

구문 분석과 Level Building을 이용한 한국어 연속음 인식

The Recognition of Korean Continuous Speech using Syntactic Analysis and Level Building

* 안 태 옥(Ann, T. O.)
** 변 용 규(Byeon, Y. G.)
*** 김 순 협(Kim, S. H.)

요 약

본 논문은 특정 화자에 대한 한국어 연속음의 효율적인 인식을 위하여, 구문분석과 OGS(Ordered Graph Search) 기법으로 변형시킨 Level Building을 이용한 인식시스템에 대해서 제안하고 있다. 본 시스템에서 사용하는 template는 연속음을 분할(Segmentation)시킨 단독음이며 소구간 경로 및 본 논문에서 제안한 선제 경로 제약에 의해 거리 계산 값이 최소인 "super" reference를 구함으로써 인식된다. 본 연구에서 사용한 연속음은 단독음 11자로 구성된 13 개의 전철역명으로서 2인의 남성과 1인의 여성화자에 의해 10번씩 발음한 130단어를 테스트하였는데 97.7%의 단어인식율을 보였다.

ABSTRACT

In this paper, in order to recognize effectively the Korean continuous speech for dependent speakers, a recognition system is proposed using syntactic analysis and level Building that is modified by OGS method. The templates used in this system are the syllable to segment the continuous speech, and input data are recognized by finding the "super reference" to minimize the distance that is computed by local path constraints and the global path constraints proposed in this paper. The continuous speeches used in this search are the subway stations, and the 130 words pronounced 10 times for the different 13 words

* 광운대학 전자계산공학과
** 경기공업 개방대학 전산과 부교수
*** 광운대학 전자계산학과 교수

consisting of 11 characters of syllable by 2 male and 1 female, are tested and the recognition rate of 97.7% is found.

I. 서 론

음성 인식은 음성을 분석하여 그 중에 포함된 언어적 특징을 추출하고 그것에 의해 언어 기호의 표시로 변환하는 처리이다. 이러한 처리를 실행하는 방법은 음성에 포함된 언어적 특징에 대한 표준 패턴을 미리 기억시켜 놓고 입력 음성과 비교하여 유사성에 의하여 판별하는 표준 패턴 대응법과 음향 분석 결과에 따라 순차적으로 분류하는 순차 분류법이 있는데 본 논문에서는 표준 패턴 대응법을 사용하였으며 두징 파라메타로 선형 예측계수와 자동상관 계수를 사용하고 있다.

그런데 음성은 동일한 피자에 있어서도 발생할 때 마다 인식 단어로 되는 음성의 지속 시간이 다르기 때문에 입력 자료와 표준 자료가 거의 유사한 패턴이라도 시간축 상의 대응 관계에 차이가 있게 되는데 이런 관계를 극복하기 위해 DP(Dynamic Programming) 알고리즘을 이용하여 음성을 인식시킨다.

음성 인식은 인식되는 음성의 종류에 따라 단속어 인식(Isolated Word Recognition), 연결어 인식(Connected Word Recognition) 및 연속음 인식(Continuous Word Recognition)으로 나눌 수 있다. 여기서, 연속음 인식은 연속된 음성을 음소(Phonemes), 음절(Syllables) 등의 음성의 기본단위들로부터 인식하는 것을 말한다.

본 연구에서는 기본 단위로 음절을 택했으며 연속음 대상으로 2, 3자의 전절역명을 선택했고 연속음을 인식 하는데 있어서 발음된 음성을 기본 단위로 나누지 않고, 단어 그 자체를 test pattern으로 하여 super reference와 matching시킴으로서 인식시키는 Level Building에 OGS 기법을 이용, 수정

하여 사용하였으며, 이 수정된 Level Building에 구문분석을 적용시킨 DTW인식 알고리즘으로 인식시켰다.

본 논문은 모두 IV절로 구성되어 있다. II절에서는 수정한 Level Building DTW 알고리즘을 토대로 해서 연속음으로 발음된 한국어 전절역명을 구문분석을 이용해서 인식시키는 DTW 알고리즘에 대해서 기술하였고, III절에서는 실험의 구성과 결과에 대해서 고찰했으며, 마지막으로 IV절에서는 결론을 내렸다.

II. 구문 분석과 Level Building을 이용한 한국어 연속음 인식

1. 알고리즘의 도입

본 연구에서는 인식대상으로 2, 3자의 연속음(전절역명) 인식에 대한 시스템을 묘사한 것으로 연속음 인식은 단속음 인식에 비해 복잡한 특성을 지닌다. 그래서 연속음을 표준 패턴으로 해서 인식시킬 수도 있지만 계산량이 커지고 대상어가 커지면 계산 용량에 문제가 생기게 된다. 따라서 문자열을 보나 작은 음성 단위의 집합으로 분할한 다음, 이들 음성단위로부터 주어진 규칙들에 의해 단어를 해석할 필요가 있다. 보통 사용되는 음성 단위는 프레임, 음소, 음절 등이 있다. 이러한 분할(Segmentation) 과정은 연속음 인식 방법의 첫 단계가 된다.

따라서, 본 연구에서는 M. K. Brown과 L. R. Rabiner가 제안한 OGS(Ordered Graph Search) 기법을 이용하고 가능한 모든 단어 계열의 Matching이 표준 패턴을 연결시켜 줌으로써 이루어 질 수 있다는 성질을 이용한 Level Building DTW 알고리즘을

수정하여 이 알고리즘에 Syntax를 적용시켰다.

결정 법칙은 Matching이 끝난 다음 어떤 Threshold이내의 전체누적된 거리값이 최소가 되는 문자열을 선택하기 위해 사용했다. 인식시스템을 여기에서 입력된 연속음의 분할점을 제공한다.

2. 구문 분석과 Level Building을 이용한 DTW 알고리즘

구문 분석과 Level Building을 이용한 DTW 알고리즘의 목적은 유한 State Parser내의 각 State에 대한 최상의 Temporal Segmentation을 가능케 하는 것과 모든 State들에 대한 Temporal Segmentation들을 가능케 하는 최적 Parser를 유지하는 것이다. 그래서 알고리즘은 구문에 따라 Reference Pattern의 열을 찾으려고 한다. 각각의 Reference Pattern은 전체 경로 제약과 소구간 경로 제약에 따라 Distance를 최소로 한다는 점에서 Test Pattern과 Match되는 최상의 것과 연속음에서 분할(Segmentation)한 단독음으로 표현된다.

Test Utterance $T(m)$ 은, $m = 1, 2, \dots, M$ 인, 다차원 특징 Vector의 Frame들의 열로 구성한다. Formal Grammar는 유한 State Automaton $A(Q, V, \delta, q_0, Z)$ 에 의해 받아들인다. 상태 천이 Diagram은 정점 즉, 상태들의 고정된 집합과 State를 연결하는 Edge 즉, Transition의 집합을 구성한다. 각기 그런 Edge들은 $v \in V$ 로써 표시된다. 여기에서 Q 는 State들의 집합이고, State의 상호연결의 정확한 방법이 천이 함수 δ 에 의해서 상징적으로 명시된다. 거기서,

$$\delta : (Q \times V) \rightarrow Q \quad (2-1)$$

즉, 만약 상태 $q \in Q$ 가 $v \in V$ 로 표시된 Edge에 의해 다른 State $s \in Q$ 와 연결된다면

$$\delta(q, v) = s \quad (2-2)$$

$q_0 \in Q$ 는 초기 상태이고 $Z \subseteq Q$ 는 Terminate state들의 집합이다. 이런 개념의 도표를 그림 1에서 보여준다. Terminate State는 Asterisk로 표시한다.

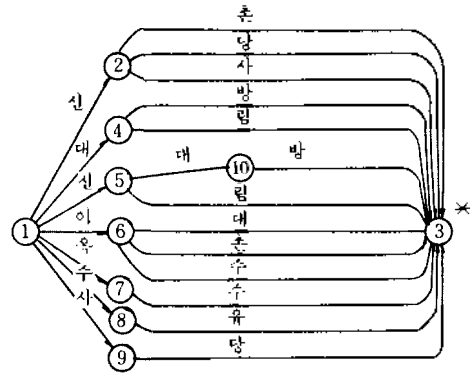


그림 1 실험에서의 상태 천이도
The state transition for the experiment.

Language는 FSA(Finite State Automaton)가 유한한 것으로 받아들여진다. 그래서 State들 $q_0, q_1, q_2, \dots, q_{|Q|-1}$ 의 수가 가능하며 만약 q_j 가 q_i 까지 가는 Transition이라면 $1 > j$ 이다. 완전하고 효율있게 Test Utterance를 Parse하기 위해서는 주어진 Test Utterance에 대해 "Super" State들의 집합 $\{[s, m]\}$ 은 $s \in Q$ 와 $m \in \{0, 1, \dots, M\}$ 을 정의한다. 만약 state q_0 에서 state s 까지 transition의 열이 존재한다면 test utterance는 "super" state $[s, m]$ 에 구문 분석된다. 그래서 reference pattern들의 열은 frame 1과 m 사이의 test pattern의 portion과 match되는 이러한 transition에 따른 단독음과 관계가 있다는 것이 포함된다. "best" distance는 각 state를 유지하고 이런 distance는, 그 state에 대한 모든 transition을 통하여, 각 transition과 연관된 distance에 original state와 관련된 "best" distance가 더해진 최소값을 생성한다.

공식적으로

$$D_s = \min_q [D_q + d_{qv}] \quad (2-3a)$$

$q, v, s. s. t. \delta(q, v) = s$

$$D_{q_0} = 0 \quad (2-3b)$$

여기서 D_s 는 State s 와 관련된 최상의 distance 이고 d_{qv} 는 단독음 v 를 사용하여 State q 로부터 state s 까지의 transition과 관련된 distance이다.

최적 parse 외에도 최적 temporal segmentation을 얻기위해서 super state $[s, m]$ 에서 "best" distance를 계산한다.

$$D_{[s, m]} = \min_q [\min_b [D_{[q, b]} + d_{q, b|v}]] \quad (2-4a)$$

$q, v, s. t. \delta(q, v) = s \quad 0 \leq b \leq m-1$

$$D_{[q_0, 0]} = 0 \quad (2-4b)$$

여기서 $D_{[s, m]}$ 은 "super" state $[s, m]$ 에서 "best" distance 이고 $d_{q, b|v}$ 는 그림 2에서 설명한 것처럼 $b+1$ 과 m frame간의 test pattern의 일부와 reference 단독음 v 의 matching과 관련된 DTWdistance이다. 그림에서 N_v 는 단독음 v 에 있어서 Reference Pattern의 Frame수이고, R_v 는 단독음 v 에 있어서 Reference template이다.

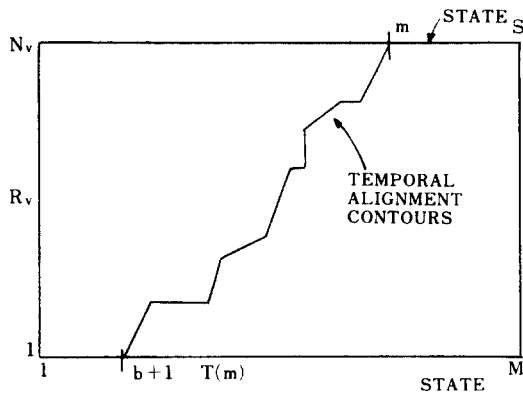


그림 2 표준 패턴과 시험 패턴의 부분간의 시간정렬
Time alignment between a reference pattern and a portion of the test pattern.

주어진 transition과 모든 ending frame m 에 대한 식(2-4)의 전대부 최소화가 그림 3에서 설명된다. 그림에서 $D_q^s(m)$ 은 State q 에서의 Super Reference의 Frame m 까지의 누적된 최소 거리를 말하고 $D_{q_0}^{s, v}(m)$ 은 State q 로부터 State s 까지에서 Reference R_v 를 사용한 Super Reference의 Frame m 까지의 누적된 거리를 말한다.

본 연구에서는 이 구문 분석에 의한 DTW알고리즘을 OGS방법을 도입한 수정된 Level Building DTW알고리즘에 적용시켰다. 그림 3을 보면 Level Building DTW알고리즘이 유용한 알고리즘이라는 것을 알 수 있다.

여기에서는 Myer와 Rabiner에 의한 Level Building procedure를 다음과 같이 수정하여 DTW알고리즘으로 사용하였다.

- 1) 각 level에서 candidate가 될 수 있는 것은 단독음의 pattern에 의해서가 아니라 단지 state에만 일치한다.
- 2) OGS 기법을 사용해서 계산 시간 및 계산량을 줄였다. 그림 4(a)에서 볼 수 있는 바와 같이 일반 DTW알고리즘에서는 점으로 표시된 모든 Grid 점을 소구간 경로 제약 안에서 계속 비교

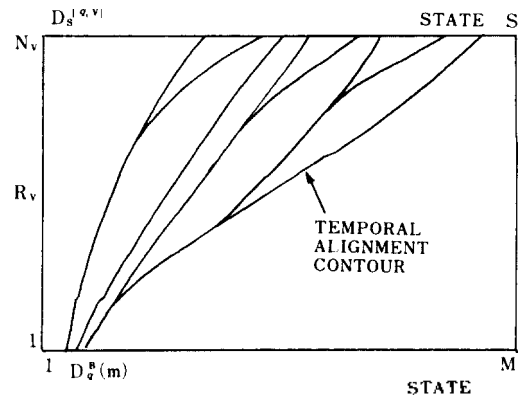
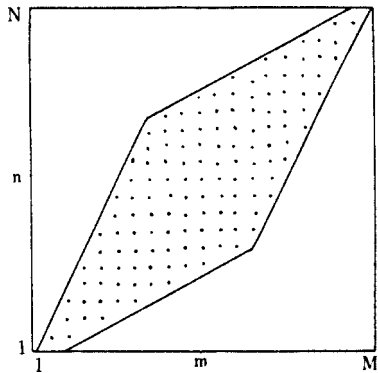
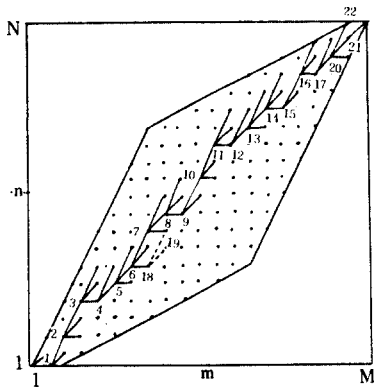


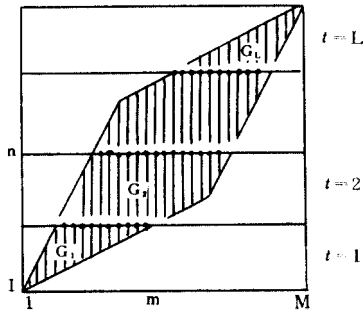
그림 3 주어진 전이에 대한 Level Building 절차의 예
Illustration of the level building procedure for a given transition.



(a) 일반 DTW 알고리즘
(a) General DWT algorithm



(b) OGS 알고리즘
(b) OGS algorithm



(c) Level Building DTW 알고리즘
(c) Level Building DTW algorithm

그림 4 DTW 알고리즘
DTW algorithm

하고 진행해 나가야 하는데 반하여 그림 4(b)는 OGS 기법에 대한 그림으로 소구간 경로 제약에 의해 Grid 점을 통하여 가장 짧은 경로를 직접 탐색해 나가는 것을 보여 주는 그림으로 모든 Grid 점을 통해 진행되는 것이 아니므로 일반 DTW 알고리즘보다 수행 시간이 3 배 빠르고 계산량도 줄어든다⁴⁾.

3) 단어는 i 번째 state에서 출발하는 그런 단독음만 i 번째에서 match 되도록 구분되었다. 따라서, 그림 4(c)에서 볼 수 있는 바와 같이 Level Building에 있어서는 Reference Pattern으로 쓰이는 단독음 모음에 대하여 각 Level 마다 비교해 주어야 함으로 Template을 No라 하면 Threshold 값을 주지 않고 Level을 L이라 했을 경우 No'번 비교해 주어야 하는데 반하여 구문 분석적인 방법을 사용하며 Syntax 관계에 따라 비교 횟수가 감소 된다.

따라서, Reference Template가 많으면 많을수록 그리고 음절의 수가 길면 길수록 Level Building은 지수적으로 증가한다.

인식효율도는 그림 5와 같다.

III. 인식실험 및 결과

본 연구에서는 인식 실험에서 단독음 11사(수, 이, 사, 신, 내, 옥, 촌, 유, 림, 당, 방)와 연속음 13자(이내, 이촌, 이수, 옥수, 수유, 신촌, 신당, 신사, 사당, 대림, 친림, 신대방, 대방)을 성인 남성 2인과 여성 1인에 의해 발음된 음성으로서 단독음은 연속음에서 segmentation시킨 단독음을 template로 쓸 때 참고로 삼기 위해서 받았으며 여기에서 음성의 발음시간은 단독음이 0.3-0.6(sec)이고 연속음이 0.6-1.0sec 정도이었다. 인식대상 어휘수는 연속음 13자에 대해 두 남성 화자가 4 번씩, 그리고 여성 화자가 2 번 반복 발성한 것으로 130개이다.

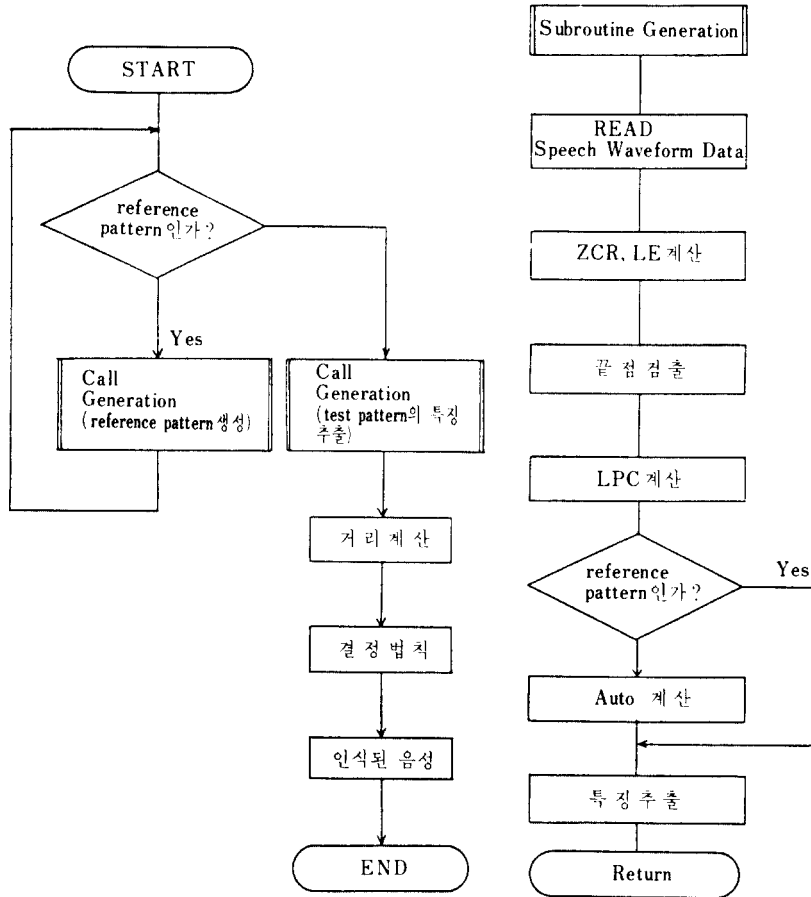


그림 5 인식 흐름도
Recognition flowchart

1. 인식시스템 구성

실험에 사용된 모든 연속음(전철역명)이 릴 테이프에 녹음되었다. 녹음은 스튜디오에서 행해졌으므로 방음장치는 양호하였다.

녹음기에서 나온 음성신호는 최대 출력 전압이 1[V]가 되도록 조절하여 차단 주파수가 4[KHz]인 대역 통과 여파기(Band Pass Filter;BPF)를 통과시킨 다음 MINC에 연결된 preamplifier를 통하여 preamp되고 12비트 A/D 변환기에 의해 12레벨로 양

사화 된다. 12비트 A/D 변환기에서 10[KHz]의 샘플링 주파수로 디지털화된 음성신호는 모뎀(Modem)을 통해 저장되었다.

본 연구에서는 이렇게 M/T에 저장된 데이터를 가지고 12차의 자동상관 계수와 12차의 LPC 계수를 표준 패턴과 시험 패턴의 특성 벡터로 설정하여 인식 실험을 수행하였다. 이때 특성 벡터 추출과 인식 실험에 사용한 컴퓨터는 VAX-11/750 미니컴퓨터이다.

본 실험의 시스템 구성은 다음과 같다.

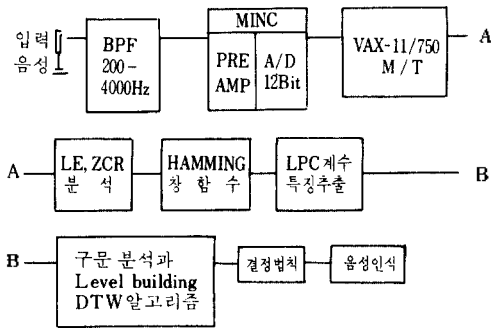


그림 6 한국어 연속음 인식에 관한 시스템 구성도.
Block diagram for the recognition of Korean continuous speech

2. 알고리즘의 수행

본 연구에서는 화자 종속 방법으로 음성 인식 실험을 다음과 같이 하였다.

a. 끝점 검출 및 연속음의 분할에 의한 표준패턴 설정

연속음을 ZCR, LE 을 이용하여 시작점과 끝점을 검출하고, 또한 이 ZCR, LE 을 이용하여 연속음을 분할 표준 패턴으로 삼았다.

b. 거리 측정법

$$d(r, t) = \log \left[\frac{a_r \cdot V_r \cdot a_r}{a_t \cdot V_t \cdot a_t} \right] = \log \frac{r}{t} \quad (3-1)$$

$a_r = (a_1, a_2, \dots, a_{12})$: 표준 패턴 분석 구간의 LPC 계수

$a_t = (a_1, a_2, \dots, a_{12})$: 입력패턴 분석 구간의 LPC 계수

V_r = 입력 패턴 분석 구간의 자동 상관행렬이다.

미리 설정된 표준 패턴과 입력 패턴간의 거리 측정은 Itakura 가 제안한 LPC 거리 측정법을 거리측

정의 척도로 삼았다. 즉, 표준패턴의 분석 구간 R 과 입력 패턴의 분석 구간 T와 의 거리는 윗식과 같이 표현된다.

c. 제약 조건

DTW 알고리즘에 의해 최적경로를 구할 경우 (m, n) 평면상에서 이 경로가 지날 수 있는 범위에 대해 몇가지 제약조건을 설정하였다.

1) 끝점 제약

$$i(1) = 1, j(1) = 1; \text{ 시작점 제약} \quad (3-2a)$$

$$i(K) = N, j(K) = M; \text{ 끝점 제약} \quad (3-2b)$$

2) 경로 제약

가. 단조 증가 조건

$$i(k+1) > i(k) \quad (3-3a)$$

$$j(k+1) > j(k) \quad (3-3b)$$

나. 소구간 경로 제약

점 $(1, 1)$ 에서부터 점 (m, n) 까지 최적 경로를 따라 더하여 누적된 거리 $D(m, n)$ 는 다음 식으로 표현된다.

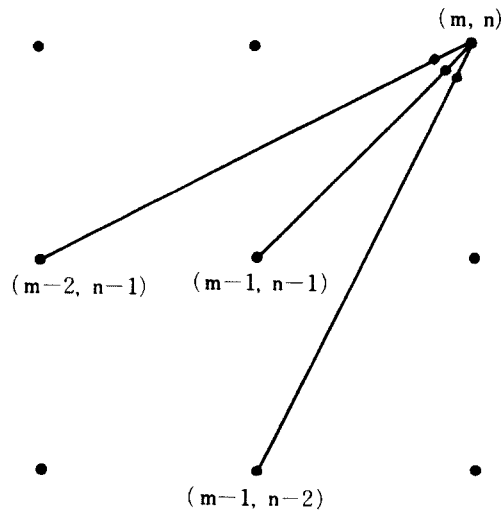


그림 7 소구간 경로 제약

The function of the local path constraint.

$$D(m, n) = \text{Min}[D(m-1, n-2) + d(m, n), \\ D(m-1, n-1) + d(m, n), \\ D(m-2, n-1) + d(m, n)] \quad (3-4)$$

다. 전체 경로 제약

전체 경로 제약 때문에 최적 워핑 경로가 진행할 수 있는 (m, n) 평면이 제한되고, 계산에 필요한 기억 용량의 감소와 수행 시간의 절약을 꾀할수 있다. 따라서 본 연구에서는 다음과 같이 전체 경로 제약을 제안하였다. 전체 경로 제약의 허용범위를 수식으로 표시하면 다음과 같다.

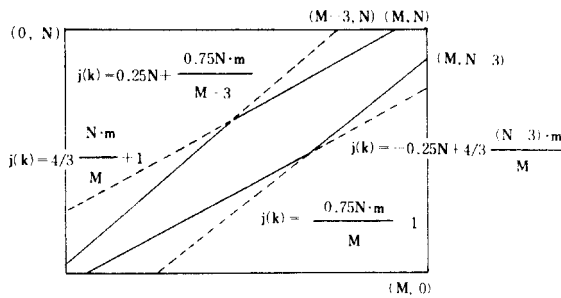


그림 8 제안된 Warping 함수의 전체 경로 제약 Wa
The legal range by proposed global range con-
straint of the warping function

$$j(k) < \text{Min} \left(0.25N + \frac{0.75N \cdot m}{M - 3}, 4/3 \cdot \frac{N \cdot m}{M} + 1 \right) \quad (3-5a)$$

$$j(k) > \text{Max} \left(0.25N + 4/3 \cdot \frac{(N - 3) \cdot m}{M}, \frac{0.75N \cdot m}{M} - 1 \right) \quad (3-5b)$$

식 (3-5)은 시작점 (1, 1)로부터 끝점 (M, N) 까지 적당한 경로를 통해 도달되는 그리드점들의 영역을 제한하는 것이다. 본 연구의 전체경로 제약에서는 시작점과 끝점의 검출시 약간의 잘못이 있다

라도 시작점과 끝점에 융통성을 주었기 때문에 오차를 줄일 수 있으며 또한 중간 과정에서도 기준치를 크게 줄이므로써 계산량을 줄였다.

3. 실험 결과

인식 시스템을 통해 성인 남성 2인 (각자 13단어를 4번씩 반복)과 여성 1인 (13단어를 2번씩 반복)이 발음한 130개의 연속음을 대상으로 구문 분석을 이용한 DTW 알고리즘을 적용시켜 실험한 결과 다음과 같다.

a. 수행 시간

수행 시간면에서 생각해 보면 OGS를 사용하면 원래 Level Building DTW 알고리즘보다 3배 정도 빠르며 본 논문에서 사용하는 시스템을 이용하면 Level Building DTW 알고리즘보다 5배에서 10배

남자A의 인식 결과

원리	이대	이준	이수	옥수	수유	진춘	진남	진사	사당	대림	진대방	내방
이대	4											
이준		4										
이수			4									
옥수				4								
수유					4							
진춘						4						
진남							4	1				
진사								3				
사당									4			
대림										4		
진대방											4	
내방												4

표 인식결과
Recognition results.

• 남성 화자 B의 인식 결과

입력 출력	이 대	이 촌	이 수	옥 수	수 유	신 촌	신 당	신 사	사 당	대 림	신 림	신 대방	대 방
이대	4												
이촌		4											
이수			4										
옥수				4									
수유					4							1	
신촌						4							
신당							4						
신사								4					
사당									4				
대림										4			
신림											4		
신대방												3	
대방													4

- Continue -

정도까지 빨라진다.

b. 인식율

남성 화자는 각기 98%의 인식율을 보였으며 여성 화자는 96%의 인식율을 보였다. 각 화자의 각 단어에 대한 인식 결과는 다음과 같다.

IV. 결 론

본 실험은 특정 화자에 대해 구문 분석 방법과 Level Building을 이용하여 한국어 연속음을 인식함에 있어서 조음 효과와 프레임 수의 차를 고려해서 연속어를 분할하여 얻어진 단독음들을 조음효과를 고려해서 만든 표준 패턴을 써서 DTW알고리즘을 적용한 것이다. 그 결과 조음효과를 고려해서 표

• 여성 화자의 인식 결과

입력 출력	이 대	이 촌	이 수	옥 수	수 유	신 촌	신 당	신 사	사 당	대 림	신 림	신 대방	대 방
이대	2												
이촌		2											
이수			2										
옥수				2									
수유					2								
신촌						2							
신당							2						
신사								2					
사당									2				
대림										2			
신림											1		
신대방												1	2
대방													2

- Continue -

준 패턴을 설정하여야 되었으므로 ‘신림’ 같은 경우는 ‘신’이 ‘실’로 발음됨으로 표준 패턴을 따로 잡아주지 않으면 안되었다.

또한 본 연구에서는 구문 분석과 수정된 Level Building을 이용하였으며 이 수정된 Level Building DTW 알고리즘은 거리를 계산하는데 많은 경로를 거쳐야 함으로 시간이 많이 걸려, 이 알고리즘에 OGS 기법을 도입해서 수행시간을 줄였다.

그런데, 여기서 구문 분석 방법과 더불어 OGS 기법은 시간을 단축시키나 오인식율을 증가시킴으로 그림 7와 같은 소구간 경로 제약과 새롭게 제안한 전체 경로 제약에 의해 오인식율을 줄였다.

본 연구에서는 13개의 서로 다른 언어에 대해 130번의 발음을 하였으며 인식율 97.7%를 얻었다. 그러나, 이 인식율은 만족할 만한 인식율이 못 되며 더 좋은 인식율을 얻기 위해서는 한국어에 있어서

의 음운학적인, 음향학적인 특성을 알아본 후에 연속음을 분할하여 표준패턴을 설정하는 것이 더 좋을 것 같으며 좀 더 연구하여 단음음을 표준패턴으로 삼아 인식해도 좋은 인식율을 얻을 수 있도록 노력해야 할 것이다. 또한 앞으로는 실시간 처리 및 강력한 인식을 위한 알고리즘이 개발되어야 할 것이다.

참 고 문 헌

- [1] F. Itakura, "Minimum Prediction Residual Principle Applied to Speech Recognition," IEEE Trans., on Acoustics, Speech and Signal Processing, Vol. ASSP-23, pp.67-72, Feb. 1975.
- [2] L.R. Rabiner, J.P. Wilpon, and A.E. Rosenberg, "A Voice Controlled Repertory-dialer System," Bell Syst., Tech., J., Vol.59, pp.1153-1163, Sep. 1980.
- [3] C.S. Myers and L.R. Rabiner, "Connected Digits Recognition using a Level Building DTW Algorithm," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-29, pp.351-363, June 1981.
- [4] M.K. Brown and L.R. Rabiner, "An Adaptive, Ordered, Graph Search Technique for Dynamic Time Warping for Isolated Word Recognition," IEEE. Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-30, No.4, pp.535-543, Aug. 1982.
- [5] S.E. Levinson, "The Effects of Analysis on Word Recognition Accuracy," The Bell System Technical Journal, Vol.57, No.5, pp.1627-1644, May-June 1978.
- [6] L.R. Rabiner and S.E. Levinson, "Isolated and Connected Word Recognition Theory and Selected Applications," IEEE Trans. on Communication, Vol. COM-29, No.5, pp.621-639, May 1981.
- [7] J.D. Markel and A.H. Gray, Linear Prediction of Speech, Spring-Verlag Berlin Heidelberg, 1976.
- [8] S.E. Levinson, A.E. Rosenberg, and J.L. Flanagan, "Evaluation of a Word Recognition System Using Syntax Analysis," The Bell System Technical Journal, Vol.57, No.5, pp.1619-1626, May-June 1978.
- [9] C.S. Myers and S.E. Levinson, "Speaker independent connected word recognition using a syntax-directed dynamic programming procedure," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-30, pp.561-565, Aug. 1982.
- [10] C.S. Myers and L.R. Rabiner, "A Level Building Dynamic Time Warping Algorithm for Connected Word Recognition" IEEE. Trans. Acoust., Speech, Signal Processing, Vol. ASSP-29, No.2 pp.284-296, April 1981.