그림 14와 그림 15에 나타내었다.

표 3. 48 Phoneme Likely Units

표기	음 소	표기	음 소	표기	음 소	표기	음 소
SIL	룩음	d	ㄷ(ㅌ)	ja	ㅈ	s	ㅅ
U	ㅡ	d+	ㄷ(ㅌ)	je	ㅊ	ss	ㅆ
W1	ㄱ	dd	ㄷ	p	ㅍ	t	ㅌ
aa	ㅏ	d~	ㄷ(ㄹ)	ju	ㅊ	uh	ㅓ
ae	ㅓ	eh	ㅓ	iv	ㅈ	wE	ㅓ
ao	ㅗ	g	ㄱ(ㅋ)	k	ㅋ	wa	ㅏ
axr	ㅑ	g+	ㄱ(ㅋ)	l	ㄹ(ㄴ)	we	ㅓ, ㅓ
b	ㅂ(ㅍ)	gg	ㅈ	m	ㅁ	wi	ㅓ
b+	ㅂ(ㅍ)	g~	ㄱ(ㄹ)	n	ㄴ	wv	ㅓ
bb	ㅃ	hh	ㅎ(ㅈ)	ng	ㅇ	z	ㅈ(ㅊ)
b~	ㅂ(ㅍ)	hh~	ㅎ	p	ㅍ	zz	ㅈ
ch	ㅊ	ih	ㅣ	r	ㄹ	z~	ㅈ

단, ㅌ : 첫음절 초성, ㅍ : 종성, ㄹ : 초성, ㅈ(ㅊ) : 445단어내 미사용

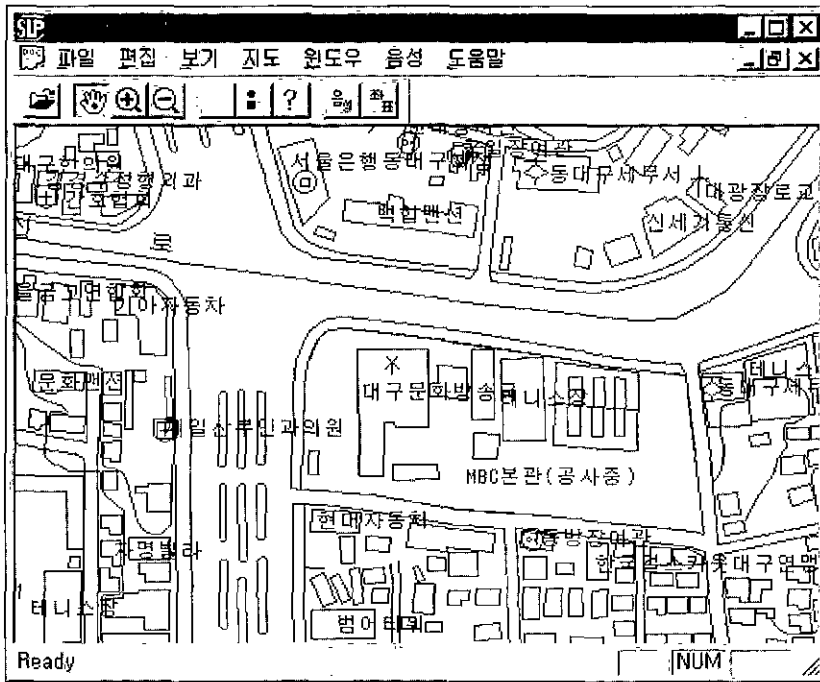


그림 14. "대구문화방송국"을 발성했을 경우

5. 결 언

최근에 가장 효율적인 공간분석기법으로 대두되고 있는 지리정보시스템(GIS)의 운용과정에서 지

표공간에 대한 정보를 마우스를 사용하여 각 위치에 대한 정보를 검색하여 왔었다. 그러나 음성구동에 의하여 지표에 대한 정보를 검색함으로써 보다 신속하고 효율적으로 수치지도상의 정보를 획득할

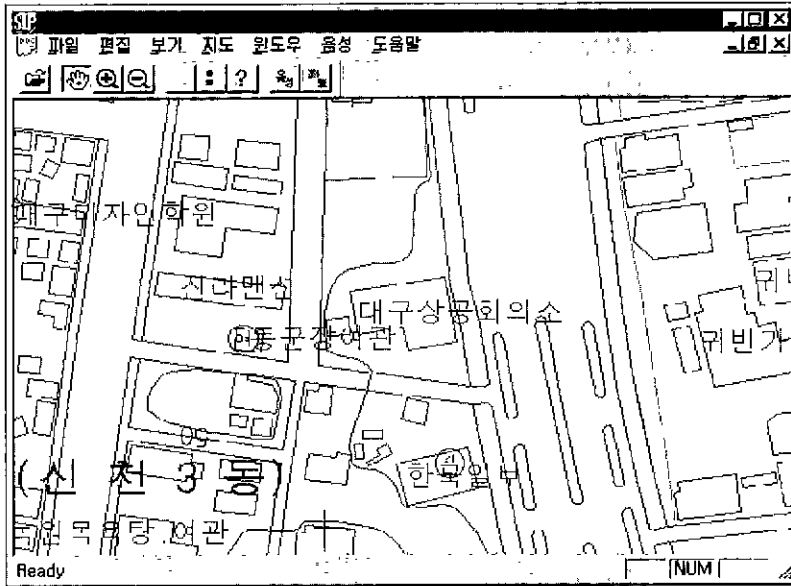


그림 15. "대구상공회의소"를 발성했을 경우

수 있음이 본연구를 통하여 가능하게 되었다. 이는 도시행정 및 여러분야에서 활용도가 높으며 특히 도시정보시스템에서 긴급한 재난 및 방재시스템에서는 필수적으로 요구되는 시스템으로 간주된다.

그러나 본연구에서는 점(point) 차원의 지형지물의 정보검색을 중심으로 한 실험적인 연구로서 대상지역내에서의 지명도가 높은 지형지물을 이미 데이터베이스화하고 특정이용자의 음성으로 음성 데이터베이스를 구축하였으나 향후로는 점차원 뿐만 아니라 선(line)과 면(area)의 검색과 주제도별 중첩분석기능을 수행할 수 있는 음성구동에 의한 공간분석시스템 개발이 바람직할 것이며 날로 발달되고 있는 멀티미디어 기능과 GIS를 통합한 보다 효율적인 종합적 정보시스템의 개발이 요망된다.

文 獻

- 김덕한, 1988, "음소 인식 단위의 Vector Quantization 과 Hidden Markov Model을 이용한 한국어 대응량 격리 단어 인식에 관한 연구." 한국 과학기술원 석사학위논문.
- 은종관, 1994, "연속분포 HMM을 이용한 한국어 연속인식시스템 개발", 한국 음향학회지 13(1), 24-31.
- 조명희, 1995, "원격탐사자료와 GIS를 이용한 라오스 남칸유역분지의 토지이용평가와 미작적지분석", 대한원격탐사학회지, 11(1),1-17.
- Anderson, J.M., 1994, Developing implementing and maintaining a map based computer-aided dispatch system for O'hare international airport, Chicago, Illinois, *URISA Conference Proceedings*, 695-705.
- Amstrong, M.P., Densham, P.J., and G.Rushton, 1992, Cartographic displays to support locational decision making, *Cartography and Geographic Information Systems*, 19(3), 154-164.
- Amstrong, M.P., Rushton, G., 1991, Decision support for regionalization : A spatial decision support system for regionalization service delivery system, *Computers, Environment and Urban Systems*, 15, 37-53.
- Beaument, J.R., 1989, Towards and integrated information systems for retail management,

- Environment and Planning*, 21, 299-309.
- Bottorff, H.R., 1995, Usage of GIS and GPS for emergency preparedness and mitigation, *GIS/LIS '95 ASPRS, ACSM Proceedings*, 78-85.
- Cartwright, W.E., and G.J.Hunter, 1995, Beyond maps : using multimedia to enrich the use of geographic information, *1995 Annual conference Proceedings, URISA II*, 420-433.
- Crawford S. C., and J. Dowdy, 1995, The Ohio Geographical Emergency Management System(OGEMS), *proceeding of GIS/CIS '95, ASPRS, ACSM*, 213-216.
- Dymon, U.J., and R.H.Platt, 1995, Mapping the geographic distributions of disasters, *Proceeding of GIS/CIS' 95, ARPRS, ACSM*, 297-301.
- Fung, D.S., and H.T.Kung, 1995, A GIS approach to natural hazard analysis, *Proceeding of GIS/CIS' 95, ASPRS, ACSM*, 338-345.
- Hin-Hui LEE and Biing-Hwang Juang, 1991, A Study on speaker adaptation of the parameters of continuous density hidden markov models *IEEE April*,
- Hirschfield, A., Brown, P. and P. Todd, 1995, GIS and the analysis of spatially-referenced crime data : experiences in mersey side, U.K., *INT.J. Geographical Information System*, 9(2), 191-210.
- Johnson, G.O., 1984 The value of computer methods for hazards vulnerability analysis, *1994 URISA Annual Conference proceeding Vol. I*, 661-674
- Lao Y., and H.P. sharma, 1995, A GIS approach to managiy gaseous hazardous spills on highways, *Proceeding of GIS/CIS' 95, ASPRS, ACSM*, 581-590.
- Nijkamp, P., and H.J. Scholten, 1993, Spatial information systems : design, modeling, and use in planning, *INT.J. Geographical Information System*, 7,(1), 85-96.
- Stein, A., Staritsky, I, Bouma, J., and J.W. Groenigen, 1995, Interactive GIS for environmental risk assessment, *INT. J. Geographical Information System*, 9, (5), 509-525.
- Timio TAKARA, Naoto MATAYOSHI, and Kasuya HIGA, 1994, Connected spoken word recognition using a many-state Markov Model, *ICSLP 94, YOKOHAMA*,