

논문 2004-41SP-3-22

구문 분석과 One-Stage DMS/DP를 이용한 연속음 인식

(Continuous Speech Recognition using Syntactic Analysis and One-Stage DMS/DP)

안 태 옥*

(Tae-Ock Ann)

요 약

본 논문은 연속음 인식에 관한 연구로써, 구문 분석을 이용한 One-Stage DMS/DP에 의한 음성 인식 방법을 사용한다. 인식 실험을 위해 우선 구간 구분화 알고리즘을 이용하여 DMS(Dynamic Multi-Section) 모델을 만들며, 구문 분석을 이용한 One-Stage DMS/DP 방법으로 연속음 데이터를 인식하게 하였다. 제안된 방법에 의한 인식 실험을 수행하는 것 외에도 비교를 위해 전통적인 One-Stage DP 방법을 같은 조건 같은 데이터를 가지고 수행하였다. 인식 실험 결과, 기존의 방법보다 구문 분석을 이용한 One-Stage DMS/DP 방법이 우수한 것으로 나타났다.

Abstract

This paper is a study on the recognition of continuous speech and uses a method of speech recognition using syntactic analysis and one-stage DMS/DP. In order to perform the speech recognition, first of all, we make DMS model by section division algorithm and let continuous speech data be recognized through One-stage DMS/DP method using syntactic analysis. Besides the speech recognition experiments of proposed method, we experiment the conventional one-stage DP method under the equivalent environment of data and conditions. From the recognition experiments, it is shown that One-stage DMS/DP using syntactic analysis is superior to conventional method.

Keywords: 구문 분석, DMS/DP, 조음 결합 현상

I. 서 론

본 연구는 연속음 인식에 대한 논문으로, Ney^[5]의 논문에 의하면 One-Stage DP 알고리즘은 Sakoe의 Two-Level DP 알고리즘의 계산 시간의 1/25 걸리고 Myers와 Rabiner의 Level-Building DTW 알고리즘^[7-8]의 계산 시간의 1/4이 소요되면서도 다른 알고리즘에 비해 인식률이 떨어지지 않으므로 One-Stage DP 알고리즘을 이용하며, 연속음이나 연결어의 기준 패턴을 생

성하는 DMS model에 의한 DMS/VQ 및 DMS/DP 인식 기법^[1]이 DP 인식 방법보다 인식률이 높으며 기억 장소 및 계산 시간이 적게 소요되므로 본 연구에서는 음성을 인식함에 있어서는 One-stage DP 방법^[4-5]에 DMS/DP 인식 기법^[2-3]을 적용하였으며 구문분석의 개념^[6]을 도입하여 One-Stage DP 방법 보다 인식률이 향상되리라는 개념에서 구문 분석을 이용한 One-stage DMS/DP 방법을 사용하였다.

따라서, 본 논문의 구문 분석 개념은 연속음을 대상으로 할 수 있는데 본 연구에서는 실험 대상으로 25개의 전철명을 택했으며, 8명의 20대 남성 화자의 발음을 이용하였다.

정희원, 호원대학교 컴퓨터학부

Division of Computer, Howon University)

* 본 논문은 호원대학교 교내 연구비에 의한 연구 논문임

접수일자: 2004년1월10일, 수정완료: 2004년4월27일

II. 이 론

1. DMS 모델의 개념 및 작성 방법

가. DMS 모델의 개념

DMS(Dynamic Multi-section) 모델^[1]은 VQ(Vector Quantization)^[9]개념을 확장시킨 MSVQ(Multi-Section VQ)방법^[10-12]이 각 구간을 등간격으로 분할하여 프레임 수를 동일한 길이로 잡아주는 것과는 달리 구간의 경계를 동적으로 잡아 각 구간의 프레임수를 가변 길이로 잡아 주는 모델이다.

즉, DMS 모델은 비슷한 특징을 나타내는 벡터들을 한 구간에 포함시키고 서로 비슷하지 않은 특징을 나타내는 벡터들은 다른 구간에 분할하는 것으로 음성 패턴의 연속된 시계열 상에서 서로 비슷한 특징을 가지는 벡터들을 모아 한 구간으로 만들어 주는 가변 길이로 구간을 구분하는 방법이다.

MSVQ는 구간을 일정하게 나누어줌으로 유사하지 않은 벡터들이 한 구간을 이룰 때 대표 특징 벡터가 서로 유사하지 않은 특징 벡터들 사이에서 구해지는 불합리성이 있게 되는데 이를 개선한 DMS 모델은 유사한 특징을 가지는 벡터들을 한 구간으로 만들기 위해 구간을 동적으로 분할하여 대표 특징 벡터를 구함으로써 짧게 발음되는 특성까지도 대표 특징 벡터로 선택 될 수 있고, 그 구간에 지속 시간 정보를 갖도록 만든 모델이다. 따라서, 단어 패턴의 연속된 대표 특징 벡터를 사용하여 음성인식을 수행함으로써 짧은 음소의 특징 벡터도 동등한 비중을 가질 수 있도록 구간을 동적으로 나누어준다는 점에서 DMS 모델이라 한다.

나. 모델 작성 방법

구간 구분화 알고리즘^[1]의 결과로 등록된 각 학습용 데이터의 단어 패턴 (T1, T2, ..., TN) 들의 구간 경계를 이용하여 각 단어 패턴들을 J 개의 구간으로 나누어 주고, 단어 패턴의 j 번째 구간에 할당된 프레임들의 특징 벡터들을 모아 중심점을 계산해 이것을 단어 모델의 j 번째 구간을 대표하는 특징 벡터로 하였다.

지속 시간 정보는 j 번째 구간의 마지막 프레임수들의 합을 단어 패턴들의 전체 프레임수로 나누어서 구한다. DMS 모델의 작성 과정은 그림 1과 같다.

2. One-Stage DP 방법^[5]

One-Stage DP 알고리즘은 하나의 지수에 의해서 시간 워핑 경로를 파라미터화하여, 최적화 기준을 직접

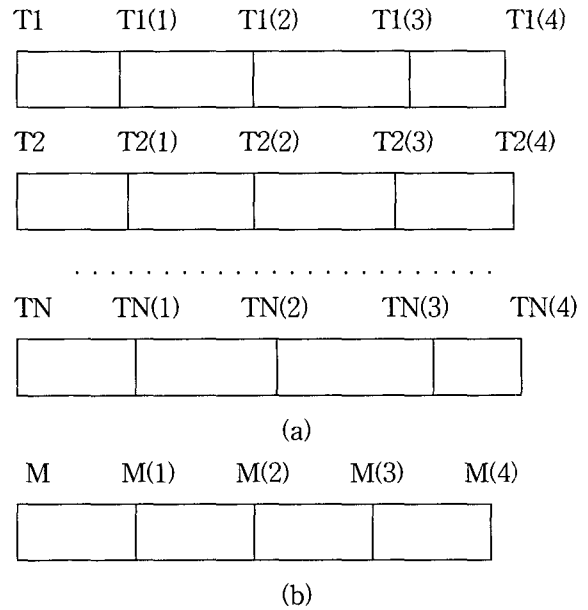


그림 1. 완성된 DMS 모델의 예
(a) 학습용 데이터 (b) DMS 모델
Fig. 1. An example of completed DMS model
(a) training data (b) DMS model

워핑 경로의 함수로 취급한다. 이 방법에서는 두 가지 전이 규칙 (transition rule), 즉, 내부 단어에 대한 전이 규칙과 단어 경계에서의 전이규칙을 사용하여 워핑 경로를 제한한다. 이러한 경로 제약조건과 DP를 사용하여 최적화를 시키면 하나의 stage를 갖는 알고리즘이 되며, 여기에서는 다른 연속음 알고리즘들과서와 같은 여러 단계의 최적화가 필요 없게 된다. One-Stage DP 알고리즘은 Level-Building DTW 방법^[7-9]과 달리 입력 문자열의 최대값을 제한할 필요가 없으며, 더욱이 이 방법은 단독어 인식에 소요되는 정도의 계산량만으로 충분하다.

미지의 입력 또는 시험패턴은 $i = 1, \dots, N$ 시간 프레임들로 구성되어진다고 가정하고, 시간 프레임은 특징 (feature) 들의 벡터에 의해서 표현되어진다. 입력 패턴은 미리 작성된 어휘로부터 선택된 개개의 단어들을 접속하는 것에 의해서 구성되어진다. 어휘의 단어들은 단독적인 발성에서 말해진 하나의 단어 발음(utterance)들로부터 달성되어지는 K개의 기준패턴들 즉, 템플리트 (template)들의 집합으로 구성되어진다. 단어 템플리트들은 지수 $k = 1, \dots, K$ 에 의해서 구별되어지고, 템플리트 k의 시간 프레임들은 $j = 1, \dots, J(k)$ 로써 표시되어진다. 여기서, $J(k)$ 는 템플리트 k의 길이이다.

연속음 인식의 궁극적인 목표는 입력 패턴을 가장 최적으로 매칭(matching)하는 템플리트들의 열 $q(1), \dots,$

q(R)를 결정하는 것이다. 템플리트들 q(1), ..., q(R)의 접속은 "super" 기준패턴^[4-5]으로 언급되어진다. 이 미지의 "super" 기준패턴은 하나의 발성 패턴같이 취급되어질 수 있으므로, 매칭 절차는 단독어 인식의 경우와 동일하게 되어진다. 이 고려사항들에 입각해서, 시간 위평 절차에서 제공되어지는 명시화와 제약조건은 분명하게 되어진다.

시험패턴의 시간 프레임 i와 각각의 템플리트 k의 시간 프레임 j는 격자점 (i, j, k)의 집합으로 정의되어진다. 각각의 격자점 (i, j, k)는 대응하는 음향적인 event들 사이에 비유사성의 척도로써 정의되어지는 소구간 거리척도(local distance measure) d(i, j, k)와 연관되어진다. 연속음 인식의 문제는 시험패턴과 템플리트들의 미지의 열사이의 가장 최적의 매칭 경로를 제공하는 격자점 (i, j, k)의 집합을 발견하는 것이다.

누적 거리 D(i, j, k)는 격자점과 격자점을 반복적으로 전이하면서 측정되어진다. 그래서, 구해진 누적 거리들의 배열 D(i, j, k)를 사용해서 최소 총 누적거리를 가지는 종단 프레임의 템플리트를 구하고 이 격자점에서부터 백트래킹해서 최상의 경로를 구한다.

3. 제안된 음성 인식 방법

가. 개념 도입

본 연구에서는 하나 이상의 음절이 다른 명칭의 음절과 같은 25개의 전철명을 그 인식 대상으로 삼았는데 실험에서 사용되어지는 음성 데이터는 20대 남성 화자 3명이 4번씩 발음한 데이터이다. 그런데, 본 연구에서 발음된 데이터의 내용은 전철명이 두자 또는 세자로 이루어졌다. 따라서, 음소 단위나 음절 단위의 음성 인식 방법을 도입하여 인식하여야 실시간 처리가 가능할 것인데 음소 단위로 인식한다는 것은 오히려 인식해야할 음소의 종류가 많아 음절 단위의 인식 방법에 비해 인식 시간이 많이 걸리고, 음소 단위의 인식이 음절 단위의 인식 방법에 비해 인식률이 현저히 저하되는 관계로 본 연구에서는 음절 단위의 인식 방법을 선택하였다.

따라서, 본 연구에서는 연결 단어 인식에 좋은 알고리즘으로 평가되고 있는 One-Stage DP 알고리즘을 음절단위의 연속음 인식에 적용시켰다. 그러나, 이 One-Stage DP 알고리즘을 연속음에 그대로 적용시키면 상호 조음결합 관계를 고려할 수 없다. 그러므로, 본 연구에서는 연속음을 3부분으로 나누어 조음 효과를 고려하였다. 즉, 첫 음절로 발음되는 것, 중간 음절로 발음되는 것들과 맨 마지막 음절로 발음되는 것으로 나누어

특징을 추출하여 구문 분석 방법을 적용하는 알고리즘을 도입하여 인식한다.

나. 연속음의 구문 분석법

본 연구에서의 구문 분석은 하향식 방법을 응용하여 사용한다. 이는 어떤 한 상태에서 한 개의 입력 심볼에 대해 한 개의 다음 상태를 가지는 결정적 유한 오토마타(Deterministic Finite Automata : DFA)를 사용하여 표시하면 식 (1)과 같다.

$$DFA M = (\{S_f, S_i, S_l\}, \{a, b, c\}, \delta, S_f, \{S_l\}) \quad (1)$$

$$\begin{aligned} \delta : \delta(S_f, a) &= S_i, \\ \delta(S_f, b) &= S_i, \\ \delta(S_i, c) &= S_l, \end{aligned}$$

연속음을 S_f, S_i, S_l과 같이 세 단계로 나눈다. 첫 음절, 중간 음절, 최종 음절로 나누는데 이는 음절 사이에 발생하는 조음 결합 현상을 고려한 것이다. 첫 음절 다음에 따라오는 중간 음절과 조음 결합을 일으키고, 중간 음절은 앞과 뒤 음절 모두와 조음 결합을 일으킨다. 최종 음절은 앞 음절하고만 조음 결합 현상을 일으킨다. 여기서 유의해야 할 사항은 첫 음절 다음에 곧바로 최종 음절이 올 수 있다는 것이다.

다. 제안된 알고리즘

(1) 알고리즘의 선행단계

제안된 구문 분석을 이용한 One-Stage DMS/DP 알고리즘을 상세히 설명하기 전에 One-Stage DMS/DP 알고리즘의 기본적인 개념들을 정리하면 다음과 같다.

- 1) i = 1, ..., N 개의 시간 프레임으로 구성된 미지의 입력 단어열의 시험 패턴의 특징 추출
- 2) 입력 단어열에 대응되는 기준패턴은 단음절로 구성된 k 개의 기준 패턴, 즉, 템플리트의 집합으로 구성되어진다. 즉, 하나의 템플리트마다 DMS 모델링에 의해 각 구간별로 구한 codeword 를 포함하고 있고 지속시간 정보는 기준 패턴을 생성할 때에는 이용하지 않는다. 이후로 부터 기준패턴 또는 템플리트를 DMS 모델링에 의해 구하기 때문에 DMS 템플리트라고 정의한다.
- 3) DMS 템플리트 k의 시간 프레임은 j = 1, ..., J로 표시되고 이때 J는 DMS 템플리트의 길이(구간수)를 의미한다.
- 4) 인식목표는 입력 단어열과 가장 잘 매칭되는 DMS

템플릿들의 접속, 즉, $q(1), \dots, q(R)$ 을 결정하는 것이다. 여기서, $q(1), \dots, q(R)$ 을 "super" 기준 패턴이라고 한다.

(2) 알고리즘

One-Stage DP 알고리즘의 기본적인 개념을 바탕으로 한 구문 분석과 One-Stage DMS/DP를 이용한 연속음 음성 인식 알고리즘은 다음과 같다.

알고리즘을 작성하는데 있어서, 레이블링된 데이터를 살펴본바 가장 짧은 데이터가 20 프레임이었고 가장 긴 데이터가 45 프레임이었으므로 이를 감안하여 실험에서 가장 짧은 데이터가 15 프레임 이하인 것은 없다는 가정하에 가장 짧은 데이터를 15 프레임으로 했고, 가장 긴 데이터가 50 프레임 이상인 데이터는 없다고 가정하여 가장 긴 데이터를 50 프레임으로 해서 알고리즘에 적용하였다. 또한, 첫 음절로 올 수 있는 음절 13개, 중간 음절로 올 수 있는 음절 3개, 최종 음절로 올 수 있는 음절 13개를 고려하여 알고리즘을 이용하였다.

단계 1) 첫 음절 부분의 초기화를 행한다.

- a) $m = 1$ 로 설정한다.
- b) $i = 1$ 로 설정한다.
- c) $k = 1, \dots, K(m)$ 에 대해서, 단계 1d-1e를 행한다.
- d) $j = 1$ 에 대해서, 다음을 행한다.

$$D(i, j, k, m) = d(i, j, k, m) \quad (2)$$

- e) $j = 2, \dots, J$ 에 대해서, 다음을 행한다.

$$D(i, j, k, m) = d(i, j, k, m) + D(i, j-1, k, m) \quad (3)$$

단계 2)

- a) $x = 1$, 및 $y = 1$ 을 설정한다
- b) $i = 2, \dots, 15$ 에 대해서, 단계 2c를 행한다.
- c) $K = 1, \dots, K(m)$ 에 대해서, 단계 2d-2e를 행한다.
- d) $j = 1$ 일 때,

$$D(i, j, k, m) = d(i, j, k, m) + \min \left\{ \begin{array}{l} D(i-1, j, k, m^*) \\ D(i-1, j, k^*, m^*) \end{array} \right. \quad (4)$$

$m^* = x, \dots, y$ 반복

$k^* = 1, \dots, K(m)$ 반복

- e) $j = 2, \dots, J$ 에 대해서, 다음을 행한다.

$$D(i, j, k, m) = d(i, j, k, m) + \min \left\{ \begin{array}{l} D(i-1, j, k, m) \\ D(i-1, j-1, k, m) \\ D(i, j-1, k, m) \end{array} \right. \quad (5)$$

- f) 2 음절 단어인 경우 단계 6으로 간다.

단계 3) 중간 음절 부분을 초기화 한다.

- a) $m = 2$ 로 설정한다.
- b) $i = 15$ 로 설정한다.
- c) $k = 1, \dots, K(m)$ 에 대해서, 단계 3d-3e를 행한다.
- d) $j = 1$ 에 대해서, 다음을 행한다.

$$D(i, j, k, m) = d(i, j, k, m) + \min \{ D(i, j-1, k^*, m-1) \} \quad (6)$$

$k^* = 1, \dots, K(m-1)$ 반복

- e) 단계 1e를 행한다.

단계 4)

- a) $i = 16, \dots, 50$ 에 대해서, 단계 4b-4c를 행한다.
- b) $m = 1, x = 1$ 및 $y = 1$ 로 설정하고 단계 2c를 행한다.
- c) $m = 2, x = 1$ 및 $y = 2$ 로 설정하고 단계 2c를 행한다.

단계 5)

- a) $m = 2, x = 2$ 및 $y = 2$ 를 설정한다
- b) $i = 51, \dots, N-50$ 에 대해서, 단계 2c를 행한다.

단계 6) 최종 음절 부분을 초기화한다.

- a) $m = 3$ 로 설정한다.
- b) $i = N-50$ 으로 설정하고 단계 3c를 행한다.

단계 7)

- a) $i = N-50, \dots, N-15$ 에 대해서, 단계 7b-7c를 행한다.
- b) $m = 2, x = 2$ 및 $y = 2$ 로 설정하고 단계 2c를 행한다.
- c) $m = 3, x = 2$ 및 $y = 3$ 로 설정하고 단계 2c를 행한다.

단계 8)

- a) $m = 3, x = 3$ 및 $y = 3$ 를 설정한다
- b) $i = N-15, \dots, N$ 에 대해서, 단계 2c를 행한다.

단계 9)

누적 거리들의 배열 $D(i, j, k, m)$ 를 사용해서 최소 총 누적 거리값을 가지는 DMS 템플리트의 중단 프레임이 있는 격자점으로부터 가장 최적의 경로를 백트래킹한다.

제안된 방법에 의한 알고리즘의 최종적인 목표는 입력 단어들의 미지의 열을 결정하는 것이고 이것은 위에서 설명된 알고리즘의 단계 9에서 백트래킹을 수행하는 것에 의해서 결정되어진다.

여기서, $m(m = 1, \dots, 3)$ 은 음절의 위치를 나타내며 1인 경우가 첫 음절을 의미하고 2인 경우가 중간 음절들, 3인 경우가 최종 음절을 의미한다. $i(i = 1, \dots, N)$ 는 입력 패턴의 특징벡터를 의미하고, $j(j = 1, \dots, J)$ 는 각 DMS 템플리트의 구간을 의미하며, $k(k = 1, \dots, K[m])$ 은 각 부분(첫 음절, 중간 음절 및 최종 음절 부분)의 거편 연속음 음절을 의미한다. 여기서, $K^{[1]}$ 과 $K^{[3]}$ 는 첫 음절과 최종 음절로 13개이고, $K^{[2]}$ 는 중간 음절이 적은 관계로 3개이다.

시간 프레임 i 에 대해서 동적 프로그래밍을 재귀적으로 수행하기 위해서, 누적 거리의 배열 $D(i, j, k, m)$ 의 집합이 필요하다. 저장 장소의 이 열은 누적 거리들의 배열으로써 간단하게 언급되었고, $D(i, k, m)$ 로서 표시되어진다. 이와 같이, 배열 $D(i, j, k, m)$ 는 단지 하나의 결만을 사용해서 동적 프로그래밍의 재귀 관계식이 시점 패턴의 시간축을 따라서 수행되어지고 격자점과 격자점을 따라서 누적 거리의 열 배열을 갱신한다.

시험 패턴과 DMS 템플리트 사이에서 최소 총 누적 거리를 구하는데 있어 연속음에서는 백트래킹이 입력 음성 단어들의 미지의 열을 발견하기 위해서 One Stage DMS/DP 알고리즘에 적용되어진다. 백트래킹 정보는 동적 프로그래밍의 재귀식이 평가되어지는 동안에 저장되어지고, 격자점 (i, j, k, m) 를 경유하는 경로상에서 DMS 템플리트 k 에 있는 j 프레임은 단일한 시작점에서 시작되어진다.

그래서, 각각의 격자점에 대해서 백포인터는 격자점 (i, j, k, m) 에 경로를 가지는 선행하는 단어의 중단 프레임(ending frame)의 값을 저장한다.

앞에서 기술했듯이, 본 논문에서는 배열 $D(i, j, k, m)$ 의 경우에서처럼 백포인터 배열을 저장할 수 있고, 가장 최적의 경로의 격자점과 격자점에서 백포인터의 값을 갱신한다. 그림 2은 제안된 구문분석과 One-Stage DMS/DP를 이용한 전체 인식 시스템을 나타낸다.

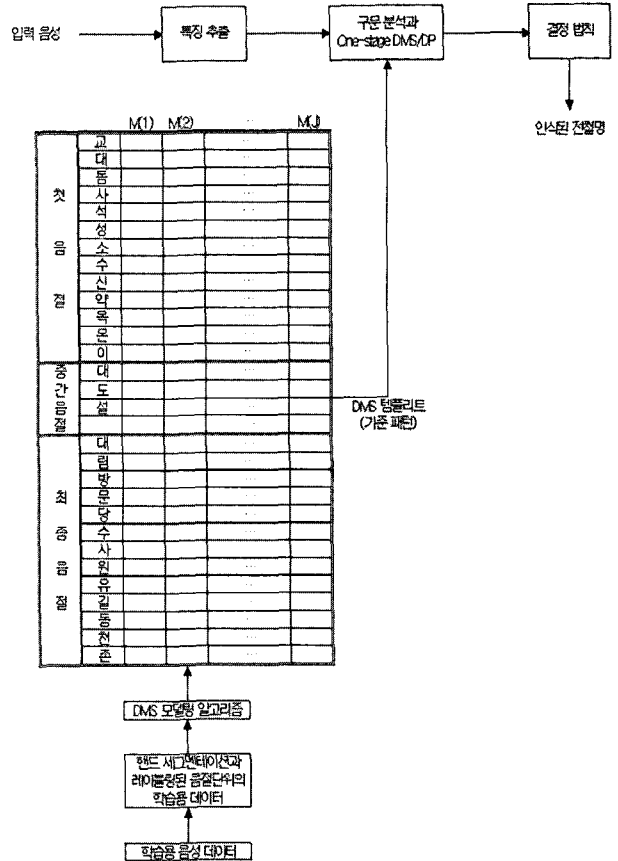


그림 2. 구문 분석과 One-Stage DP를 이용한 전체 인식 시스템의 블록도

Fig. 2. Block Diagram of overall recognition system using One-Stage DP and syntactic analysis

표 1. 연속음 데이터 목록

Table 1. Data list of the continuous speech

순번	전철명	순번	전철명	순번	전철명	순번	전철명	순번	전철명
1	교대	6	석수	11	신길	16	신사	21	옥수
2	대림	7	성수	12	신당	17	신철동	22	온수
3	대방	8	소사	13	신대당	18	신천	23	이대
4	동대문	9	수원	14	신도림	19	신촌	24	이수
5	사당	10	수유	15	신림	20	약수	25	이촌

III. 실험 결과 및 고찰

1. 실험 결과

본 논문에서 사용되어지는 대상 어휘는 전철역명 상호간에 연속해서 일어날 수 있는 모든 경우의 수를 조합해서 구성한 Balanced list이다. 데이터 목록은 표 1과 같이 25개의 대상 어휘이다.

실험에 사용되어지는 음성 데이터는 25개의 전철역

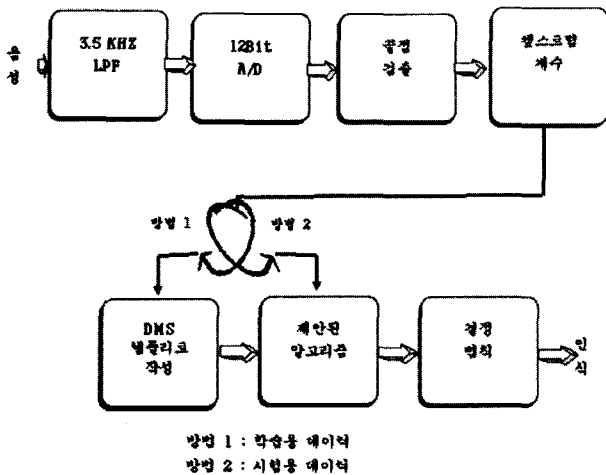


그림 3. 인식 시스템의 구성도
Fig. 3. The block diagram of recognition system

표 2. 구간수 변화에 따른 실험 결과
Table 2. Recognition result according to the section variation

구간수 \ 방법	6	8	10	12	14	16
기존의 방법	86	89.5	90.5	91.5	91.5	91
제안된 방법	91	95	97.5	98.5	98	98

표 3. 구간수 변화에 따른 실험 결과
Table 3. Recognition result according to the section variation

구간수 \ 방법	6	8	10	12	14	16
기존의 방법	82	85.5	87.25	88.25	88	88
제안된 방법	89	93.25	94.5	95.5	95	94.75

명을 20대 남성 화자 8인에 의해서 4번씩 발성한 음성 데이터를 이용해서 수행하였다. 4명이 2번씩 발성한 음성 데이터를 DMS 템플릿과 화자 종속의 실험을 수행하기 위해서 사용하였고 4명의 화자가 4번 발성한 음성 데이터는 화자 독립의 실험을 수행하기 위해서 사용되어졌다. 그리고, 전체적으로 음성 데이터들은 보통의 발성 속도에서 녹음되어 저장되어졌다.

가. 인식 시스템의 구성

인식 실험을 위한 음성 인식 시스템은 그림 3와 같이 구성하였다.

나. 인식 실험

8명의 화자 중 4명의 화자가 발성한 2회 발음으로 모

델을 만들고 나머지 발음으로 실험을 행한 화자 종속의 실험 결과가 표 2에 나타내어졌다. 비교를 위하여 기존의 One-stage DP 방법에 의한 실험도 같이 행하였다. 실험 결과 12구간으로 나누었을 때 가장 좋은 98.5%의 인식률을 보였다.

또한, 나머지 4명의 화자의 발음으로 실험을 행한 화자 독립의 실험 결과가 표 3에 나타내어졌다. 실험 결과 화자 종속 실험과 마찬가지로 12구간으로 나누었을 때 가장 좋은 95.5%의 인식률을 보였다.

2. 고찰

위의 실험 결과에서 보여지듯 본 실험은 두가지로 나누어서 비교 실험하고 있다. 기존의 방법과 제안된 방법 둘 다에서 공히 DMS의 구간수를 증가하면 할수록 인식률이 증가하다가 떨어지는 경향이 있다. 기존의 방법에서나 제안된 방법에서나 화자 종속 및 화자 독립의 경우 모두 구간수 12에서 인식률이 가장 좋았다.

실험 데이터에서, 오인식되는 전철명들은 '신촌'이 '신천'으로 '옥수'가 '온수'로, '석수'가 '성수' 등으로 오인식되고 있다.

각각의 인식율은 기존의 방법인 경우 화자 종속인 데이터는 91.5%, 그리고 화자 독립인 데이터는 88.25%의 인식률을 보이고 있다. 기존의 방법에 비해 구문 분석 기법을 도입한 One-Stage DMS/DP 방법에서는 화자 종속인 데이터는 98.5%의 인식률을 나타내고 화자 독립인 데이터는 95.5%를 나타낸다. 기존의 방법과 제안된 방법과의 차이는 약 7% 정도로 제안된 방법이 인식률의 증가를 가져왔다.

IV. 결 론

본 논문은 구문 분석을 이용한 One-Stage DMS/DP에 의한 연속음 인식에 관한 연구로써, 기존의 One-Stage DP 알고리즘에서 문제가 제기된 바 있는 상호조음결합 현상을 적절히 처리하기 위해 구문 분석의 개념을 도입하였다.

실험한 데이터로는 20대 남성 화자 8명이 4회 발성한 25개의 전철명을 사용하였다. 이 중 4명의 화자가 2번 발음한 데이터로 모델을 만들고, 나머지 2번 발음한 데이터로는 화자 종속의 실험을 하였고, 나머지 4명의 화자가 4번 발음한 데이터로는 화자 독립의 실험을 행하였는데 기존의 방법에서는 각각 인식률이 91.5%,

88.25%로 나타났으며, 제안된 방법에서는 각각 인식률이 98.5%, 95.5%로 나타나 제안된 시스템의 우수성이 입증되었다.

연속음 인식이 현 실생활에 적용되면 ARS(Audio Response System), 장애인 시스템, 음성 다이얼링, 승강기 운용 등 응용할 수 있는 분야가 무수히 많다. 그리고 인간의 가장 간편한 통신 수단인 음성을 이용하면 편리한 점이 많다. 그러나, 아직도 실생활에 적용하기 위해서는 더 높은 인식율을 요구한다.

상호 조음결합 현상으로 인한 오인식율을 줄이고, 인식률 향상을 목적으로 한 이번 실험은 실험 결과 약 7%의 인식률 향상이 있었다. 그러나, 앞으로 더 인식율을 높이는 데는 상호 조음 효과를 완화시킬 수 있는 더 새로운 알고리즘의 연구와 기준 패턴을 정하는 방법의 연구가 우선되어야 할 것이다.

참고 문헌

- [1] Tae Ock Ann and Young Kyu Byun, "A Study on Speech Recognition using DMS Model," Acoustical Society of Korea, Vol. 13, No 2E, pp.41-50, 1994. 12.
- [2] 이항섭, "실시간 음성 다이얼링 시스템 구현을 위한 단독어 인식에 관한 연구", 광운 대학교 대학원, 1991. 2.
- [3] 안태욱 외, "실시간 음성 다이얼링 시스템 구현을 위한 연결어 인식에 관한 연구", 한국 음향학회지, 제12권 제3호, pp.13-25, 1993. 6.
- [4] Hiroaki Sakoe and Seibi Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition," IEEE. Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-26, No.1, pp. 43-49, Feb. 1978.
- [5] H. Ney, "The Use of a One-Stage Dynamic Programming Algorithm for Connected Word Recognition," IEEE Trans. on ASSP, Vol. ASSP-32, No.2, pp.263-271, April 1984
- [6] 안태욱, 변용규, 김순협, "구문 분석과 Level Building을 이용한 한국어 연속음 인식", 한국 음향학회지, 제5권 제4호, pp.27-36, 1986. 12.
- [7] C. S. Myers, R. R. Rabiner, "Connected Digit Recognition Using a Level-Building DTW Algorithm", IEEE Trans. Acoust. Speech, signal processing, Vol. ASSP-29, pp 351-363, June 1981.
- [8] C. S. Myers and L. R. Rabiner, "A Level Building Dynamic Time Warping Algorithm for Connected Word Recognition," IEEE Trans. on ASSP, Vol. ASSP-29, No.2, pp.284-296, April 1981
- [9] Y. Linde, A. Buzo, and R.M. Gray "An algorithm of Vector Quantizer Design", IEEE Trans. Commun, Vol. COM-28 pp.84-95, Jan 1980
- [10] D.K. Burton, J.E. Shore, J.T. Buck "Isolated-Word Speech Recognition Using Multisection Vector Quantization Codebooks" IEEE Trans. of Acoustics, Speech, Signal Processing Vol ASSP-33. No 4. August 1985
- [11] 이성권 외, "시간 정보와 VQ를 이용한 DDD 지역명 인식에 관한 연구", 한국 음향학회지, Vol.8, No.5, pp.102-112, 1989. 10.
- [12] 안태욱, 김순협, "TSVQ를 이용한 Computer 자동 음성 인식", 대한 전자 공학회지, 제 27권, 제 7호, pp.157-165, 1990. 7.

저자 소개



안 태 옥(정회원)
제 40권 SP편 제 6호 참조

