

음성인식을 이용한 차량 번호 조회 시스템의 구현

윤철중^{*}, 윤재선, 홍광석
성균관대학교 전기전자컴퓨터공학부 HCI연구실

An Implementation of Car Number Retrieving System with Speech Recognition

Chul-Joong Yoon, Jeh-Seon Yoon, Kwang-Seok Hong

HCI Lab, Electrical & Computer Engineering, Sungkyunkwan University

endxrein@ece.skku.ac.kr, sunhci@ece.skku.ac.kr, kshong@yurim.skku.ac.kr

요약

음성인식 기술은 사용자의 편리성을 제공하는 인터페이스로 많이 활용이 된다. 또한 음성이라는 특성상 새로이 기계를 작동하는 방법을 익히지 않아도 되며 빠르게 정보를 전달 할 수 있다.

본 논문에서는 음성인식 기술을 차량 번호를 조회하는 단말기에 적용하였다. 이것은 기존의 단말기보다 입력이 간편하여 사용자에게 편리함을 제공한다. 또한 잦은 오류를 피할 수 있으며 오류가 발생했다 하더라도 쉽게 수정할 수 있는 기능을 제공한다.

I. 서론

음성은 인간에게 있어서 가장 편리한 정보전달의 도구이며 의도하는 내용을 빠르게 전달 할 수 있는 도구이기도 하다. 정보화 사회가 이루어지면서 정보의 양이 많아지고 또 대용량 정보 이동의 빠른 전달이 요구되고 있다. 이러한 점에서 음성을 이용한 빠른 전달과 편리한 사용의 요구는 더욱 커지게 되며 그 필요성 또한 증대되고 있다.

기계가 인간의 음성을 인식하는 음성인식 기술은 이미 광범위한 응용분야를 이루고 있다. 휴대폰에 전화번호를 대체하는 방법으로써 사용된 분야에서부터 컴퓨터 응용

프로그램의 음성인식에 이르기까지 적용분야는 한계가 없다고 할 수 있다.

본 논문에서는 음성인식 기술을 차량의 번호를 조회하는 단말기에 적용하였다. 차량의 번호조회에 사용되는 단말기는 사용자가 느끼기에 불편할 만큼의 작고 많은 수의 자판 때문에 입력의 오류와 사용의 불편을 야기한다. 본 시스템에서는 지금까지의 단말기에서 지적되어 왔던 입력의 오류와 불편을 줄이고자 음성을 사용하여 정보를 전달한다. 음성으로 입력할 수 있는 분야는 차량 번호 조회, 운전면허번호 조회, 주민등록번호 조회이다.

기존 단말기의 많은 수의 자판을 눌러 정보를 입력하는 것보다 음성인식 기술을 이용한 단말기는 빠른 속도와 정확한 입력을 가능하게 한다. 그리고 여기에 보완하여 음성 입력 시 오류가 발생할 경우 정정할 수 있는 버튼을 두었다. 오류가 발생한 부분을 수정 가능하도록 한 것인데 이 버튼만 이용하여서도 정보의 입력이 가능하도록 하였다. 기존의 단말기에 비해서 음성인식기술을 적용하여 사용자에게 간편한 인터페이스를 제공하였으며 자판의 수도 대폭적으로 줄어들게 되었으므로 오류 발생의 소지도 적을 것으로 기대된다.

II. 차량 번호 조회 시스템

시스템의 특징으로는 기존의 버튼으로 입력을 하는 방식에서 음성을 이용한 입력방식이라고 할 수 있다. 이것

은 버튼을 일일이 누르는 것보다 빠른 입력이 가능하며 입력되는 버튼의 수가 적으므로 입력 시 오류발생의 소지를 줄일 수 있다. 또한 버튼으로 직접 원하는 번호를 누르던 방식에서 Up, Down버튼을 조작하여 데이터를 입력하는 간접적인 방법을 채택하고 있다. 이것은 음성 인식의 방법을 채택함으로써 조작을 위한 버튼의 수가 감소하여 많은 버튼을 일일이 누를 필요성이 적고 음성 인식 시 오류가 발생하면 오류를 수정하기 위한 방법으로 채택하였다. 이것은 입력하는 방식중 보조적인 방법이 된다. 이렇게 음성인식 방법을 채택하여 조작을 위한 버튼수가 적어져 오류의 발생 소지를 줄였다.

본 시스템은 차량번호조회, 운전면허조회, 주민등록번호를 조회하는 기능을 가지고 있다. 각각 조회를 한 후 입력된 데이터와 비교를 하여 결과 값을 출력하는 과정이 있으며 이 과정은 추후 데이터베이스와 연동에 의해서 확장될 수 있다. 본 시스템에서는 임의의 대상을 저장하고 이것과 조회를 하였다.

차량번호의 인식절차로 차량 번호는 "서울 12 나 1234"로 구성되어 있으므로 지역명 인식, 숫자인식, 단어인식, 숫자인식의 네 가지의 인식대상을 가지고 있다. 운전 면허 번호의 경우는 "서울 12 123456 12"와 같은 형태이므로 지역명 인식, 숫자인식, 숫자인식, 숫자인식의 네 가지의 인식대상을 가지고 있으며 마지막으로 주민 등록 번호 인식은 "123456 1234567"과 같이 숫자인식의 두 가지 대상을 가지고 있다.

인식에 필요한 단어는 지역명 서울, 인천, 대전, 대구, 부산, 광주, 울산, 경기, 충북, 충남, 전북, 전남, 강원, 경북, 경남, 제주 16개, 자음 ㄱ(기역), ㄴ(니은), ㄷ(디귤), ㄹ(리을), ㄹ(미음), ㅂ(비읍), ㅅ(시옷), ㅇ(이응), ㅈ(지읒), ㅊ(치읓), ㅋ(키읓), ㅌ(티읕), ㅍ(피읖), ㅎ(히읇) 14개와 모음 ㅏ(아), ㅑ(어), ㅓ(오), ㅕ(우), ㅡ(으) 5개의 총 19개와 숫자음 품, 일, 이, 삼, 사, 오, 육, 칠, 팔, 구의 10개 총 45개의 단어를 인식하고 있다. 표 1에서는 각 인식 단계별 사용되는 인식 데이터의 분류를 나타내었다.

III. 음성인식

음성인식 시 먼저 인식데이터의 분류와 인식 단계를 결정, 인식 방법을 결정한다.

인식데이터 분류는 숫자음 인식대상 10개, 지역명 16개, 자음과 모음 19개로 나눈다. 다음으로 인식 단계를 분류하게 되는데 차량번호인식에서는 지역명, 숫자음, 자모음이 인식이 되어야하고 운전 면허 번호 인식에서는 지명과 숫자음, 그리고 주민 등록 번호 인식에서는 숫자음이 인식되도록 단계를 분류하였다.

인식에 사용된 데이터는 70명 남성화자가 발생하였으

표 1 state별 인식 데이터

state 1	state 2	state 3	state 4	state 5
차량번호 인식	지역명	숫자 2자리	자음,모음	숫자 4 자리
운전면허 번호인식	지역명	숫자 2자리	숫자 6자리	숫자 2자리
주민등록 번호	숫자 6자리	숫자 7자리	.	.

며, 11.025kHz 샘플링 주파수, 16bit 양자화를 하였다. 인식 알고리즘으로써는 K-Means Algorithm과 DHMM을 사용하였고 K-Means Algorithm level은 128로 하였다.

3.1 인식모델

인식에 사용한 모델은 단어 모델, 음절 모델, 반음절 모델로 나누어진다. 단어 모델은 지명과 자음을 인식하는데 사용하였다. 입력받은 음성 데이터의 시작점과 끝점을 추출하여 훈련데이터로 사용하였으며, 음절모델은 모음과 숫자음의 경우, 시작점과 끝점을 추출한 후 데이터를 저장하였고 지명 중 충북, 충남, 전북, 전남, 경북, 경남의 인식률을 높이기 위하여 "충", "전", "경" 음절을 규칙에 의해 분할 저장하였다. 반음절 모델은 숫자음의 안정된 모음 구간에서 분할하여 시작점에서 분할 구간까지는 CV, 분할 프레임과 끝점까지는 VC로 저장하였으며, 모든 자음 데이터는 "ㄱ"은 "기역", "ㄴ"은 "니은" 과 같이 발생하게 되므로 앞쪽 음절 모음은 "이"모음이 된다. 따라서 규칙에 의해 시작점과 앞쪽 모음의 안정 구간, 즉 CV 데이터를 분할하여 저장하였다. 반음절+음절형태의 모델은 자음의 데이터의 앞쪽 모음의 안정구간을 시작점으로 자음데이터의 끝점 즉 V+CVC의 모델을 구성하였다.

3.2 인식절차

(1) 차량 번호 인식

차량 번호 인식은 지명, 숫자음 1자리 또는 2자리, 자음, 모음, 숫자음 4자리의 순으로 인식한다. 지명과 숫자음의 인식방법은 각 자리의 데이터를 입력받은 후 안정 구간을 검출하여 앞쪽 반음절과 뒤쪽 반음절로 구분하여 각각 반음절 모델과 비교하여 결정한다. 지명 다음의 reference 후보는 숫자음 10개이나 다음의 인식 후보는 숫자음이 올 수 있고 자음이 올 수도 있기 때문에 인식 후보의 수는 숫자음 10개와 자음 14개 총 24개가 온다.

숫자음 인식방법은 숫자음 분할규칙을 적용하여, CV, VC 반음절로 구분한 후, reference model과 비교하여 결정한다. 그리고 자음 인식방법은 자음의 분할 규칙을 적용하여 앞쪽 CV 단위와 뒤쪽 V+CVC 단위로 구분하고 각각 reference model과 비교하여 확률 값을 구한 후 두 방법에 의해 결정된 최소의 확률 값을 인식결과로 결정한다.

(2) 운전 면허 번호 인식

운전 면허 번호의 인식절차는 지명, 숫자음 2자리, 6자리, 2자리의 순으로 인식을 한다. 지명을 인식하는 모델은 단어 단위의 지명 reference model과 비교한 후, 인식된 지명이 "충북, 전북, 경북, 충남, 전남, 경남"일 경우에는 입력 데이터에 규칙을 적용하고 앞쪽 음절을 분할하여 "충, 전, 경"의 음절 reference model 과 비교하여 결정한다. 그 이외의 지명인 경우 단어 단위의 인식 결과 값에 의해 인식을 결정한다. 숫자음 인식은 차량 번호 인식의 방법과 동일하다.

(3) 주민 등록 번호 인식

주민 등록 번호는 6자리, 7자리의 숫자음으로 구성되어 있으며, 인식방법은 각 자리의 데이터를 입력받은 후 안정구간을 검출하여 반음절 모델과 비교하여 결정한다.

IV. 차량 번호 음성인식 시스템의 구현

인식 시스템은 윈도우즈 응용프로그램으로 시뮬레이션을 하였고 실제 구현했을 때와 동일한 형태로 구성되어 있다. 시스템의 구동 시 그림 1과 같은 초기 화면이 나타난다. 시스템은 조회하고자 하는 부분을 선택할 수 있게 되어 있으며 초기에는 차량 번호 인식이 기본으로 선택되어 있다.

(1) 차량 번호 조회

"1. 차량 번호 인식"을 선택하고 "OK"버튼을 누르면 그림 2와 같은 인식 형태로 바뀌게 된다. 여기에서 지역명과 차종 번호, 자음 모음, 차량번호를 순차적으로 발생하고 오류가 발생하였을 때에는 "LEFT, RIGHT" 버튼을 눌러 오류가 발생한 곳까지 이동시킨 후 "UP, DOWN"버튼을 눌러 수정을 한다.

그림 3은 입력된 데이터와 저장된 데이터를 비교하여 결과를 나타낸 그림이다. 입력된 차량번호에 관한 정보를 볼 수 있으며 "OK"버튼을 누르면 다시 처음으로 돌아간다. 입력 데이터가 수배 데이터와 같다면 그림 4와 같이 수배번호라고 결과를 나타낸다.

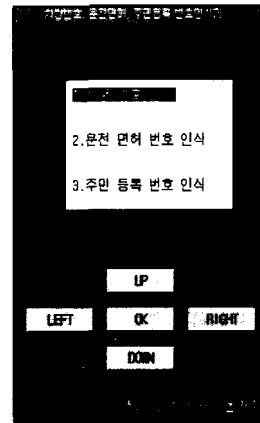


그림 1 차량번호 인식시스템

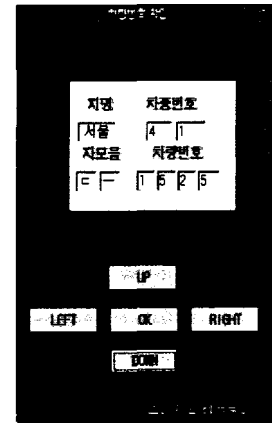


그림 2 차량 번호 인식 형태

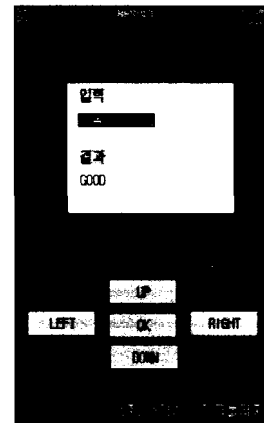


그림 3 차량 번호 조회 결과

(2) 운전 면허 번호 조회

초기 화면에서 "2. 운전 면허 번호 인식"을 선택하고 "OK" 버튼을 누른다. 차량 번호 조회와 마찬가지로 두 번째에서는 운전 면허 번호를 입력할 수 있는 그림으로 바뀐다. 오류가 발생했을 시 수정 방법은 차량 번호 조회의 방법과 동일하다. 마지막에는 입력한 운전 면허 번호에 대한 결과를 보여준다.

(3) 주민 등록 번호 조회

초기 화면에서 "3. 주민 등록 번호 인식"을 선택하고 "OK" 버튼을 누른다. 앞의 두 과정과 마찬가지로 두 번째는 주민 등록 번호를 입력할 수 있고, "OK" 버튼을 누르면 입력한 번호에 대한 결과를 알 수 있다.

(4) 인식성능

테스트에는 남성화자 5명이 일반 사무실 환경에서 2번씩 발생한 결과를 가지고 인식률을 측정하였다. 표 2에

V. 결론

본 논문에서는 음성인식 기술을 차량 번호 조회단말기에 적용하였다. 이 시스템은 기존의 손으로 일일이 입력을 해야 하는 불편을 해소하고 입력의 오류를 줄여 빠른 검색을 가능하게 한다. 또한 입력 시 오류의 수정도 간단하다. 이러한 장점을 가지고 있어 사용자에게 편리한 환경을 제공한다.

인식성능에 있어서도 자음인식률을 제외한 나머지 부분은 인식률이 높게 나타난다. 자음 부분의 인식률이 낮은 이유는 자음의 조음위치가 같은 곳에서 나는 자음은 구별이 어렵기 때문이다. “ㄸ”, “ㄲ” (두 입술 사이), “ㄷ”, “ㅌ”, “ㄴ”, “ㄹ” (혀끝과 윗잇몸), “ㄱ”, “ㅋ” (연구개), “ㅈ”, “ㅊ” (경구개)과 같은 소리는 그 구별이 어렵다. 향후 이러한 자음을 구별 할 수 있는 특징을 찾아내어 자음 부분의 인식률을 높일 수 있다면 더욱 효율적인 시스템이 될 것이라고 생각한다.

현재 DSP를 이용한 인식 단말기를 제작중에 있다.

참고문헌

- [1] 윤재선, 홍광석, “반음절 단위 HMM을 이용한 연속 숫자 음성인식”, 한국음향학회지 제 17권 제 5호, pp.73-78, 1998.
- [2] 윤재선, 최광표, 홍광석, “ Windows환경에서 음성인식을 이용한 멀티미디어 게임의 구현”, 한국음향학회 학술발표대회 논문집 제 17권 2호, pp335-338, 1998.
- [3] 김순협 외 4인, “음소 단위에 의한 한국어 연속 숫자 음 인식에 관한 연구”, 한국음향학회지 제 8권 3호, pp.5-15, 1989

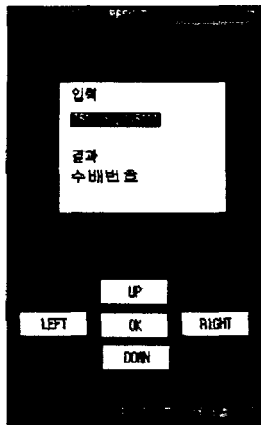


그림 4 수배 번호 인식결과

차량 번호 인식, 운전 면허 번호 인식, 주민 등록 번호 인식의 결과와 전체적인 인식률을 나타내었다. 인식 결과를 구하는 과정에서 지역명을 발성한 경우는 차량 번호 인식과 운전 면허 번호 인식의 경우에만 있었고, 자음과 모음을 발성하는 경우는 차량 번호 인식의 경우만 존재한다. 따라서 지역명과 자음, 모음의 발성 횟수가 숫자음에 비해 상대적으로 적었다. 자음의 경우 인식률이 낮은 것은 발음이 나는 곳이 같은 자음들 즉 조음위치가 같은 자음들 사이에서 오인식이 많았기 때문이다. 두 입술 사이에서 나는 소리인 “ㄸ”, “ㄲ” 혀끝과 윗잇몸이 만나서 소리가 나는 음인, “ㄷ”, “ㅌ”, 경구개와 혀바닥 사이에서 나는 소리인 “ㅈ”, “ㅊ” 그리고 연구개에서 나는 소리인 “ㄱ” “ㅋ”와 같은 자음과 조음위치가 같고 울림 소리인 “ㄴ”, “ㄹ”인 경우는 그 구별이 매우 어렵다.

표 2 인식률

(단위:%)

인식후보 종류	지역명	자음	모음	숫자음
차량번호	90	70	100	96.6
운전면허번호	100	.	.	97
주민등록번호	.	.	.	97.6
총 계	95	70	100	97.1