

반음절 단위 HMM을 이용한 연속 숫자 음성인식

Continuous Digits Speech Recognition using Semisyllable Unit HMM

윤 재 선*, 홍 광 석*
(Jeh Seon Youn*, Kwang Seok Hong*)

*이 논문은 성균관대학교의 1997년도 석급학술연구비에 의하여 연구되었음.

요 약

본 논문에서는 조음 효과에 대처할 수 있는 새로운 음성인식 단위로 반음절, 반음절 + 반음절 단위 HMM을 제안하여 연속 숫자 음성인식을 하였다. 반음절 단위는 무음과 안정구간으로, 반음절 + 반음절 단위는 안정, 천이, 안정구간으로 구성 되어 있고, 음성인식 단위 분할시 비교적 스펙트럼의 변화가 안정한 모음구간에서 분할하므로 분할 위치가 약간 변하여도 인식성능에는 큰 영향을 주지 않게 된다. 또한, 제안된 반음절, 반음절 + 반음절 인식단위는 그 패턴 안에 다음 숫자열의 정보를 포함하고 있기 때문에 모든 HMM 패턴들과 비교하는 것이 아니라, 다음 숫자열의 정보를 포함한 HMM 패턴들과 비교한다. 인식실험결과 제안된 방법이 효율적임을 확인하였다.

ABSTRACT

In this paper, we present new speech recognition units, semisyllable and semisyllable + semisyllable unit HMM that can cope effectively with the coarticulation effects and recognize the continuous digits speech. The semisyllable unit consists of silent and stationary parts. The semisyllable + semisyllable unit has stationary part at each end of the unit and transient part of the middle. Even though the segmented position may happen to shift, there is no effect on recognition performance because segmentation happens to be a stable vowel section which has little variation of spectrum. Since the proposed semisyllable and semisyllable + semisyllable units are involved the information of subsequent sequences, segmented units solely compare with the semisyllable and semisyllable + semisyllable HMM patterns including the information of the subsequent sequence in recognition process. The recognition results confirm the effectiveness of the proposed method.

1. 서 론

컴퓨터의 발전과 통신을 이용한 정보 및 금융 서비스의 확대됨에 따라서 주민등록번호, 비밀번호, 통장번호, 회원번호 등 많은 분야에서 무제한 연속 숫자열에 대한 인식을 필요로 하고 있다. 이들 연속 숫자열에 대한 인식은 키보드 입력뿐만 아니라 음성인식 등의 수단에 의해서 확인할 필요성이 증가하고 있다.[1][2] 연속 숫자 음성인식은 숫자음의 경계가 불명확한 곳이 많고, 조음 현상으로 인해, 같은 음소라도 다르게 발음되는 경우가 있다. 따라서 오인식율도 고립 숫자음의 경우에 비해서 훨씬 높게 된다. 그러나, 이런 문제에도 불구하고 연속 숫자 음성인식의 중요성은 명백하다. 그 이유는 연속 음성인식을 통해서만 인간과 기계사이의 통신수단의 자연스러움과 원하는 속도를 얻을 수 있기 때문이다. 기존의 숫자음

인식은 주로 음절 또는 음소 등의 부단어 단위를 이용하여 인식을 하여 왔지만[3][4] 연속 발성된 경우 음성의 특성상 음절, 음소 단위로 정확한 분할을 하기가 매우 어려우며, 음성인식에서 오인식률의 절반정도가 부정확한 경계 분할이 원인이다.[5]

본 논문에서는 반음절, 반음절 + 반음절을 인식단위로 하는 새로운 연속 숫자 음성인식 시스템을 제안하고 그 성능을 평가하고자 한다. 인식 단위로서는 부정확한 분할문제를 해결하기 위하여 발생된 음절 중에 음향특성이 안정된 모음 부분에서 분할 처리하는 방법을 제시한다. 이 방법을 연속 숫자 음성인식에 적용하기 위하여 단독 숫자음 0/공/에서 9/구/까지의 앞부분과 뒷부분의 반음절 각각 10개와, 연속 발성음 00/공공/, 01/공일/, ..., 99/구구/의 연결부분의 반음절 + 반음절 100개를 기본 인식단위로 제안하고, 인식 알고리즘은 인식 성능이 뛰어난 HMM[6]을 이용하였다. 연속 숫자음 인식 방법은 반음절 패턴을 분할하여 인식한 후, 인식된 패턴의 정보로부터 다음 반

* 성균관대학교 전기 전자 및 컴퓨터 공학부
접수일자 : 1998년 6월 8일

음절+반음절 패턴을 추적하는 방법을 제안하여 그 성능을 확인하였다.

II. 반음절, 반음절+반음절 패턴과 HMM

연속 음성인식에서 단어 단위로 인식을 할 경우, 그 표준패턴을 위한 기억용량 및 계산 시간을 상당히 많이 필요로 한다. 그러나, 음소나 반음절 등의 부단어 단위로 경계를 구분할 수 있다면 대용량의 고성능 인식 시스템의 구성이 가능하다.



그림 1. 숫자음 /육육/의 파형과 스펙트로그램
Fig. 1 Waveform and spectrogram of digit /yuk-yuk/.

한국어 연속 숫자 음성의 경우 0에서 9까지 10개의 단음절의 조합으로 구성되어 있다. 그러나, 연속된 숫자음의 발생시에는 숫자와 숫자사이에 음절의 구분이 없이 연결되어 나타나는 경우를 음성 파형 관찰에서 쉽게 확인할 수 있다. 일례로써 그림 1에 숫자음 66/육육/이라고 발생한 음성 파형과 스펙트로그램을 나타내었다. 그림 1과 같이 자연스러운 발생으로 인하여 육과 육 사이가 연결되어 있고, 이와 같은 경우 정확한 음절 구분이 쉽지 않게 된다.

2.1 반음절, 반음절+반음절 숫자 패턴 분할

0에서 9까지의 단독 발생음에서 시작 부분과 끝 부분 각각 10개씩 반음절 20개의 분할은 음성 신호의 프레임당 에너지[2]와 Delta 에너지, 에너지로 정규화된 영교차율을 기준으로 안정된 모음구간의 평균을 분할하였다. 그림 2는 단독 숫자 발생음 질에 대한 반음절, 반음절 분할의 예로써 실선은 에너지, 점선은 영교차율, 대시는 Delta 에너지값을 나타낸 것이며, 이 파라미터들로부터 모음의 안정구간을 찾아낸 후, 안정구간의 중간 위치에서 반음절 /치/, 반음절 /일/을 분할하였다.

반음절+반음절 단위의 분할은 Delta 에너지를 이용하여 숫자음의 시작 부분을 검출하고, 다음에 에너지와 Error의 문턱값 이상이 되는 공통 프레임을 저장한다[7]. 앞음절과 뒤쪽 음절 각각의 모음구간 평균 프레임을 상하고, 그 구간을 반음절+반음절 단위로 정의했다. 그림 3은 두자리 숫자음 /공육/을 분할하는 방법을 나타낸 것으로 실선은 에너지, 짧은 실선은 Error값이며, 공과 육에서 음절의 안정된 모음 구간을 구하여 반음절+반음절 단위

/용유/를 분할한 것을 나타내었다.

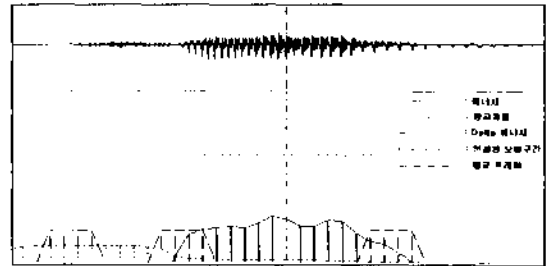


그림 2. 단독숫자음 질의 반음절 /치/와 /일/
Fig. 2 Semisyllable /fj/, fil/ of digit /gil/.

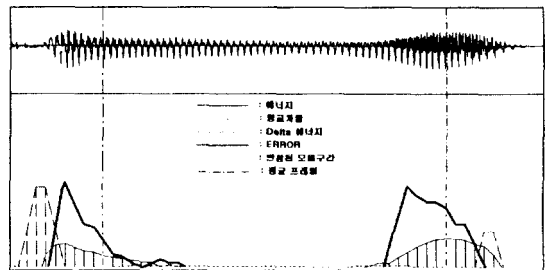


그림 3. 숫자음 /공육/의 반음절+반음절 분할
Fig. 3 Semisyllable + semisyllable segmentation of digit /g5-yuk/.

일반적으로 HMM에 많이 사용되는 triphone은 대부분 천이-안정-천이의 세 가지 상태로 구성되는 것에 반하여 [8], 본 논문에서 정의된 단위는 음성이 시작하는 반음절은 무음-안정의 상태를 갖고, 반음절+반음절인 경우 안정-천이-안정의 상태와 음성이 끝나는 반음절은 안정-무음의 상태로 구성된 세 종류의 인식단위로 구성한다. 이 방법은 인식단위 분할시 천이구간보다는 스펙트럼의 변화가 적은 안정된 정상 상태에서 분할하면, 분할 구간을 쉽게 찾을 수 있고 약간의 분할 위치 차이가 있어도 인식성능에는 큰 영향을 주지 않게 된다. 또한, 분할된 인식 단위는 음성의 천이부분을 잘 표현할 수 있다.

2.2 반음절, 반음절+반음절 HMM

HMM을 이용한 음성 인식 시스템에서 인식 대상 어휘들에 대한 HMM들이 블랙박스과 같은 역할을 하여 미지의 데이터를 각 모델에 적용했을 때 가장 큰 값의 확률을 나타내는 것을 인식된 것으로 한다. HMM의 성능에 영향을 미치는 것은 상태전이(state transition)모델, 상태수, VQ의 코드워드수이다. 본 논문에서 상태 천이는 상태 자체에 시간적 순서를 부여할 수 있는 left-to-right 모델을 사용하였고, 0부터 9까지의 단독 발생된 숫자음에 대한 반음절 패턴은 앞쪽 패턴의 무음-안정, 뒤쪽 패턴의 안정-무음의 특성을 모델링하기 위해 상태수를 2개씩 두었으며, 00에서 99까지 발생된 100개의 반음절+반음절 패턴의 상태수는 안정-천이-안정 상태를 나타내기 위해 상태수를 5개씩 두었다. 분할된 인식단위를 VQ에 의해 코드

값으로 이산화한 후 반음절, 반음절+반음절 단위에 따라 상태수가 2개, 5개로 하는 HMM 모델의 예를 그림 4와 그림 5에 나타내었다.

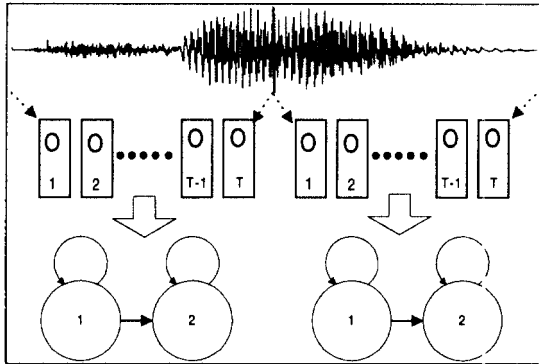


그림 4. 단독 숫자음 /칠/에서 상태수가 2개인 반음절 HMM /치/와 /일/
 Fig. 4 Semisyllable HMM /ʃi/, /il/ with states 2 in digit /ʃil/.

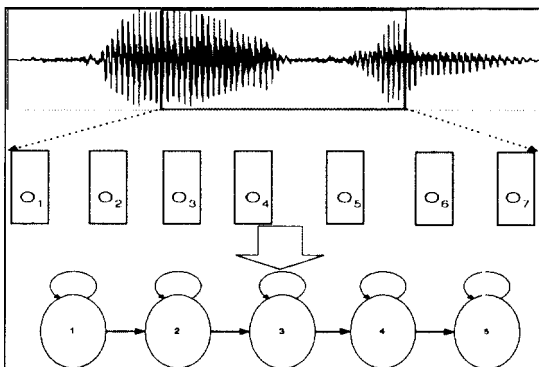


그림 5. 두자리 숫자음 /사삼/에서 상태수가 5개인 반음절+반음절 HMM /아사/
 Fig. 5 Semisyllable+semisyllable HMM /a-sa/ with states 5 in two digits /sa-sam/.

III. 연속 숫자 음성인식

3.1 반음절 Model 분할 인식

본 논문에서 제안한 연속 입력된 숫자음의 인식방법은 입력음성의 모음 구간 분할을 통해 반음절, 반음절+반음절, ..., 반음절+반음절, 반음절 단위로 구분한 후 각각의 분할된 반음절 단위들은 반음절, 반음절+반음절 HMM과 비교하며 인식하는 것이다. 이 방법은 그림 6에 나타내었다. 그림에서 입력 숫자음의 시작 부분의 반음절과 끝나는 부분의 반음절은 각각의 반음절 HMM에 의하고, 분할된 반음절+반음절 단위들은 반음절+반음절 HMM에 의해 인식하는 과정을 나타내었다.

3.2 규칙을 이용한 반음절 패턴 인식

여러 자리수의 숫자음인 경우 규칙을 이용하여 최적의 반음절 열을 찾아내어 연속 숫자음을 인식한다. 일반적인

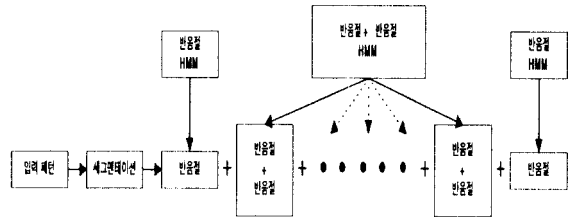


그림 6. 반음절, 반음절+반음절 단위의 인식 과정
 Fig. 6 Recognition process of semisyllable and semisyllable+semisyllable unit.

로 연속 숫자음을 인식하고자 할 경우, 모든 인식 단위들과 비교를 통해 시퀀스를 비교하지만[1], 제한한 인식 단위를 이용하여 인식하는 경우 모든 반음절을 인식후보로 두고 인식을 하는 것이 아니라, 첫 반음절 구간에서 확률값이 큰 3개의 값을 선택한 후, 각각의 선택된 다음의 반음절+반음절 패턴 10개와 비교를 한다. 그 30개의 확률값 중 가장 큰 3개를 선별한 후, 선택된 뒤쪽의 반음절 정보와 연결된 반음절+반음절 패턴들은 두 번째 반음절+반음절 구간과 인식을 수행한다.

마지막 반음절 구간에서는 선택된 3개의 반음절+반음절의 마지막 음소를 반음절 패턴과 비교한 후 세 개의 확률값 중 가장 큰 확률값을 가지는 반음절 열을 결정한다. 그림 7은 규칙을 이용한 세자리 숫자음 인식 과정의 예를 나타내었다.

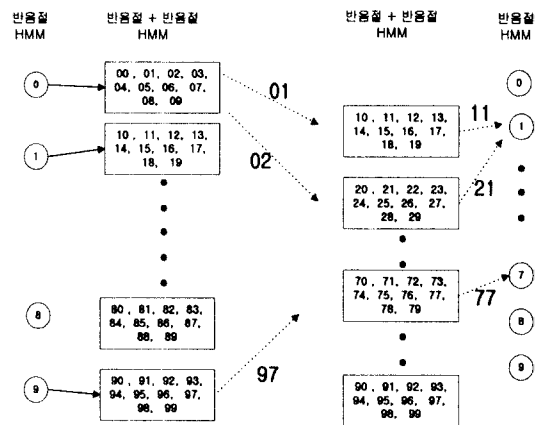


그림 7. 규칙을 이용한 반음절 열의 인식 과정
 Fig. 7 Recognition process of semisyllable sequence using the rule.

그림 7에서 처음에 세그멘테이션된 반음절은 표준 반음절 패턴 10개와 비교하여 확률값이 큰 3개, 0, 1, 9를 선정하고, 다음 인식구간은 0의 반음절로부터 얻은 정보를 가지고 다음 반음절+반음절 (00~09), 1은 반음절+반음절(10~19), 9는 반음절+반음절(90~99)표준패턴과 비교하여 확률값을 구한다. 구한 열 30개(00-09, 10-19, 90-99)의 확률값 중 가장 큰 값 01, 02, 97을 선택한다. 다음 인

식구간에서는 01은 반응절+반음절(10~19)패턴, 02는 반응절+반음절(20~29)패턴, 97은 반응절+반음절(70~79)패턴과 비교한다. 두 번째 반응절+반음절 인식구간에서 11, 21, 77이 결정되면, 반응절의 뒷부분의 1과 7의 패턴에 대한 확률값을 구한 후 인식 후보로 결정된 011, 121, 977 열 가운데에서, 가장 높은 확률값을 가지는 열을 최종 인식 숫자열로 결정한다.

제한된 반응절, 반응절+반음절 인식단위는 그 패턴 안에 다음 숫자열의 정보를 포함하고 있기 때문에 모든 인식대상과 비교하는 것이 아니라, 다음 숫자열의 정보를 가지고 있는 HMM과 비교하므로 인식 속도의 향상을 가져오게 된다.

IV. 실험 및 결과

본 논문에서 사용한 음성 데이터는 소음이 45~50dB 정도 되는 사무실 환경에서 남성 화자 30명이 0~9의 단독 발성을 3회씩 발성한 것과 00~99까지 두자리 숫자음 1회씩 자연스럽게 발성한 음성 신호를 16bit, 11.025kHz로 샘플링하여 저장하고, 이로부터 반응절 20종 각각 90개와 반응절+반음절 100종 각각 30개를 훈련 데이터로 사용하였다.

음성신호는 20ms의 프레임 단위로 Hamming window를 사용하였고, 10ms씩 중첩하여 14차 가중켄트럼계수를 구하였다. K-means 알고리즘[9]을 이용한 벡터 양자화에서 코드북은 32, 64, 128, 256 레벨로 하였다.

HMM에 의한 학습은 Baum-Welch 알고리즘을 사용하였고, zero의 확률을 방지하기 위해 HMM의 모든 파라미터 값에 대해 임의의 최저치 이하의 값을 최저치로 대체하여 계산하였고, 계산의 편이성을 위하여 log함수로 대체하여 계산했다. 또한, 전향 후향 확률을 계산할 때 시간이 증가할수록 파라미터 재추정시 사용하는 α, β 값이 작아져 underflow가 발생하기 때문에 α, β 값에 대한 스케일링 상수는 임의의 상태 i 에서 발생하는 모든 α, β 값을 관측 심벌열의 길이에 대해 합한 값을 사용하여 해결하였다.[10]

인식 실험 데이터는 종속 화자 5명, 독립 화자 5명이 한 자릿수(0~9) 10개, 두 자릿수(00~99) 100개, 세 자릿수, 네 자릿수, 다섯 자릿수는 각각 임의의 수 100개씩의 숫자음을 1번씩 자연스럽게 발성하여 실험하였다.

입력된 음성은 자동으로 끝점 추출 및 반응절 분할을 하였고, 인식대상이 되는 음의 처음 반응절 인식 후 예측된 다음 반응절+반음절 10개를 결합해 가면서 인식하는 방법으로 실험하였다.

한자리 숫자음의 실험 데이터는 종속, 독립실험에서 각각 50개, 그 이상의 자릿수에서는 각각 500개의 데이터를 사용하였고, 인식률은 연속 숫자음을 기준으로 한 연속 숫자 인식률과 단음절을 기준으로 한 단위 숫자별 인식률로 나누어 결과표를 작성하였으며, 인식률은 소수점 둘째 자리에서 반올림하였다.

한자리 숫자음에 대한 인식 결과는 표 1, 두자리 숫자

음에 대한 인식 결과는 표 2, 세 자리 숫자음에 대한 인식 결과는 표 3, 네 자리 숫자음에 대한 인식 결과는 표 4, 다섯 자리 숫자음에 대한 인식 결과를 표 5에 나타내었다.

표 1. 한자리 숫자음의 인식 결과

Table 1. Recognition results of a single digit.

(단위:%)

VQ Level	인식방법	화자 종속	화자 독립
32		96	96
64		98	90
128		94	92
256		98	92

표 2. 두자리 숫자음의 인식 결과

Table 2. Recognition results of two digits.

(단위:%)

VQ Level	인식방법	화자 종속		화자 독립	
		연속두자리 인식률	단위숫자별 인식률	연속두자리 인식률	단위숫자별 인식률
	32	86.2	94.1	73.6	86.0
64	89.0	94.4	81.2	89.7	
128	86.2	93.4	80.2	89.4	
256	89.6	94.6	80.0	89.6	

표 3. 세 자리 숫자음의 인식 결과

Table 3. Recognition results of three digits.

(단위:%)

VQ Level	인식방법	화자 종속		화자 독립	
		연속세자리 인식률	단위숫자별 인식률	연속세자리 인식률	단위숫자별 인식률
	32	83.6	93.5	74.8	90.2
64	88.0	95.4	75.8	89.9	
128	86.6	95.2	72.4	89.3	
256	86.8	95.2	75.6	90.7	

표 4. 네 자리 숫자음의 인식 결과

Table 4. Recognition results of four digits.

(단위:%)

VQ Level	인식방법	화자 종속		화자 독립	
		연속네자리 인식률	단위숫자별 인식률	연속네자리 인식률	단위숫자별 인식률
	32	73.0	92.4	58.6	86.1
64	77.4	93.5	65.0	88.8	
128	72.4	91.8	62.8	88.6	
256	75.4	93.2	63.0	88.7	

표 5. 다섯 자리 숫자음의 인식 결과
Table 5. Recognition results of five digits.

(단위: %)

VQ Level	화자 종속		화자 독립	
	연속다섯자리 인식률	단위숫자별 인식률	연속다섯자리 인식률	단위숫자별 인식률
32	72.2	92.7	57.2	87.3
64	73.8	93.8	61.6	89.3
128	73.0	92.8	60.4	89.4
256	70.8	92.7	63.2	90.9

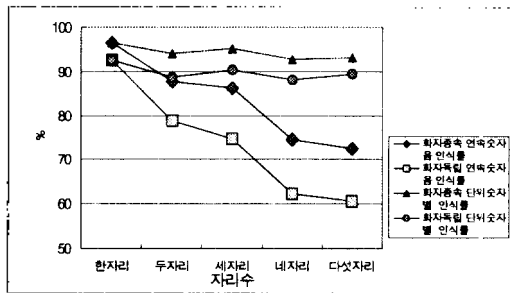


그림 8. 연속 숫자 인식률과 단위 숫자별 인식률
Fig. 8 Total recognition rate and unit digit recognition rate.

또한 자릿수에 따른 연속음 인식률과 단위숫자별 인식 결과를 그림 8에 나타내었다. 한자리 숫자음은 단일 자릿수이기 때문에 자릿수에 따른 인식률과 단위 숫자별 인식률이 같으며, 화자종속 단위숫자별 인식률은 98~91.8%, 화자독립 단위 숫자별 인식률은 96~86%를 나타내기 때문에 자릿수가 많아짐에 따라 모든 자릿수를 맞추는 연속 숫자 음성인식률이 낮아짐을 알 수 있다.

자릿수에 따른 화자 독립 실험의 단위 숫자별 인식률의 평균을 나타낸 표 6을 보면, 일반적으로 높은 자릿수보다는 낮은 자릿수의 인식률이 낮아지는 것을 볼 수 있다.

그림 9와 그림 10에는 코드북의 크기와 자릿수에 따른 단위 숫자별 인식률을 나타내었다. 화자종속의 경우 코드북의 크기가 64인 경우에 가장 좋은 인식률을 보였으며, 화자독립인 경우에는 코드북의 크기가 256인 경우 한자릿수를 제외하고 좋은 인식률을 나타내었다.

표 6. 자릿수에 따른 단위 숫자별 인식률
Table 6. Unit digit recognition rate classified by digits.

단위숫자 인식률 VQ Level	두자리			세자리				네자리					다섯자리					
	자릿수에따른인식률			자릿수에따른인식률				자릿수에따른인식률					자릿수에따른인식률					
	10자리	1자리	평균	100자리	10자리	1자리	평균	1000자리	100자리	10자리	1자리	평균	10000자리	1000자리	100자리	10자리	1자리	평균
VQ32	86.2	85.8	86	92.4	90.4	88	90.2	87.4	87	85.2	84.8	86.1	90.6	87.6	88.2	84.6	85.4	87.3
VQ64	92	87.4	89.7	93	90.6	86.2	89.9	89	90.6	87.6	88	88.8	91.8	89.4	88.4	89	88	89.3
VQ128	91.6	87.2	89.4	93.8	89	85.2	89.3	91.2	89.8	87.4	86	88.6	91.8	90	90.2	89	86.2	89.4
VQ256	92.4	86.8	89.6	95.6	89.6	87	90.7	92.6	89.2	86.8	86.2	88.7	93.2	91.8	93	88.4	88	90.9

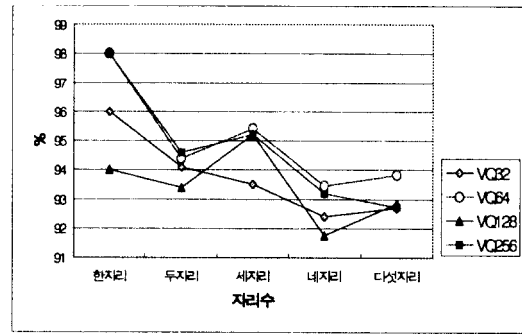


그림 9. 화자종속 단위 숫자별 인식률
Fig. 9 Speaker dependent unit digit recognition rate.

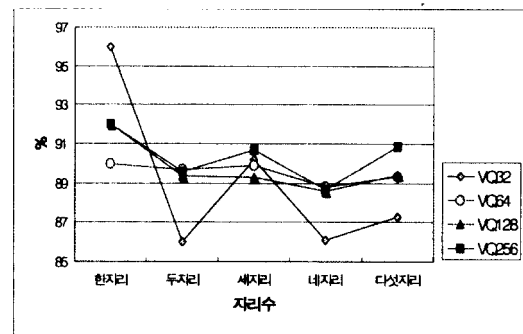


그림 10. 화자독립 단위 숫자별 인식률
Fig. 10 Speaker independent unit digit recognition rate.

긴 숫자열의 낮은 자릿수가 높은 숫자음의 인식률을 높이기 위해서는 긴 자릿수를 갖는 숫자음의 반음절+반음절 단위를 추가로 훈련 데이터로 사용하여 숫자음의 자릿수를 고려한 반음절+반음절 HMM을 이용하고, 화자독립 인식률을 높이기 위한 주파수 Warping기법[11]과 GDP를 도입한 화자적응기법[12]을 활용하면 더욱 좋은 연속 숫자 음성인식 시스템을 구성할 수 있을 것으로 생각된다.

V. 결 론

본 논문에서는 음향학적 특징이 안정된 모음구간의 분

할에 의한 반음절, 반음절+반음절 단위 HMM에 의한 연속 숫자음 인식 방법을 제안하고 인식 시스템을 구축하여 그 성능을 확인하였다.

한자리 숫자 단위로 인식할 경우에는 음소간의 조음 현상이 숫자를 구별하는 데 큰 어려움을 주지 않지만, 연속어로 인식 범위가 확장될 경우에는 음소간, 단어간 조음 현상 때문에 단위숫자의 경계가 상당히 애매모호해진다. 이러한 조음 현상을 수용하기 위해서는 음소 또는 단위 숫자보다 개념적으로 확대된 인식 단위가 필요로 하게 된다. 따라서, 음향학적으로 안정된 모음 부분에서 분할할 수 있는 반음절, 반음절+반음절 인식 단위를 제안하고, 인식열을 반음절, 반음절+반음절 패턴으로 분할하여 규칙에 따라 반음절열을 추적 결합하는 방법을 제안하여 그 성능을 확인하였다. 한국어 숫자음의 인식률은 다른 인식 대상에 비해 상대적으로 저조한 인식률을 보인다. 이것은 한국어 숫자음이 모두 단음절이라는 점 외에도 숫자음 조합에는 규칙성이 존재하지 않기 때문이다. 본 논문에서는 인식 단위로 반음절, 반음절+반음절 단위 패턴을 제안하여 이용하였기 때문에, 연속 발성의 경우에 음절 또는 음소 등의 부단어 단위의 분할시에 발생하는 오인식을 피할 수 있게 된다.

또한, 제안한 방법은 인식단위가 반음절, 반음절+반음절 단위에 다음에 올 반음절 패턴 정보가 포함되어 있기 때문에 규칙을 적용하여 실시간 숫자음 인식이 가능했다.

제안한 방법은 전화번호, 주민등록번호 등과 같은 통신망을 이용한 정보 및 금융 서비스에서 키보드의 입력이 아닌 음성만을 이용한 연속 숫자음 인식에 적용이 가능하고 차후에 한국어 내어줘 연속 음성인식 시스템에 적용이 가능하리라 생각된다.

참고 문헌

1. 양진우, 김순협, "HMM과 연결 숫자음의 후처리를 이용한 음성 다이얼링에 관한 연구," 한국음향학회지 VOL.14 No.5, pp. 74-82, 1995.
2. J.G. Wilpon and L.R. Rabiner "Improvement in Connected Digit Recognition Using Higher Order Spectral and Energy Features," Proc. ICASSP VOL.1, pp. 349-352, 1991.
3. 김순협 외 4인, "음소 단위에 의한 한국어 연속 숫자음 인식에 관한 연구," 한국음향학회지 VOL.8 No. 3, pp. 5-15, 1989.
4. 박현상 외 3인, "Diphone단위의 hidden Markov model을 이용한 한국어 단어인식," 한국음향학회지 VOL.13 No. 1, pp. 14-23, 1994.
5. J.C. Junqua, J.P. Haton, *Robustness in Automatic Speech Recognition*, Kluwer Academic Publishers, Boston, 1996.
6. P.C. Woodland, C.J. Leggetter, J.Odell, V.Valtchev, S.J.Young, "The 1994 HTK Large Vocabulary Speech Recognition System," Proc ICASSP, Vol.1, pp. 73-76, 1995.
7. M.Hahn and C.K. Park, "An Improved Speech Detection Algorithm for Isolated Korean Utterance," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 525-528, 1992.
8. Kai-Fu Lee, *Automatic Speech Recognition: The Development*

of the SPHINX system, Kluwer Academic Publishers, Boston, 1989.

9. J.G. Wilpon, L.R. Rabiner, "A Modified K-Means Clustering Algorithms for use in Isolated Word Recognition," IEEE Trans. on Acoust. Speech and Signal Proc., Vol. 33, No. 3, pp. 587-594, 1985.
10. X.D.Huang, Y.Ariki, M.A. Jack, *Hidden Markov Models for Speech Recognition*, Edinburgh University Press, Edinburgh, 1990.
11. L.Lee, R.Rose, "A Frequency Warping Approach to Speaker Normalization," IEEE Trans. on Speech and Audio Proc. Vol. 6 No. 1, pp. 49-60, 1998.
12. Rucy-Chung Shyu, Jhing-Fa Wang, "A Vowel-Driven Mandarin Speech Autodialer with Adaptation Ability," IEEE Signal Proc., Vol. 4 No. 6, pp. 164-166, 1997.

▲ 윤 재 선 (Jeh Seon Youn)



1996년 2월 : 성균관대학교 전자공학과
공학사

1998년 2월 : 성균관대학교 전자공학과
공학석사

1998년 3월 ~ 현재 : 성균관대학교 전
기 전자 및 컴퓨터 공
학부 박사과정

※주관심분야: 음성 인식 및 통신 신호 처리

E-Mail: sunhci@bible.skku.ac.kr

▲ 홍 광 석 (Kwang Seok Hong)



1985년 2월 : 성균관대학교 전자공학과
공학사

1988년 2월 : 성균관대학교 전자공학과
공학석사

1992년 2월 : 성균관대학교 전자공학과
공학박사

1990년 3월 ~ 1993년 2월 : 서울보건
문대학교 전산정보처리
과 전임강사

1993년 3월 ~ 1995년 2월 : 제주대학교 정보공학과 전임강사

1995년 3월 ~ 1996년 2월 : 성균관대학교 전자공학과 조교수

1996년 2월 ~ 현재 : 성균관대학교 전기 전자 및 컴퓨터 공
학부 조교수

※주관심분야: 음성 인식 및 신호 처리

E-Mail: kshong@yurim.skku.ac.kr