

퍼지를 이용한 음성 인식 시스템의 설계

○
장형중*, 임준식*
*경원대학교 전자계산학과

Design of Speech Recognition System Using Fuzzy

○
Hyoung-Jong Jang*, Joon-Shik Lim*
*Dept. of Computer Science, Kyungwon Univ.

요 약

본 논문에서는 전화나 교환기에서, 사람의 음성으로 이름을 식별하는 시스템을 구축하고 있다. 기존의 음성 인식 방식은 DTW, HMM, LPC, NN 등의 방식을 사용하였으나, 본 논문에서는 퍼지를 이용하고 있다. 퍼지를 이용하여 음소에 관한 룰베이스를 작성하고, 이를 토대로 가장 유사한 이름을 찾아준다. 즉 성을 모르고 이름만 안다고 하여도 목록에 있는 사람의 이름 중에서 가장 근접한 사람을 찾아줄 수 있도록 한다. 또한 음소 단위의 인식을 통하여 인식할 이름의 수가 증가한다고 하여도 별도의 인식 작업이 없이 첨가가 가능하도록 하고 있다.

1. 서 론

음성 인식은 인공 지능의 다른 여러 분야같이 인간의 감각, 사고, 학습 능력을 이해하고 모방하여 컴퓨터로 구현하고자 하는 응용 분야 중 하나이다.

본 연구의 연장된 목표는 전화를 이용한 음성 교환기 설계이다. 기업이나 연구실에 전화를 걸었을 때, 내부 번호를 모르면 교환을 통해야 하는 불편이 있고, 교환원을 따로 두거나 사원들이 일일이 전화를 바꾸어 주어야 한다. 이때, 교환기가 사람의 이름을 인식한다면 간단히 해결할 수 있다.

이런 목표를 위해 사람의 이름을 인식하는 시스템 구축하고 있다. 전화 시스템을 감안하여 입력 값은 8kHz 대역으로 한정하였고, 실제로는 마이크로폰을 이용하였다. 또한, 현재의 음성 인식 장비가 고가의 Unix 계열의 장비를 사용하고 있는데, 이를 상대적으로 저렴한 Pentium 급 컴퓨터와 간단한 Sound Card에서 운영이 가능하도록 하였다.

본 논문은 1장 서론에 이어, 2장 연구 배경에서는 인간의 발성 기관과 발생된 음성의 특징 그리고 우리말에서의 음소 구성과 인식 단위로서의 음소들에 대하여 설명하였고, 3장 관련 연구에서는 음성 인식에 관한 설명을 하였다. 4장에서는 본 시스템의 설계에 대한 설명이 있으며, 5장 결론에서는 예상되는 결과와 추후 과제에 대한 제시로 끝을 맺었다.

2. 연구 배경

2.1 음성 신호

음성 신호는 인간의 발성 기관에 의해서 생성된다 [1]. 음성의 생성과정은 음원의 생성, 발생기관의 조음, 음의 방사의 세 부분으로 구분 할 수 있다.

음성은 연속성, 개별성, 동일성이라는 특징을 가진다. 음성은 입에서 만들어지고 귀에서 듣게 되므로 음성이 존재하는 구간은 입에서 귀까지의 공간이라 할 수 있다. 이 때, 존재 양상은 연속적인 음파로서

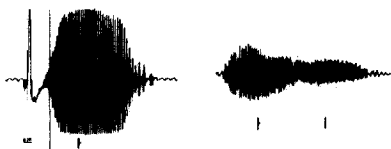
나타난다. 또한, 일회적이며 개별적인 존재 양상도 갖는다. 어떤 사람이 똑같은 말을 반복했다 할지라도, 처음 발음한 것과, 나중 발음한 것은 그 자체가 같을 수 없고 단지 같다고 확인하는 것은 우리의 귀일 뿐이다.

무한히 많은 음성은 다른 점보다 더 많이 서로 같은 점을 가지고 있다. 이러한 점들을 음성의 동일성이라고 부르며, 동일성은 바로 음성이 발음되는 조음위치와 조음방식에 기인한 것으로 보통 같은 말이라고 말할 때는 이 위치와 방식의 동일성에 근거를 두고 하는 말이다. 음성은 그 개별성에서 보면 무한히 존재하지만 그 동일성에서 보면 몇 개 안되는 것이 반복 사용되고 있는 것이다.

2.2 음소의 구분

단어를 인식 단위로 하는 인식 시스템은 하드웨어의 제한된 메모리와 계산 속도로 인하여 이들 시스템의 실시간 처리 시스템의 구현을 어렵게 하고 있다는 것이다. 그러므로, 단어보다 더 기본적인 음성 단위들, 다시 말하면, 음소 등을 인식단위로 하여 인식을 수행하는 연구들[2][3][4][5][6]이 활발하게 이루어지고 있다. 특히 40여개의 음소[7][8]를 갖는 한글에서는 음소들을 인식단위로 사용함으로써 이러한 한계들을 극복할 수 있을 것이다.

음소의 경계 추출은 Basseville과 Benveniste, Andre-Obrecht 들에 의하여 여러 가지 알고리즘들이 연구되었다.[9][10] 그러나 이러한 경계 추출은 그림1 (a)에서의 예('타')와 같이 무성음('ㄷ')과 유성음('ㅏ')의 음소 사이에서는 뚜렷하게 나타나는 경우가 있지만, (b)에서의 예('아이')와 같이 유성음('ㅏ')과 유성음('ㅣ') 혹은 유성음과 무성음들 사이에서 상호간의 조음현상 때문에 정확한 경계를 찾기가 어려울 뿐만 아니라 이러한 경계부근의 천이구간에서의 음성 샘플들은 음소 단위 인식에서는 별다른 의미를 갖지 못한다. 따라서 이러한 음소의 경계를 정확히 찾는 것보다는 특성의 변화가 적은 안정된 대표구간을 찾는 것이 더 효율적이라 할 수 있다.



(a) 무성자음과 모음 (b) 모음과 모음
 그림 1. 음소 경계의 예

2.3 음소의 구분

2.3.1 우리말의 음운론적 음소 구성

우리말의 음운은 자음과 모음으로 설정된다.

(1) 자음

우리말 자음에 속하는 음운들은 조음점과 조음 방법에 따라서 조음되며 반모음을 제외하면 총 19개의 자음들이 있다. <표1>은 조음 계열에 속한 자음들을 보인 것이다. 자음 음운들의 쓰이는 위치를 보면 받침으로 쓰이는 자음은 'ㄱ', 'ㄴ', 'ㄷ', 'ㄹ', 'ㅁ', 'ㅂ', 'ㅇ' 이렇게 7자이다.

<표 1> 자음 계열 음운들의 조음

조음방법 \ 조음점	양순		치경		연구개		성문
	양순	치경	구개 (파찰)	구개 (마찰)	연구개		
단순	p(ㅍ)	t(ㅌ)	c(ㅊ)	s(ㅅ)	k(ㄱ)		
긴장	pp(ㅃ)	tt(ㅆ)	cc(ㅉ)	ss(ㅆ)	kk(ㄲ)		
대기	ph(ㅑ)	th(ㅓ)	ch(ㅕ)		kh(ㅋ)		
비강	m(ㅁ)	n(ㄴ)			ɱ(ㅇ)		h(ㅎ)
설측		l(ㄹ)					

(2) 모음

모음에 속하는 음운들은 두 입술의 모양, 혀의 높낮이 및 혀의 앞과 뒤에 따라서 조음된다. 모음의 구조적 계열을 <표2>에 보였다.

<표 2> 모음 계열 음운들의 조음

혀의 앞뒤 입술 모양	전설		후설	
	평순	원순	평순	원순
혀의 높낮이				
고	ㅣ	ㄱ	ㅡ	ㅜ
중	ㅑ	ㅓ	ㅕ	ㅛ
저	ㅗ		ㅝ	

반모음 'w'와 'y'들은 'w'의 경우는 조음은 'ㅜ'와 거의 같고, 'y'의 경우는 조음이 'ㅣ'와 거의 같으나, 이들은 독립적으로 쓰이는 것이 아니라 항상 다른 모음 앞에서만 쓰이므로 반모음(semi-vowel)이라 부른다.

2.3.2 음소 인식 관점에서의 음소

이들 변이음들은 조음점 변이에 의한 변이음, 조음 방법의 변이에 의한 변이음으로 구분할 수 있는데, 조음 방법의 변이에 의한 변이음은 한 음소가

다른 음소로 음의 변동이 일어나는 경우이므로 음소 단위의 인식에서는 고려 대상에서 제외되고, 조음점 변이에 의한 변이음만이 고려 대상이 된다.

(1) 모음

앞에서 살펴본바와 같이 우리말의 음소 수는 복모음을 제외하면 총 29개이고, 변이음들을 포함하는 경우는 총 59개이다. 이들 음소는 복모음 10개(혹은 12개)를 포함하지 않는 것으로 복모음을 반모음+단모음(ㄱ, ㅋ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ) 혹은 단모음+반모음(ㅡ)이라는 관점에서 분리하여 인식할 경우, 이들 복모음은 음소의 수에서 제외될 수 있다. 또한 단모음 중 'ㄱ'과 'ㄱ'의 경우는 다차원 척도상의 거리가 매우 근접하여 있어[11] 단모음의 인식에서는 동일 모음으로 취급하고, 상위 레벨의 처리에서 구문론적인 분석에 의하여 구분하는 것이 더 효율적이라고 판단되어 'ㄱ'의 동일 모음으로 취급하도록 한다. 따라서, 'ㄱ', 'ㄱ'등도 복모음으로 간주할 경우 모음 음소의 수는 단모음 7개 즉, 'ㅏ', 'ㅑ', 'ㅓ', 'ㅕ', 'ㅗ', 'ㅛ', 'ㅜ', 'ㅠ'와 반모음 'y', 'w'의 2개가 된다.

<표 3> 인식 단위로서의 모음

	단모음	반모음
인식 단위로서의 음소	ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ, ㅚ, ㅜ	y, w

(2) 자음

조음점과 조음 방법은 일치하나, 쓰이는 위치에 따라 다른 특성을 갖도록 조음된다. 하지만, 본 논문에서는 사람의 이름을 인식하는 경우에 한하므로 조음 방법 중 긴장에 해당하는 음은 인식 단위에서 삭제하고, 특성상 일정정도 고립음에 가까우므로 실시간 처리를 위하여 'ㄹ', 'ㄴ'을 'ㄹ' 하나로 인식하면, 인식 단위로서의 자음 음소는 <표4>에서와 같이 20개가 된다.

<표 4> 인식 단위로서의 자음 계열 음운들의 조음

조음점 조음방법	양순 음	치경 음	치경 구개 (파찰)	치경 구개 (마찰)	연구 개	성문
단 순	ㅍ(ㅍ)	ㅌ(ㄷ)	ㅈ	ㅊ	ㅋ(ㅋ)	
대 기	ㅍ	ㅌ	ㅈ		ㅋ	ㅎ
비 강	ㅁ(ㅁ)	ㄴ(ㄴ)			(ㅇ)	
설 측		ㄹ(ㄹ)				

3. 관련 연구

음성 인식의 기술은 최근 크게 부각되는 기술로서 적용할 수 있는 범위가 매우 넓다. 특히 통신 분야는 음성인식 기술이 사용된다면 그 장점이 더욱 크게 발휘될 것이다.[12]

3.1 음성인식 시스템의 분류

음성 인식은 사용하는 기술의 종류에 따라 몇 가지로 분류할 수 있다.

첫째로 사용자의 형태에 따라 누구나 인식 가능한 화자 독립 인식 시스템과 특정 화자만을 인식 할 수 있는 화자 종속 인식 시스템을 구성할 수 있다.

둘째로 화자의 발성에 따라 단어 단위로 발성하여 인식하는 고립단어 인식, 단어들을 연결하여 발성하는 연결단어 인식, 화자의 말(문장)에서 특정 단어만을 인식하는 핵심어 인식, 그리고 평상시 말하는 자연스러운 대화체를 인식할 수 있는 연속음 인식 등으로 분류할 수 있다.

마지막으로 인식 가능한 어휘 수에 따라 대용량 단어인식과 소용량 단어인식이 있다. 소용량의 단어를 인식하는 경우는 단어 단위로 할 수 있기 때문에 고립어나 연결단어 인식 등에 사용되어 왔다. 그러나 대용량 단어는 수천 단어에서 수만 단어를 인식하여야 하기 때문에 이러한 방법으로는 곤란하다. 그래서 음소 단위나 그와 비슷한 단위로 인식하는데 그 방법이 어려워 아직까지 많은 제품들이 있지는 못하다.

3.2 음성인식 시스템의 구성

모든 음성 인식 시스템들이 다음과 같은 구성을 가지고 있다. 우선 입력된 신호 중에서 실제 사람이 발성한 음성 신호만을 검출하여 특징을 추출하고 이것으로 룰 베이스와 비교를 측정해 최종적으로 인식하게 되는 것이다. 그러나 시스템마다 전처리 단계를 가지거나 조금씩 다른 특징들을 추출함으로써 서로 다른 특성을 지니는 시스템을 구성하게 되는 것이다.

음성 신호를 입력받게 되면 실제 화자가 발성한 음성 부분만을 검출하여야 한다. 아무리 좋은 음성 인식 알고리즘을 사용한다하더라도 음성 검출이 제대로 이루어지지 않았다면 좋은 인식률을 기대하기 어렵기 때문에, 음성 인식에 앞서 정확한 음성 검출이 요구된다.

실음성 구간을 검출한 다음에는 음성 특징을 추출

한다. 이 부분은 일종의 음성 압축 부분으로서 불필요하게 중복되는 음성 정보를 없애고 동일 음성 신호들간의 일관성을 높임과 동시에 다른 음성 신호와는 변별력을 높일 수 있는 정보를 추출한다.

이렇게 구한 특징 계수들을 가지고 실제 인식에 들어가게 되는데 이때 사용되는 음성인식 알고리즘에는 여러 가지가 있으나 대표적인 것으로 Vector Quantization, Dynamic Time Warping, Neural Network, Hidden Markov Model 등이 있다.

그리고 최종적으로 인식한 결과가 올바른 것인지를 확인하는 부분이 있어야 한다. 위의 과정을 거쳐 인식한 음성이 기대하는 말이 아니었을 경우 이를 오인식하는 것이 아니라 거절할 수 있어야 한다. 즉, 잘못된 음성 입력에 대한 검증 부분까지 포함되어야만 신뢰성이 있는 시스템이라고 할 수 있는 것이다.

3.3 음성 인식 시스템의 조건

음성인식이란 간단히 말해 사람의 말을 기계에게 전달해 원하는 동작을 유도하는 것이다. 결국 사람과 기계의 자연스러운 통신이 궁극적인 목적이라고 볼 수 있다. 따라서 불특정 화자가 언제, 어디에서 무슨 말을 하더라도 이를 인식하는 것이 최종적인 목표이다. 다시 말해, 사용환경에 영향을 받지 않는 화자 독립 무한 연속음 인식이 되어야만 한다. 또한 이러한 시스템을 사용하는 사용자의 부주의로 인한 영향도 고려해야 한다. 그러나 이러한 것은 최상의 조건으로 아직까지 이러한 제품은 없다.

하지만 사용환경에 적합한 시스템을 고려할 수는 있다. 사람과 사람의 통신에도 100% 인식이 불가능하다고 한다. 실험에 의하면 인간의 고립 단어 인식 능력은 98%라고 하는데 이 것과 비슷한 정도의 인식 성능을 가진 시스템이라면 사용자가 불편함 없이 사용 가능하다고 볼 수 있다. 그리고 인식 대상 어휘도 사용 목적에 따라 다르므로 제한된 어휘 내에서 서비스 가능한 시스템을 구성할 수 있다.

4. 설계

시스템의 전체 구조는 <그림2>와 같다.

음성 수집 장치는 음성을 디지털 신호로 변환한다. 이 때, 전화 음성에 대비하여 표준적인 전화 통신 관행인 8kHz를 사용하며, 16-bit로 샘플링한다.

디지털이 된 음성 신호는 저차수의 디지털 시스템에 입력된다. 이렇게 함으로써 신호를 스펙트럴 측면에서 평평하게 하여 이후의 신호 처리 과정에서

정확도가 떨어지는 것을 어느 정도 방지 할 수 있다. 직접적인 효과는 배경 잡음을 감소시키고, 신호 스펙트럼을 평준화하는 것이다. 프리엠퍼시스에 가장 많이 쓰이는 시스템은 고정 1차 시스템이다.

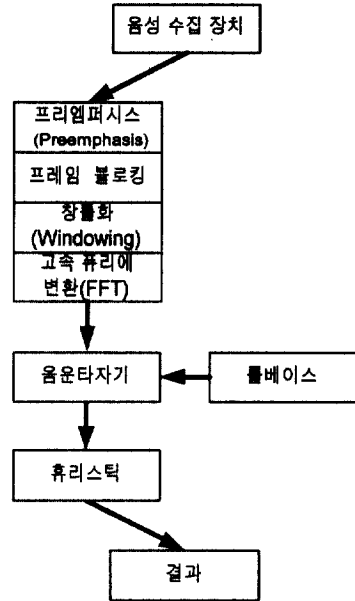


그림 2. 전체 시스템 구성도

프리엠퍼시스 필터는

$$H(z) = 1 - \alpha z^{-1}, \quad 0.9 \leq \alpha \leq 1.0 \quad (1)$$

을 사용하며, 입력 샘플에 관하여서는

$$\tilde{s}(n) = s(n) - \alpha s(n-1) \quad (2)$$

의 차함수로 표현된다. 가장 보편적인 α 값인 0.95를 사용한다.

프레임 블로킹 단계에서는 음성신호 $\tilde{s}(n)$ 을 M 샘플 간격으로 N 샘플씩 프레임으로 블록화한다.

창틀화(windowing) 단계에서는 구분된 각각의 프레임들이 시작과 끝에서 갖는 신호의 불연속성을 최소화하기 위하여, 다음 식과 같이 창틀 함수 $\omega(n)$ 을 신호에 곱하여 줌으로써 각 프레임의 시작과 끝에서 0에 점점 가까워지도록 창틀화 한다.

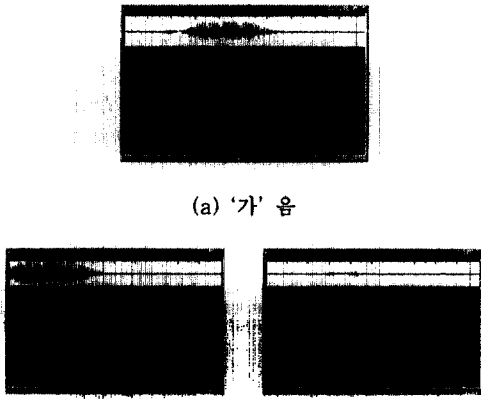
$$\tilde{x}_i(n) = x_i(n)\omega(n), \quad 0 \leq n \leq N-1 \quad (3)$$

본 논문에서는 다음과 같은 방법을 사용한다.

$$\omega(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n \leq N-1 \quad (4)$$

고속 푸리에 변환(Fast Fourier Transformation)을 실시한다.

다음 과정은 툴 베이스를 만든다. 각 음소 별로 기준 모델을 10ms 단위로 10개씩 만든다. 여기서 10ms를 단위로 한 것은 모음의 경우 일반적 발음이 길지만, 자음의 경우 특히 무성음의 경우는 50~60ms로 짧기때문이다. 총 29개의 음소와 공백 즉 300개의 기준모델을 작성한다. 여기서 자음의 예를 들면 '가'의 경우 '가'에서 '아'를 제거하고 단위 별로 기준 모델을 작성한다. 예를 보면 <그림3>과 같다.



(a) '가' 음 (b) '가' 음 추출 (c) '가' 음 추출 후
그림 3. 툴베이스 작성 예시

<그림3>의 (a)는 초기 '가'음에 해당하는 샘플이고, (a)에서 '가'음만을 추출한다. 이때 남은 음을 가지고, '가'에 해당하는 음을 샘플링한다. 이때, '가'음의 경우 '가'와 '거'의 경우 자음과 모음 사이의 차이 구간이 존재하므로 여러 모음에서 샘플링 작업을 한다.

입력되는 값을 10ms로 나누어 툴 베이스와 비교한다. 같은 값이 없을 경우 최대한 유사한 값이 나오는 모델로 선택한다. 이때 음운 타자기를 사용한다.[13]

음소 인식 과정에서는 불완전할 수 있는 음운 타자기의 결과를 사용하여 정확한 인식을 유도하는 다음과 같은 휴리스틱을 사용한다. 그림으로 표현하면 <그림4>와 같다.

음성 인식기에서 음운 타자기에 음성 파일을 전파하면 철자법대로 완벽하게 인식된 결과가 출력되는 것이 아니다. 그러므로 음운 타자기의 결과를 가지고 이름을 인식하기 위한 방법이 필요하다.

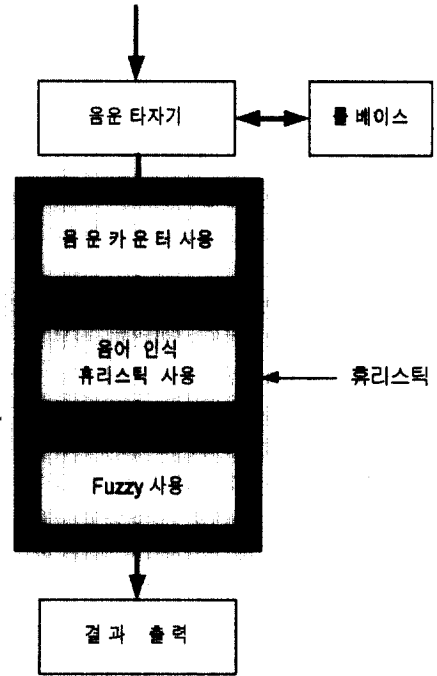


그림 4. 휴리스틱

이에 대한 휴리스틱은 다음과 같다.

음소 인식을 위해 동일 음소 전파의 연속성을 측정하는 단순 카운터를 사용한다. 자음의 경우 파열음의 발음 길이는 다른 자음보다 짧아 모든 자음을 단순한 카운터만을 사용하여 인식한다면 파열음은 인식이 잘 되지 않는다. 그래서 파열음이 전파되는 중간이나 후에 'ㅎ'이 전파되는 경우 이것 역시 파열음 카운터 증가에 사용하여 인식을 돕도록 한다. 또 발음 지속 시간이 긴 음운의 인식에서 짧은 거리의 비연속성을 극복하기 위해 오버룩 카운터를 사용한다.

각 음소를 인식하는 방식으로 1차적으로 각 음소의 발음 지속 시간을 이용한다. 자음은 모음의 경우보다 지속 시간이 짧다. 그러므로, 자음중에서 유성 자음은 무성 자음에 비해 많은 카운트를 한다.

카운터 사용 후 인식 결과를 가지고 다시 출력 결과에서의 보정을 위한 2단계의 출력 보정 방법, 음어 인식 휴리스틱을 사용한다.

2차 출력한 결과는 완벽한 이름으로 인식되는 것이 아니다. 그러므로 목표 음소의 배열과 인식된 음소의 배열 사이의 거리를 측정하는 것이 필요하다. 마지막으로 퍼지룰 이용하여 등록된 이름과의 비교

를 통해 가장 높은 값을 가진 이름을 최종적으로 인식한다.

5. 결 론

본 연구는 전화 음성 교환기 구축에 관한 소프트웨어를 목적으로 하여 만들고 있으며, 전화를 컴퓨터와 연동하거나, 교환기와의 연동 문제는 차후 연구 과정으로 남겨 놓았다.

기존의 음성 인식 방식에서는 단순한 원음 인식에만 중요하였으나, 인식된 값만으로 이름을 식별할 수 있도록 퍼지를 이용한다. 음성은 사람마다 개인 차이가 심하고, 그 경우가 너무 많아 경우에 따른 분류작업을 할 수 없기 때문에 퍼지를 사용함으로써 효율성과 속도 면에서 유리하게 한다. 음운 타자기에서 모음과 자음을 차례로 인식하고, 휴리스틱을 적용하는 가운데 퍼지를 이용한다. 예를 들면, 성을 모르고 이름만 안다고 하여도 목록에 있는 사람의 이름 중에서 가장 근접한 사람을 찾아줄 수 있도록 하였다. 또한 음소 단위의 인식을 통하여 인식할 이름의 수가 증가한다고 하여도 별도의 인식 작업이 없이 입력이 가능하도록 한다.

부가적인 이득은, 고가의 장비를 사용하던 음성 인식 분야를 저가의 장비로도 사용 가능하게 할 수 있다. 또한 별도의 DSP 장치 대신에 Sound Card를 사용하여 호환성을 높인다.

참고문헌

[1] J. L. Flanagan, *Speech Analysis, Synthesis, and Perception*, 2nd. ed., Springer-Verlag, New York, 1972

[2] 전용구, 양진우, 김순협, "SOFM 신경회로망을 이용한 한국어 음소 인식", *한국음향학회지*, 제 14권 제 2호, pp. 101-112, 1995년

[3] 김선인, 이행세, "분산 신경망을 이용한 고립단어 음성에 나타난 음소 인식", *한국음향학회지*, 제 14권 제 6호, pp. 54-61, 1995년

[4] A. Waibel et al., "Phoneme Recognition Using Time-Delay Neural Networks", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, No.3 pp. 328-339, Mar. 1989.

[5] A. Ljolje and S. Levinson, "Development of an Acoustic-Phonetic Hidden Markov Model for Continuous Speech Recognition", *IEEE Trans. on*

Signal Processing, vol. 39, No.1 pp. 29-39, Jan. 1991.

[6] X. Huang, "Phoneme Classification Using Semicontinuous Hidden Markov Models", *IEEE Trans. on Signal Processing*, vol. 40, No.5 pp. 1062-1067, May 1992

[7] 정연찬, *한국어 음운론*, 개문사, pp. 127-151, 1980

[8] 박창배, *한국어 구조론 연구*, (주)탑출판사, pp. 11-37, 1990

[9] M. Basseville and Benveniste, "Sequential Detection of Abrupt Changes in Spectral Characteristics of Digital Signals", *IEEE Trans. on Inform Theory*, vol. 29, pp. 709-723, Sept. 1983.

[10] R. Andre-Obrecht, "A New Statistical Approach for the Automatic Segmentation of Continuous Speech Signals", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, No.1, Jan. 1988.

[11] 권영욱, 정현열, "다차원 척도 구성법을 이용한 한국어 음소의 분석", *대한전자공학회논문지*, 제29권 B편, 제11호, pp. 22-30, 1992년 11월.

[12] 감지은, 이문형, "음성인식 증권 정보 서비스 구축 사례" *정보처리학회지* 제6권 제 4호 1999.7

[13] 이아정, "음운 타자기에 근거한 한국어 음어 음성 인식", *홍익대학교 석사학위논문*, 1994.12