

부호패턴에 의한 음성표현과 인식방법

A Speech Representation and Recognition Method using Sign Patterns

김 영 화*, 김 운 일*, 이 희 정*, 박 병 철**

(Young Hwa Kim, Un Il Kim, Hee Jeong Lee, Byung Chul Park)

〈요 약〉

본 논문에서는 새로운 음성표현 방법의 하나로 멜켄스트럼 계수에 대한 부호패턴(+,-)을 이용하는 방법을 제안한다. 부호패턴은 모음이나 비음과 같이 정상성이 강한 음성신호에 대해서 비교적 안정된 형태를 얻을 수 있으며, 음운성을 크게 손상시키지 않고 화자간 개인차를 흡수할 수 있다. 본 논문에서는 이러한 부호패턴을 이용하여 한국어 음소를 표현함으로써 음소모델의 작성과 음소의 인식절차를 크게 감소시킬 수 있음을 보인다.

〈ABSTRACT〉

In this paper the method using a sign pattern(+,-) of Mel-cepstrum coefficients as a new speech representation is proposed. Relatively stable patterns can be obtained for speech signals which has strong stationarity like vowels and nasals, and the phonemic difference according to the individuality of speakers can be absorbed without affecting characteristics of the phoneme. In this paper we show that the reduction of recognition procedure of phonemes and training procedure of phoneme models can be achieved through the representation of Korean phonemes using such a sign pattern.

I. 서 론

연속 발생된 음성신호를 음소단위로 인식하기 위해

서는 각 음소를 명확히 구분지어 표현할 수 있는 효율적인 방법을 모색하지 않으면 안된다. 음소를 표현하는 방법으로는 spectrogram, FFT 스펙트럼, 에너지, ZCR, 그리고 선형예측법에 의해 유도된 파라미터(LPC계열 파라미터)등이 많이 이용되고 있다. 이들 중에서 LPC계열 파라미터들은 음성 스펙트럼을 효율

*성균관대학교 대학원 전자공학과

**성균관대학교 전자공학과 교수

적으로 기술할 수 있을 뿐 아니라 유도과정이 간단하여 음성인식에 매우 유용하게 쓰인다. 그러나 이들 파라미터들은 음운환경이나 화자의 변동에 극히 민감하여 동일음소에 대해서 공통된 표현을 나타내는데 큰 어려움이 따른다.

본 논문에서는 이런 문제를 해결하기 위한 방법으로 음성신호의 단시간 구간에 대한 스펙트럼을 나타내는 멜켵스트럼 계수의 부호만을 이용하여 음소를 표현하는 방법을 제안한다. 이 방법은 적은 파라미터만으로 음소의 표현이 가능하고 모음과 같이 정상성이 강한 음소에 대해서는 안정된 부호패턴을 얻을 수 있으므로 인식절차를 매우 단순화시킬 수 있다. 또한 부호패턴에 의한 음소표현을 음운성을 크게 손상시킴없이 음운환경의 변화와 화자간 개인차를 잘 흡수할 수 있다.

본 논문에서는 한국어의 음소들에 대해 부호패턴을 이용하여 hidden Markov 모델을 구성함으로써 부호패턴이 기존의 LPC 파라미터에 의한 거리척도에 비해 모델 training과 인식절차에서 단순하여 보다 효과적이며 인식률도 크게 떨어지지 않음을 보이고 있다.

II. 부호패턴

1. 부호패턴의 성질

음성 분석을 행하여 얻은 특징 파라미터들은 음운환경과 화자의 변동에 의하여 그 변화의 폭이 다양하다. 이것은 파라미터의 값을 비교하여 음성인식을 행하는 기존 음성인식 방법에서 인식 오류를 발생케 하는 주된 요인이 된다. 따라서 파라미터 값의 변동을 감소시킴으로써 안정된 특징을 나타낼 수 있는 새로운 음성의 표현방법이 요구된다. 부호패턴은 음성신호의 LPC 분석을 통하여 얻어진 파라미터 값중 (+) 또는

(-)의 부호만을 이용하여 음성의 특징을 기술하는 방법으로 음성의 특징을 수치적변동에서 부호적 변동으로 바꾸어 표현함으로써 안정되고 뚜렷한 음성특징을 기술하려는 기법이다.

부호패턴은 기존의 수치적인 계산을 통해 최적값을 찾는 패턴매칭방법과는 달리 눈으로 확인되는 특징을 부호를 이용하여 표시함으로써 처리과정을 단순화하는데서 그 의미를 찾을 수 있다. 여러 복잡한 수치적인 계산으로도 정확한 인식을 하기 어려운데 단순히 2개의 부호만을 이용하여 정확히 음성을 인식하는 데는 많은 어려움이 따른다. 그러나 멜켵스트럼 파라미터의 분포를 보면 차수별로 음소, 음절에 대한 특징의 구분이 가능하다. 그림1은 단어 /노가리/에 대한 멜켵스트럼 계수의 크기를 차수별로 나타낸 것이다. 점선은 각 음소를 구분하고 있는데, 여기서 보면 모음의 안정 구간에서 모든 계수의 크기가 차수별로 양 또는 음으로 일괄적으로 나타나고 있다. 즉 5차 계수까지 음소 /오/는 (+ + - - -), /아/는 (+ - - - -), 그리고 /이/는(+ + + - -)로 일정한 형태를 나타낸다. 이와 달리 자음구간에서는 계수의 크기가 양, 음의 변화를 보여 일정한 부호형태를 나타내기 힘들다. 본 논문에서는 이와 같이 수치적인 계산에 의하지 않고 부호패턴 그 자체만의 확인으로 구분이 가능한 특징을 음성 처리에 이용하고자 한다. 표1에 /노가리/에 대한 모든 계수를 부호로 바꾸어 차수별로 시간에 따른 변화를 보이고 있다. 일정한 모음의 부호패턴이 한눈에 들어오는 것을 볼 수 있다. 이 자,모음에 대한 부호형태의 특징을 채우면 많은 효과적인 음성처리를 할 수 있을 것이다. 이와 같은 부호패턴의 눈으로 나타나는 특징이 없다면 두 값만으로 음성을 처리하기는 상당한 어려움이 따를 것이다. 음성을 부호패턴

표 1. 멜켵스트럼 계수의 부호패턴(/노가리/)
Sign pattern of Mel cepstrum coefficients (/ NXOH GXAA RXIH /)
time(frame)

	123456	78901234567	890	12345678901	234567	8901234567	890123456
order	1	-----+	+++++++	+++	+++++++	+++++	+++++++
	2	+++++	+++++++	+++	-----	+++++	+++++++
	3	++----	-----	---	-----	++----	+++++++
	4	+++++	-----	---	-----	+++++	+++++++
	5	-----	-----	+	-----	-----	-----
	o	o	g	a	l	i	

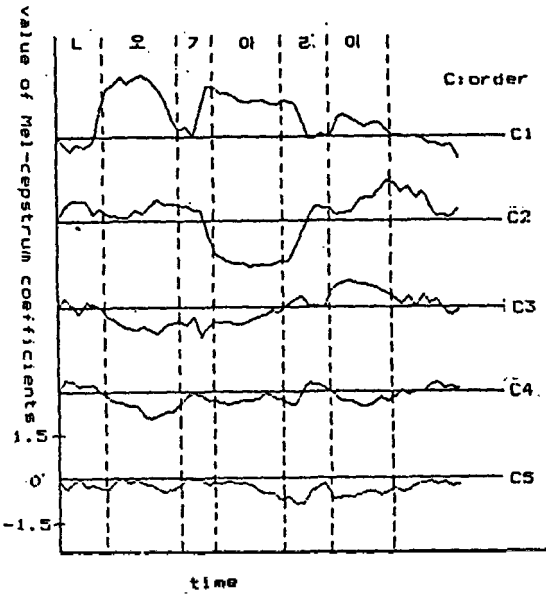


그림 1. 멜켄스트럼 계수값의 시간적 변화(단어 /노가리/) Variation of values of Mel-cepstrum coefficients in time (Word speech /NXOH GXAA RXIH).

표 2. 음성 분석 조건 Conditions of speech analysis.

샘플링 주파수	: 10kHz
A/D 변환	: 8bit
LPE	: 4kHz
중첩	: 10ms
프레임 길이	: 20ms
분석 차수	: 15차
window	: Hamming
Preemphasis	: $1-z^{-1}$

C5:(+), C6:(-)로 세 명의 화자에 대해 파라미터의 차수별 분포가 모두 일정하게 나타남을 확인할 수 있다. 많은 음성 분석 파라미터를 이용한 알고리즘들이 화자적응에 상당한 어려움을 가지고 있는 것을 고려할 때 부호패턴의 화자 적응 가능성은 큰 의미가 있는 것이다.

본 논문에서 이용한 음성의 분석조건은 표2와 같다. 멜켄스트럼 계수는 음성의 LPC분석을 거쳐서 얻은 켈스트럼 계수를 이용하여 구한다. 이산적음성신호에서 유한한 길이를 갖는 프레임 구간에서 구한

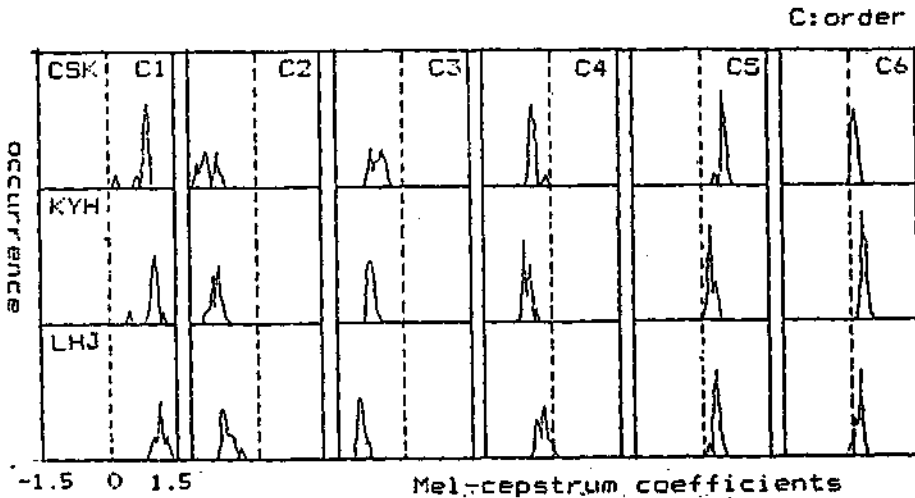


그림 2. 모음 /아/의 멜켄스트럼 계수의 분포 Distribution of Mel-cepstrum coefficients for vowel /AA/.

으로 표현하면 각 음소에 대한 특징을 변화시키지 않고 처리과정을 단순화시킬 수 있다.

또 모음의 부호패턴은 화자의 변동에 따르는 영향을 효과적으로 흡수하고 있다. 그림2에 3명의 남성 화자가 발성한단모음/아/에 대한 멜켄스트럼 파라미터의 차수별 분포도를 나타내고 있다. 그림2에서 모음 /아/의 부호는 C1:(+), C2:(-), C3:(-), C4:(-),

m차의 켈스트럼을 이용하여 Oppenheim Johnson의 점화식 변화에 의해 멜켄스트럼 계수를 얻는다.⁽⁴⁾

2. 부호패턴의 음운 표현

2.1. 모음의 부호패턴

모음은 음운적 변화와 화자 등에 관계없이 일정한 부호패턴을 가지고 있어서 다른 음운 구간과는 확실한

구분을 갖는다. 부호패턴의 이러한 특성은 각 모음들을 쉽게 인식할 수 있고, 또 비슷한 형태를 갖는 모음끼리 몇개의 그룹으로 묶어서 유효하게 이용할 수 있음을 암시하는 것이다.

단모음에 있어서 단독모음(V)의 경우 저차에서 매우 안정된 부호패턴을 보이고 초성+모음(CV), 모음+종성(VC)의 경우는 조음결합의 영향으로 한모음에 대해서 몇가지의 다른 부호패턴을 나타낸다. 표3은 모음의 부호패턴과 조음요소를 보인 것이다. 이 표에서 X는 부호를 명확히 규정짓기 어려운 경우를 나타내고 있다. 여기서 부호패턴A는 어떠한 조음결합의 영향에 대하여도 변하지 않는 부호형태를 표시한 것이고, 부호패턴B는 표시한 부호형태가 나타날 확률이 70% 이상인 경우를 표시한 것이다.

표 3. 조음요소에 따른 모음의 부호패턴
Sign pattern of vowel from viewpoint of articulation factors.

Table with 4 columns: 조음요소 (phonological factors), 모 음 (vowel), 부호패턴A (Sign pattern A), 부호패턴B (Sign pattern B). Rows include vowels like 비, 이, 에, 에, 아, 우, 오 and their articulation factors like 전설고, 편반, 전, 중, 반, 개, 원, 후.

순: 입술모양·원: 원순모음, 비: 비원순모음
설: 혀의 위치·전: 전설모음, 중: 중설모음, 후: 후설모음
고: 설고·개: 개모음, 반: 반개(또는 반폐) 모음, 폐: 폐모음

모음의 발성은 입술모양, 혀의위치, 혀의높이의 세가지 조음요소에 의해 이루어진다. 모음들간에 조음요소가 많이 겹칠수록 부호패턴의 유사도가 그에 따라 더욱 커지는데, 그 예로 부호패턴의 유사도가 큰 두쌍인 /아/와 /어/, 또 /오/와 /우/의 경우 각각 ‘중설 개모음’과 ‘후설 폐 원순모음’의 조음요

소가 겹치고 있다. 단모음의 경우 /애/와 /에/, /아/와 /어/, 그리고 /오/와 /우/는 부호패턴이 유사하게 나타난다. 이 세그룹은 서로의 모음들이 각각 전설, 중설, 후설의 특징을 갖고 있는데, 이런 부호패턴의 유사성으로 볼 때 모음은 혀의높이와 입술모양보다는 혀의위치에 의해 영향을 많이 받는 것을 알 수 있다. 이것은 단모음에서는 부호패턴만으로는 신뢰성있는 음소인식을 행하기는 어렵지만 위와 같이 조음위치에 따라 서로 유사하게 나타나는 음소들을 묶어서 한 음소군으로 한다면 다른 계산이나 처리과정 없이 부호패턴만을 조사함으로써 그 음소군들을 나눌 수 있다. 이 음소군은 음소인식이나 어휘처리를 하는데 있어 그 과정을 단순화시키면서 효과적으로 행할 수 있을 것이다.

2.2. 자음의 부호패턴

자음은 그 자신만의 단독 발성이 불가능하여 반드시 모음이 뒤따라야 하므로 후속모음의 영향을 강하게 받게 된다. 이 때문에 자음은 음운환경에 따라 음의 변형이 많이 일어나게 된다. 따라서 자음에서는 공명음인 비음과 유음을 제외하면 모음과 같이 일정한 부호패턴을 찾기 힘들다. 특히 종성자음의 경우는 그 변화가 더 심하다. 그러므로 모음과 같이 부호의 대조만으로 음소를 판별하기는 힘들다. 표4는 후속모음의 변화에 의하여 변하는 유성자음의 부호패턴을 나타내고 있다. 이 표에서 동일 자음이라 할지라도 뒤에 오는 모음이 달라지면 전혀 다른 부호패턴을 갖게 되는 것을 확인할 수 있다. 그러나 /ㄱ/, /ㄴ/ 과같이 비음의 경우는 모음과 유사한 특성에 의해 후속모음의 변화에 대해서도 비교적 일정한 부호패턴을 나타내는데 이것이 자음들을 분류할 수 있는 한 특징이 될 수 있다.

표 4. 자음에서 후속모음 효과
Effect of succeeding vowel in preceding consonant.

Table with 7 columns: 선행자음 (preceding consonant), BX, DX, ZX, GX, MX, NX. Rows show the effect of succeeding vowels H, A, O on the preceding consonants.

Ⅲ. 부호패턴과 확률 모델에 의한 음성신호 표현

제 II 장에서 음성신호의 표현 방법으로 부호패턴을 제안하였다. 음성신호는 모음과 비음, 설측음같은 공명음에서는 비교적 안정된 부호패턴을 갖지만 자음에서는 일정한 부호패턴을 설정하기가 힘들다. 설사 설정할 수 있다하여도 그 패턴을 일일이 일거하기에는 많은 어려움이 따른다. 그러나 모음과 조음결합되는 과정에서 그 음들의 특징적인 면들을 발견할 수 있으므로 확률적인 수법을 이용하면 더 효과적으로 음성을 표현할 수 있다. 이 장에서는 Markov source에 의해 부호가 발생하는 것을 가정하여 음소를 확률적으로 모델화하는 방법을 설명한다. HMM 방식은 많은 양의 트레이닝 데이터로 부터 Markov chain을 구성하는 상태 전이 확률, 상태 심볼 확률을 추정하여 모델링을 행하여 음성 특징을 나타내는 방법이다. 이 모델은 음성신호의 시간적인 변화를 잘 나타낼 수 있고, 비교적 적은 기억용량으로 원하는 음성신호의 각 패턴들을 기억시킬 수 있다. (N2X3)

한 프레임의 음성신호를 나타내는 파워 스펙트럼을 부호패턴으로 표시할 때 이 부호패턴을 이용하여 Markov 모델을 만들기 위해서 적절한 규칙을 적용하여 부호패턴에 대응하는 관측심볼로 변화시킬 필요가 있다. 부호패턴을 s로 볼때, 이것에 일대일 대응하는 실수값이 반드시 존재하도록 규칙X를 설정한다. 이 규칙X에 의해 부호패턴 s는 이에 유일하게 대응하는 하나의 실수값 X(s)를 갖게 되는데 본 논문에서 사용한 규칙은 식(1)과 같다. p는 이용하는 파라미터의 최고차수를 나타낸다(6차까지 이용하는 경우 p는 6)

$$X(s) = \sum_{i=1}^p W(i) * sgn(i) \quad (1)$$

여기서

표 6. 음소 분류
Phoneme classification.

단 모음(7)	아(AA), 어(AX), 우(UH), 오(OH), 에(EH), 애(AE), 이(IH)
이중모음(5)	야(J-AA), 여(J-AX), 유(J-UH), 요(J-OH), 예(J-EH)
초성자음(9)	ㄱ(GX,KX), ㄴ(NX), ㄷ(DX, TX), ㄹ(LX, RX), ㅁ(MX), ㅂ(BX, PX), ㅅ(SX), ㅈ(ZX, CX), ㅎ(HH)
종성자음(7)	ㄱ(KQ), ㄴ(NX), ㄷ(TQ), ㄹ(LX), ㅁ(MX), ㅂ(PQ), ㅇ(NG)

$$W(i) = 2^{-(i-1)} \quad (2)$$

$$sgn(i) = 1 \text{ (파라미터의 } i\text{차 계수가 양수일 때)} \quad (3)$$

$$sgn(i) = 0 \text{ (파라미터의 } i\text{차 계수가 음수일 때)} \quad (4)$$

이다.

이 규칙을 이용해서 확률 음소모델을 만들때 X(s) 자체가 하나의 심볼을 나타내기 때문에 제한된 관측심볼을 만들어내기 위한 벡터양자화가 필요하다. 음소모델은 X(s)를 이용해서 구한 파라미터 (π, A, B), 즉 상태 초기확률, 상태 전이확률, 상태 심볼확률로 나타낼 수 있는데, 하나의 관측심볼에 대하여 많은 트레이닝 과정을 통해서 얻어진다. 이와같이 얻은 음소모델들에 의해 입력음소의 확률값을 구하여 그 값이 가장 큰 모델을 찾음으로써 각 음소에 대한 인식을 행할 수 있다.

Ⅳ. 실험 및 고찰

1. 실험고찰

이 실험에서는 15차로 분석한 파라미터 중에서 멜컨스트럼의 저주파 특성을 고려하여 6차까지만을 이용하였다. 실험에 이용된 데이터는 한명의 남성화자가 발성한 890단어(178개의 단어를 5회 발음)에서 음소를 선택하였다. 자, 모음 실험에 사용된 음소의 갯수를 표5에 나타냈다. 이와함께 실험에 이용된 음소의 분류를 표6에 표시하였다.

표 5. 자,모음 실험 데이터
Experiment data of consonants and vowels.

음소분류	사용단어	사용음소	비고
단 모음	145	294	145개 단어 1회씩 발음
이중모음	145	145	29개 단어 5회씩 발음
초성자음	890	1400	178개 단어 5회씩 발음
종성자음	310	310	62개 단어 5회씩 발음

멜켵스트림 계수는 인간의 청각 특성을 잘 나타내는 파라미터로 특히 저주파수 분해 능력이 좋게 나타나므로 일반 켈스트림보다 적은 차수의 계수로도 청각적으로 같은 질의 합성음을 얻을 수 있어 음성의 정보 압축에도 유효하다.⁽⁶⁾ 따라서 본 논문에서는 멜켵스트림 계수를 부호패턴의 특징관찰 및 인식실험을 위해 사용하였다. 입력 음성 신호에 대한 부호패턴을 추출하고 난 후, 각 음소별로 세그멘테이션을 하고, 그림 3의 HMM을 트레이닝하여 음소모델을 작성한다. 실험에 사용된 음소모델은 그림3과 같이 5상태로 제한된 Markov 모델을 사용했다. 이것은 left-to-right 모델로 시작은 항상 상태 q_1 에서 이루어진다. 즉 $\pi=(1,0,0,\dots,0)$ 으로 된다. 상태 심볼확률 B의 초기값은 모든 관측 심볼에 관하여 동일한 확률을 갖는 것으로 하였다. 만약 행렬 $B=[b_{jk}]$ 의 원소에 대하여 $b_{jk} < e$ 이면 $b_{jk} = e$ 로 하였다. 실험에서 사용된 e의 값은 10^{-6} 이다. 모델의 트레이닝은 확률이 1%이상 개선되지 않거나 바로 전 확률보다 낮아질 때 중단하였다.

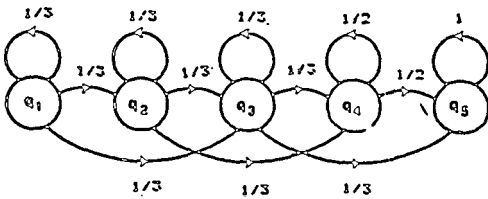


그림 3. HMM 음소 모델
HMM phoneme model.

2. 인식실험 결과 및 고찰

2.1. 단모음

단모음에 대해 II장에서 설명한 것처럼 조음위치(혀의위치)에 따라 부호패턴이 비슷하게 나타나는 음소끼리 묶은 음소군에 대한 인식실험결과를 표7에 나타냈다. 이 표에서 보면 평균 인식률이 89.4%로 단모음에서는 순수하게 부호패턴만으로도 특징있는 음소들을 무리지을 수 있다. 즉 단모음의 부호패턴은 그것만으로 신뢰성있는 음소인식을 행하기는 다소 어렵지만 유사한 패턴을 갖는 음소끼리 묶을 수 있기 때문에 음소군내에서 새로운 음향적 지식을 추가한다면 시간을 단축시키면서 신뢰할만한 인식 효과를 누릴

수 있을 것이다.

표 7. 모음의 음소군에 대한 인식결과
Recognition result for the group of vowel phonemes.

O/I	아	어	우	오	애	에	이
아,어	48	36	4	14			
오,우		1	42	35			1
애,에	2	1	4	1	17	36	2
이					2	1	47
인식률(%)	96.0	94.7	84.0	70.0	89.5	97.3	94.0

단모음에 대한 화자간의 개인차 흡수를 알아보기 위해 38개의 단어에서 7개의 단모음에 대해 각각 10개씩의 음소를 선택하였다. 음소모델을 작성한 화자를 포함하여 4명의 서로다른 남성화자의 음성을 가지고 비교하였다. 인식된 음소의 갯수와 인식률을 표8에 나타냈다. 음소모델은 화자A의 음성을 가지고 작성하였다. 인식결과에서 나타나듯이 특정화자A와 3명의 불특정화자간에 인식률의 차이가 작음을 확인할 수 있다. 비록 각음소가 아닌 음소군에 대한 인식결과이지만 화자적응의 실현 가능성을 보여주는 것이다.

표 8. 화자간의 인식결과와 비교
Comparison of recognition result of speakers.

화자/모음	아,어(20)	오,우(20)	애,에(20)	이(10)	인식률(%)
A	19	16	15	9	84.3
B	18	20	10	10	82.9
C	18	19	14	10	87.1
D	20	19	14	10	90.0
인식률(%)	93.8	92.5	66.3	97.5	86.1

2.2. 이중모음

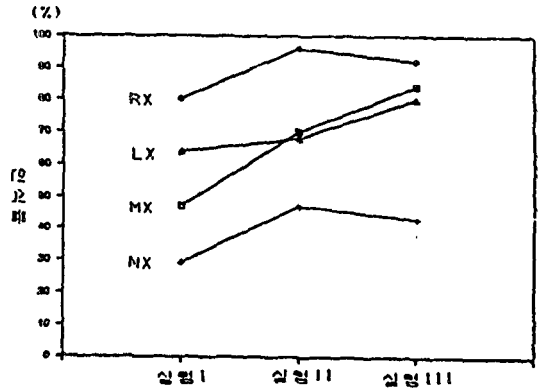
음소모델을 이용한 이중모음의 인식률을 표9에 나타냈다. 이중모음은 단모음에 비해서 대체로 높은 인식률을 나타내는데, 이것은 모음의 구간내에서 두 음의 천이과정을 통해 조음되는 이중모음의 특성때문에 좀더 특징있는 음소모델을 만들기 때문이다.

표 9. 이중모음의 인식결과
Recognition result of diphthongs.

I/O	J-AA	J-AX	J-EH	J-UH	J-OH	인식률(%)
J-AA	29	1				96.7
J-AX	3	27				90.0
J-EH			25	5		83.3
J-UH	1	1		19	4	76.0
J-OH		1	3	2	24	80.0
평균						85.5

2.3 초성자음

초성자음의 인식실험은 /ㅃ, ㄷ, ㄱ, ㅅ, ㅎ, ㅈ, ㄴ, ㄹ, ㅁ, /의 9개의 음소에 대해서 실시하였는데 /ㅃ, ㄷ, ㄱ, ㅈ/에 대해서는 유성음과 무성음을 구분하여 모델을 작성하고 /ㄹ/에 대해서는 실측음과 탄설음을 구분하여 모델을 작성하여 모두 14개의 음소모델을 만들었다. 이때 음소모델은 각각 100개씩의 음소를 이용하여 만들었다. 실험은 음운성질을 고려하여 세단계의 과정으로 나누어 행하였다. 실험 I에서는 모델을 만든 14개의 음소 모두에 대해서, 실험 II에서는 음소를 4개의 공명음과 10개의 비공명음의 모델을 분리해서, 그리고 실험 III에서는 길이의 정규화를 행한 후 실험 II의 과정을 되풀이 하였다. 즉 음소를 시간길이에 따라 4프레임 미만과 4프레임 이상으로 음소를 구분하여 음소모델을 만들고, 다시 공명음, 비공명음으로 분리해서 실험을 행하였다. 위의 세단계에 걸쳐



(c)

그림 4. 자음의 인식결과 (a) 파열음 (b) 마찰음/파찰음 (c) 공명음
Recognition result of consonants.
(a) Plosives (b) Fricatives/Affricates (c) Sonorants

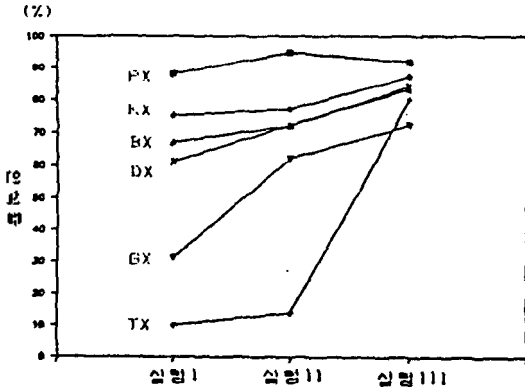
행한 인식결과를 그림 4에 파열음, 마찰음(파찰음), 공명음으로 나누어서 비교하였다. 길이의 정규화를 행한 실험 III에서의 인식률은 공명음의 경우 평균 74.7%, 비공명음의 경우 76.7%로 음소들이 길이에 따른 영향이 크다는 것을 확인할 수 있다. 간단화된 인식과정을 행하면서도 이런 인식률을 얻을 수 있는 것은 큰 의의가 있다고 생각된다.

2.4 종성자음

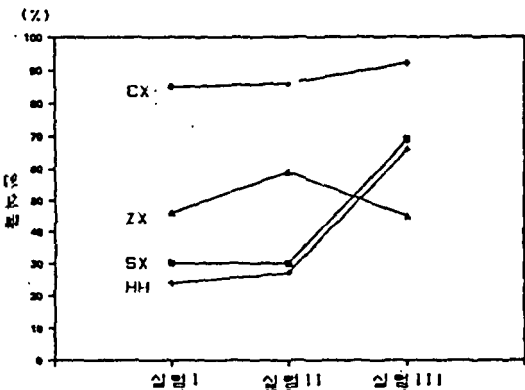
종성자음 /ㅃ, ㄷ, ㄱ, ㅁ, ㄴ, ㅇ, ㄹ/에 대한 인식률을 표 10에 나타냈다. 모든 음소모델은 각각 50개씩의 음소를 사용하여 만들었다. 자동 새그멘테이션이 어려운 파열되지 않는 내파음 /ㅃ, ㄷ, ㄱ/은 인식률이 매우 낮게 나타나고 있다. 나머지 자음들은 초성자음과 비슷한 인식률을 나타내고 있다.

표 10. 종성자음의 인식결과
Recognition result of final consonants.

자 음	PQ	TQ	KQ	MX	NX	NG	LX
인식률(%)	70	20	46	66	76	78	92



(a)



(b)

V. 결 론

음성신호를 표현하는데 있어 real time을 실현하고

높은 인식효과를 얻기 위해서 계산량의 증가를 막아 시간의 소비를 줄일수 있고, 또 모든 화자에 공통된 음운특징을 갖는 파라미터의 발견이 필요하다. 본 논문에서는 이러한 문제점을 해결할 수 있는 음성 표현방법으로 부호패턴을 제안하였다. 그리고 부호패턴을 이용하여 hidden Markov 모델을 작성함으로써 한국어 음소를 표현하면 음소의 인식절차를 크게 단순화시킬 수 있음도 보였다. 부호패턴은 모음과 비음, 유음같은 공명자음에서는 안정된 부호형태를 나타내지만 비공명자음에서는 후속모음의 영향으로 부호의 변화가 심하게 나타났다. 단모음의 경우는 다른방법의 도입없이 부호패턴 그 자체로 유사 음소군의 인식률이 89.4%를 나타냈다. 자음의 경우는 음운성질에 따라 공명음과 비공명음의 구분, 길이의 정규화를 통해서 각 음소의 인식 결과를 비교하였다. 길이정규화를 해서 공명음과 비공명음으로 구분한 경우, 각 음소의 인식률이 약 76%로 거리 척도를 HMM에 이용해서 한 인식결과와 비슷하게 나타나고 있다.¹⁷⁾ 또 개인성을 내포하는 모음에 있어서 여러화자의 부호패턴이 유사하게 나타나 화자적응이 가능함을 보였다.

부호패턴을 음성 표현방법으로 이용함에 있어서 부호 자체의 조사만으로 분류할 수 있는 음소군내에서 개개의 음소를 구분할 수 있는 음향적 특징을 찾는다 면 상당히 높은 인식률을 얻을 수 있을 것이다. 또 이 음소군을 단어 가설에 이용해도 효과적이라고 생각 된다. 그리고 부호패턴은 대체로 모음과 자음과의 사이는 정확히 구분하므로, 이것을 자동 세그멘테이션 에도 적용해 볼 수 있을 것이다. 이에 앞서서 하나의 부호패턴에 여러 개의 모음이 대응되어 음소인식 효율을 적게하는 것을 막기 위하여 부호패턴과 음성 스펙트럼 사이의 관계를 보다 명확히 규명하여 멜켵스트럼 계수값의 기준을 일률적으로 고정시키지 않고 차수마

다 각 음소의 구분에 미치는 영향을 고려하여 기준점을 선택하는 방안도 검토해 볼 필요가 있으리라 생각 된다.

참 고 문 헌

1. S.E.Levinson, L.R.Rabiner, M.M.Sondhi, "An Introduction to the Application of the Theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition," The Bell System Technical Journal, Vol. 62, No.4, pp.1035-1074, April 1983.
2. L.R.Rabiner, B.H.Juang, "An Introduction to Hidden Markov Models," IEEE ASSP Magazine, pp.4-16, January 1986.
3. L.R.Rabiner, S.E.Levinson, M.M.Sondhi, "On the Application of Vector Quantization and Hidden Markov Models to Speaker Independent Isolated Word Recognition," The Bell System Technical Journal, Vol.62, No. 4, pp.1075-1105, April 1983.
4. Masanori AKITA, Satoshi IMAI, "A Segmentation Method for Continuous Speech Using Mel-Cepstrum," 일본 전자 통신 학회 논문지, Vol.J68-D, No.3, pp.323-329, March 1985.
5. Tadashi KITAMUKA, Satoshi IMAI, Chieko FURUICHI, Takao KOBAYASHI, "Speech Compression Using Mel-Cepstrum," 일본 전자 통신 학회 논문지, Vol.J67-A, No.11, pp.1092-1093, November 1984.
6. 김운일, "부호패턴을 이용한 모음인식," 성균관 대학교 전자공학과 석사논문, 1988.
7. 유현창, 이회정, 박명철, "음소 길이를 고려한 3-State Hidden Markov Model에 의한 한국어 음소 인식," 한국 음향 학회지, Vol.8, No. 1, 1989.

▲김영화



1962년 11월 7일생
 1988년 2월 성균관대학교전
 자공학과 졸업
 1988년 3월~현재: 성균관대
 학교 대학원 전자공학
 과 석사과정

▲이희정 8권1호 참조

▲박병철 8권1호 참조

▲김운일



1963년 3월 18일생
 1987년 2월 성균관대학교전
 자공학과 졸업
 1989년 2월 동대학교대학원
 석사과정 졸업
 1989년 8월~현재: 군복무