

선형예측계수에 기초한 퍼지추론 단어인식

Word Recognition using Fuzzy Inference based on LPC

최 승 호*, 김 형 근**
(S. H. Choi*, H. K. Kim**)

요 약

LPC열로 구성된 음성패턴의 주파수변동을 해결하기위해 LPC와 스펙트럼, LPC차수와 스펙트럼의 관계를 고찰하여 새로운 형태의 멤버쉽함수를 제안하였다.

또한, 시간변동을 해결하기위해서는 음성구간을 여러구간으로 등간격분할하는 다구간 등분할법을 사용하였으며, 이때 오인식은 주로 동일음절이 같은 발성위치에 있을때 발생되었다. 이러한 오인식을 줄이기위해 제안된 멤버쉽함수로 퍼지추론한 뒤 구간별 확신도에 가중치를 부여하고, 세번째후보까지를 인식대상으로 하는 판정알고리즘을 제안하였다.

본 방법의 타당성을 검증하기위해, DDD지역명 28개를 대상으로 인식실험한결과, 삼각형멤버쉽함수에 의한 퍼지추론은 92.0%, 삼각형멤버쉽함수에의한 퍼지추론과 판정알고리즘은 92.9%, 제안된 멤버쉽함수에의한 퍼지추론과 판정알고리즘은 93.8%의 인식률을 보였다.

ABSTRACT

To solve the frequency variation of speech patterns which consist of LPC sequences, new membership function view from LPC, spectrum and the relations between the order of LPC and spectrum is proposed.

To solve the time variation, multi-sectation equi-segmentation method which equally divide the speech section into several section are applied. False recognition mainly occur at time when the same syllable is placed at the same utterance. To reduce the error, fuzzy inference is executed using the proposed membership function and weights are assigned into sectional certainty and then the decision method for recognized the section up to the third candidate.

To testify the validation of this method, we experimented the recognition test of 28 DDD area names. The recognition rate of the fuzzy inference by the triangle membership function is 92%. That of the combined method of the fuzzy inference and the decision method is 92.9% and that of fuzzy inference by the proposed membership function is 93.8%.

I. 서 론

음성인식의 분석방법 중 선형예측분석법이 가장 널리

사용되고 있다[1-2]. 이 분석법이 사용된 초기에는 알고리즘의 복잡성 때문에 큰 실효를 거두지 못했으나, 보다 간단한 거라적도인 log likelihood ratio 거리를 F. Itakura가 음성인식에 적용한 후, 음성인식에 선형예측계수를 이용하려는 연구가 활발히 전

* 동신대학교 정보통신공학과
** 한국방송통신대 전자계산학과
접수일자: 1993년 8월 6일

개되고 있다[3-5].

또한, 선형예측모델이 all-pole 모델이므로 고주파 성분을 갖고 있는 음성에 대해서는 계수가 부적절하고 계수와 스펙트럼간에 일대일로 대응하지 않는 등의 단점이 있으나, 계수들을 다른 특징량으로 변형시켜 이들을 인식파라미터로 사용하여 좋은성과를 얻고 있다[6].

음성정보를 패턴인식하기 위해서는 음성의 유향학적 특성을 고려하여 패턴의 변동요인을 제거할 필요가 있다. 이러한 패턴의 변동요인은 시간변동과 주파수 변동으로 구분되며, 시간변동은 화자마다 달라지는 발생속도와 동일 화자라 할지라도 같은 단어를 재발성할때 발생 지속시간의 차이에 의해 발생하고, 주파수변동은 발생상황과 성도길이가 화자마다 달라짐에 따라 발생한다.

본 연구에서는 주파수변동을 해결하기 위해서 퍼지이론을 도입하였으며, 특히 멤버쉽함수의 변동폭을 결정하기위해 LPC계수와 스펙트럼, LPC차수와 스펙트럼의 관계를 고찰하여 전력의 변화율인 PDR(power difference rates)을 정의하였으며, 이를 근거로 새로운 형태의 변동폭을 갖는 멤버쉽함수를 제안하였다. 또한, 시간변동은 음성구간을 8구간으로 일정하게 나누는 다구간 등분할법을 사용하였다[13]. 이때 다구간 등분할법은 같은 발생위치에 동일 음절이 있을때 오인식이 주로 발생하기 때문에 이를 개선하기 위해 인식과정에 구간별 확신도를 세번째 후보까지 정한뒤 이 팩턴들을 다시 분리하여 인식단어를 판정하는 판정방법을 제안하여 삼각형 멤버쉽함수와 제안된 멤버쉽함수의 오인식율을 고찰하였다.

II. 선형예측계수와 멤버쉽함수

2-1. 선형예측계수와 주파수 스펙트럼의 관계[8-10]

개인간 음성신호의 차이는 발생기관의 해부학적 구조와 발생기관을 움직이는 습관의 차이에 기인된다. 음성신호에 포함된 음향학적인 정보를 추출하는 일반적인 방법은 음성의 발생과정을 하나의 동적시스템으로 보고 관측된 입출력데이터를 기초로 모델링된 시계열 모델이다.

시계열 모델에는 AR(auto regressive)모델, MA(moving average)모델, ARMA(auto regressive moving average)모델 등이 있으며, AR모델은 현재의

출력을 과거의 출력과 현재의 입력으로, MA모델은 현재의 출력을 현재와 과거 입력들의 선형결합으로 나타낸다. 그리고, ARMA모델은 현재 출력을 과거 출력과 현재와 과거 입력들의 선형결합으로 나타낸다.

이 중에서, AR모델은 음성신호를 분석하는데 주로 사용되며, 이 분석법은 다음과 같다. 현재의 시계열 음성신호 $x(n)$ 을 선형시스템의 출력이라고 가정하면 선형예측계수를 추정하기 위한 관계는 식(1)과 같다.

$$\hat{X}(n) = - \sum_{k=0}^{K-1} A_k x(n-k) \quad (1)$$

여기에서 K 는 예측차수, A_k 는 선형예측계수이고, $\hat{x}(n)$ 은 선형예측신호이다.

이때 파워스펙트럼을 데시벨[dB]단위로 나타내면 식(2)와 같다.

$$L(f) = 10 \log_{10} |P_x(f)| \text{ [dB]} \quad (2)$$

여기에서 $P_x(f)$ 는 $\sigma^2 \Delta t H[\exp(j2\pi f \Delta t)]$ 이고, σ^2 는 모델입력의 분산이다.

이로부터 계수와 스펙트럼의 관계를 포락선의 형태로 부터 고찰한다.

그림 1은 단어 "서울"중 "어"의 일부분에 해당되는 23번째 프레임에 대한 포락선이다. 이 포락선에서는 제 1 포먼트주파수 495[Hz]에서 나타나기 때문에 모음의 주성분이 저주파임을 알 수 있다. 이러한 사실은 계수로부터 정확한 스펙트럼이 추정됨을 알 수 있다.

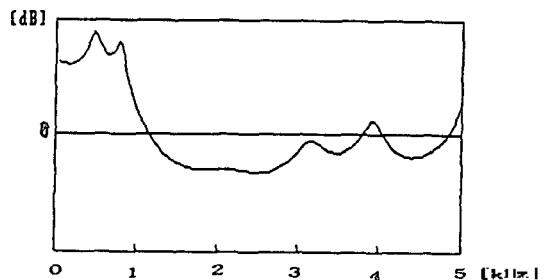


그림 1. "서울"의 23번째 프레임에 대한 스펙트럼 포락선.
Fig 1. Spectrum envelope for the 23th frame of "Seoul".

그림 2는 포락선과 차수와의 관계를 고찰하기위해 계수값을 차수별로 변화시킨 포락선들이며, 이때 사용된 서울의 23번째 프레임의 차수별 계수값은 표 1과 같다.

그림 2의 (a)는 A_1 를 영으로, 나머지 계수값은 원래의 값으로 하여 스펙트럼 포락선을 구한 것으로 높은 주파수보다는 낮은 주파수에서 포락선이 크게 변한것을 알 수 있다. 이로부터 A_1 는 높은 주파수보다는 낮은 주파수에 영향을 주고 있음을 알 수 있다. (b)는 A_2 를 영으로, 나머지 계수값은 원래의 값으로 하여 스펙트럼 포락선을 구한 것이다. 이 경우도 (a)에서 마찬가지로 높은 주파수에서 보다는 낮은 주파수에서 스펙트럼의 모양의 크게 변화였으며 A_2 도 A_1 과 마찬가지로 낮은 주파수에 많은 영향을 주고 있음을 알 수 있다. A_3 만 영으로하여 스펙트럼포락선을 구한 (c)는 (a), (b)와 같이 높은 주파수보다는 낮은 주파수에서 크게 변화였지만 (a), (b)만큼 변화를 주지 못하였다. 이것은 A_3 가 A_1 , A_2 보다 낮은 주파수에 영향을 작게 줌을 나타내는 것이다. A_4 만을 영으로하여 구한 스펙트럼포락선인 (d)는 (c)에 비하여 스펙트럼포락선의 변화가 더 심하다. 이는 A_4 가 A_3 의 값보다 2배 이상이 크기 때문에 나타난 현상으로 보아진다. 따라서 스펙트럼포락선은 선형예측계수의 차수보다는 크기에 영향을 더 받는 것으로 사료된다. (e)는 A_5 만 영으로 하여 구한 스펙트럼포락선이다. (d) 보다 원 스펙트럼포락선에 비슷한 스펙트럼포락선을 보이고 있다. 이는 A_5 의 값이 A_1 , A_2 , A_3 , A_4 의 값보다 작은 원인으로 사료된다. (f)는 A_6 을 영으로 한 경우로서 (a), (b), (d)보다는 원 스펙트럼포락선에 유사하고, (c), (d), (e)보다는 낮은 주파수에서 덜 변화하였다. 이러한 결과는 선형예측계수가 스펙트럼포락선에 영향을 주는 원인이 선형예측계수 차수보다 계수의 크기임을 알 수 있다. (g), (h), (i), (j), (k), (l), (m), (n)로 갈수록 원래의 스펙트럼포락선에 가까워지고 있다. 이는 선형예측계수의 차수가 증가할수록 스펙트럼포락선에 주는 영향은 적어짐을 알 수 있었다. 이는 차수가 증가할수록 선형예측계수의 값이 작아지기 때문으로 사료된다. 이상과 같은 고찰로부터 다음과 같은 사실을 알 수 있다. 선형예측계수의 변화는 대역내의 높은 주파수보다는 낮은 주파수에 보다 많은 영향을 준다. 이는 선형예측계수가 AR모델로부터 얻어진 계수이고, AR모델

은 높은 주파수보다 낮은 주파수에 보다 더 많은 영향을 주는 all pole 모델이기 때문이다.

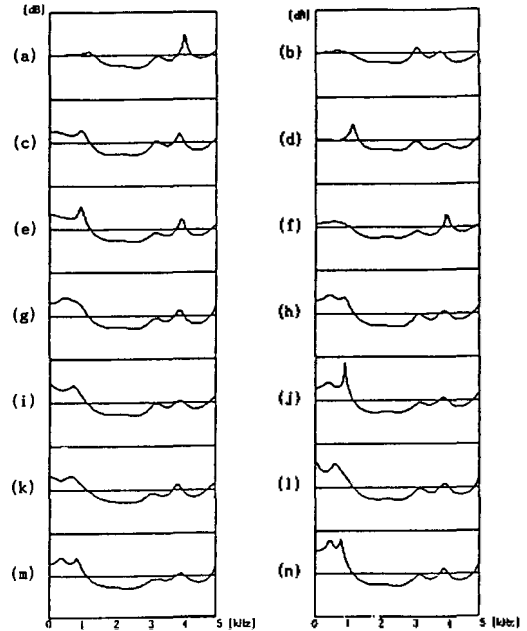


그림 2. 차수별 계수값을 영으로 대신한 경우의 스펙트럼 포락선.

Fig 2. Spectrum envelope in case of replacement coefficient of each order with "zero".

- (a) $A_1 \rightarrow 0$ (b) $A_2 \rightarrow 0$ (c) $A_3 \rightarrow 0$ (d) $A_4 \rightarrow 0$
- (e) $A_5 \rightarrow 0$ (f) $A_6 \rightarrow 0$ (g) $A_7 \rightarrow 0$ (h) $A_8 \rightarrow 0$
- (i) $A_9 \rightarrow 0$ (j) $A_{10} \rightarrow 0$ (k) $A_{11} \rightarrow 0$ (l) $A_{12} \rightarrow 0$
- (m) $A_{13} \rightarrow 0$ (n) $A_{14} \rightarrow 0$.

표 1. "서울"의 23번째 프레임에 대한 선형예측계수.
Table 1. The LPC for the 23th frame of "Seoul".

차 수	계 수 값
A_0	1.00000
A_1	-0.98094
A_2	-0.92206
A_3	0.42604
A_4	0.97262
A_5	0.30376
A_6	-0.62164
A_7	-0.17825
A_8	-0.18155
A_9	0.24053

A_{10}	- 0.20902
A_{11}	0.32190
A_{12}	0.150554
A_{13}	- 0.22693
A_{14}	- 0.01223

음성인식에 있어서 스펙트럼의 사용은 포맷트주파수를 알아보기 위함이 대부분이다. (i), (j), (m), (n)의 경우 선형예측계수가 다름에도 불구하고 스펙트럼포락선의 포맷트주파수는 동일하다. 이는 음성인식을 포맷트주파수만으로 인식할 때 오인식의 원인이 될 수 있음을 의미한다. 몇 개의 포맷트 주파수의 비교보다는 주파수 전체를 비교함으로써 이러한 문제를 해결할 수 있다. 왜냐하면, 선형예측계수는 주파수와 일대일 대응하는 것이 아니라 각 계수의 크기에 따라 정도의 차이는 있지만 모든 주파수 성분을 갖고 있기 때문이다.

2-2. 전력 변화율과 멤버쉽함수

1) 전력 변화율(PDR)

계수값의 변화가 포락선에 어떤 영향을 주고 있는지를 고찰하기 위하여 포락선의 전력 변화율(PDR : Power Difference Rate)을 식(3)와 같이 정의하여 사용하였다.

$$PDR = \frac{\text{원래의 power 값} - \text{변화시킨 power 값}}{\text{원래의 power 값}} \times 100[\%] \quad (3)$$

예를들면, 그림 3은 "인천"중 "어"의 일부분인 42 프레임에서 첫번째 계수만을 가변하고 나머지 계수들은 변화시키지않고 변화량은 첫번째계수 A_1 에 대한 백분율로 환산하여, A_1 의 값을 그대로 한 경우 100[%], A_1 의 값을 영으로 한 경우를 0[%]로 나타내었다.

그림3의 (a)는 원 파형의 첫번째 계수값을 100[%]로 했을 때의 원 포락선이며, (b)는 첫번째 계수의 0[%]로 PDR은 129[%]이고, 제 1, 2포맷트의 식별이 불가능해져 포락선의 형태가 달라진다. (c)는 첫번째 계수의 90[%]로서 포맷트는 변하지 않지만 그 크기는 크게 변하고, PDR은 84[%]이다. (d)는 계수의 95[%]로 포락선에는 영향을 주지만 포맷트와 크기는 거의 같고, PDR은 12[%]이다. (e)에서 (i)까지

는 첫번째 계수의 97[%], 99[%], 99.5[%], 100.5[%], 101[%]로 구한 포락선들로 원 포락선과 거의 같은 모양이며, 이 때의 PDR은 4.4[%], 0.54[%], 0.14[%], 0.15[%], 0.65[%]이다.

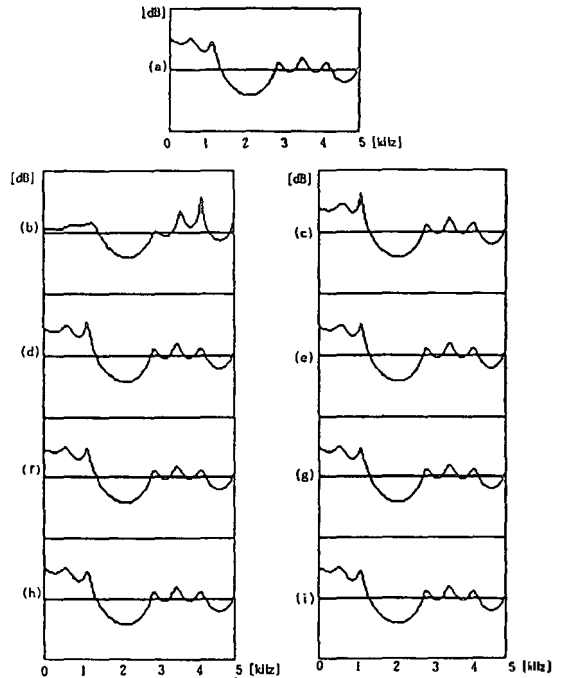


그림 3. 계수의 변동에 따른 스펙트럼 포락선의 예.
 (a) 원파형 (b) 0[%] (c) 90[%]
 (d) 95[%] (e) 97[%] (f) 99[%]
 (g) 99.5[%] (h) 100.5[%] (i) 101[%]

Fig 3. An example of the spectrum envelope according to the variation of coefficient.

그림 4는 6대 도시명에서 무작위로 50 프레임을 선택하여 계수값의 변화에 따른 전력 변화율의 평균값을 나타낸 것이다. 이를 고찰해보면 PDR이 1[%]미만인 경우는 계수의 변동율이 상위 0.5[%]와 하위 1[%] 이내인것을 볼 수 있으며, 또한 이들의 포락선의 형태가 원파형의 포락선과 거의 유사함을 알 수 있었다. 이러한 사실로부터 각 차수의 계수들은 포락선과 직접적으로 일대일 대응하지는 않지만 포락선의 형태에 전체적으로 영향을 주기 때문에, 한 프레임을 구성하고 있는 계수들은 차수에 무관하게 하

나 하나가 포괄적으로 모든 주파수 성분을 갖고있는 것으로 사료된다.

따라서 개인차에 따른 주파수변동의 흡수는 계수의 변동을 흡수하는 것과 같기 때문에, 계수 자체에 퍼지개념을 도입하고, 전력 변화율의 결과를 적용한 새로운 형태의 멤버쉽함수를 제안하여 퍼지추론에 의해 주파수변동을 해결하고자 한다.

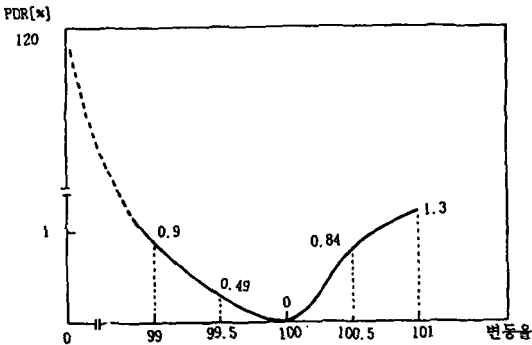


그림 4. 계수의 변동율에 따른 전력 변화율의 평균값
Fig 4. Mean of the PDR according to the variation of coefficient.

2) 멤버쉽함수[11-13]

주파수변동을 흡수하기 위해서 표준패턴의 평균값을 중심값으로, 분산값을 변동폭으로 하는 삼각형 멤버쉽함수와 전력 변화율로부터 작성된 변형 삼각형 멤버쉽함수를 제안하여 사용하였다.

삼각형 멤버쉽함수 $\mu_T(X)$ 와 변형된 삼각형 멤버쉽함수 $\mu_M(X)$ 를 식(4), 식(5)에 각각 나타내며, 여기에는 M은 표준패턴의 평균값, D는 분산값이며, α 는 평균값의 상위 0.5[%] 값이고, β 는 평균값의 하위 1[%]값을 나타낸다.

$$\mu_T(X) = \begin{cases} (X-M)/4D+1, & M-4D < X < M \\ 1, & X=M \\ (X-M)/4D+1, & M < X < X+4D \\ 0, & X \leq M-4D, X \geq M+4D \end{cases} \quad (4)$$

$$\mu_M(X) = \begin{cases} (X-M)/4D+1, & M-4D < X < M-\beta \\ 1, & X-\beta \leq X \leq M+\alpha \\ (X-M)/4D+1, & M+\alpha < X < X+4D \\ 0, & X \leq M-4D, X \geq M+4D \end{cases} \quad (5)$$

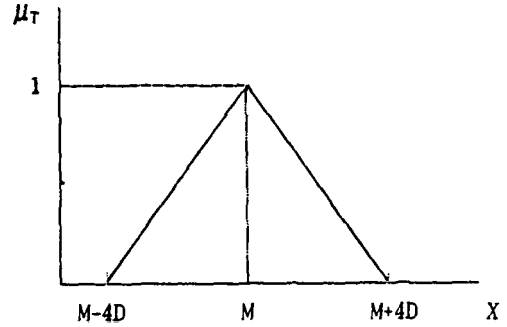


그림 5. 삼각형 멤버쉽함수(M=평균값, D=분산값).
Fig 5. Triangle membership function (M=mean, D=variance).

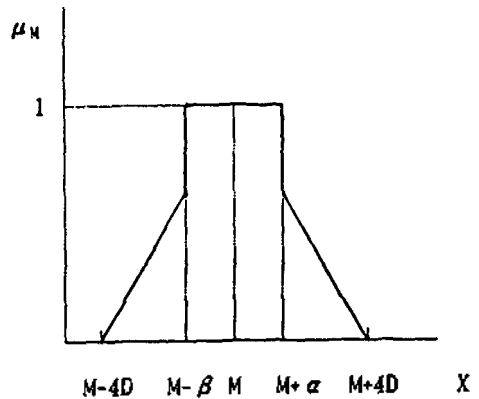


그림 6. 변형된 삼각형 멤버쉽함수.
Fig 6. Modified triangle membership function.

Ⅲ. 퍼지추론과 판정방법

본 연구에서는 시간변동과 주파수변동의 해소를 위해 다구간 등분합법과 변형된 멤버쉽함수를 이용한 퍼지추론 방법을 각각 사용하고, 인식단어를 결정하는 과정에 오인식을 개선하기 위해 구간별로 세 번째 후보까지 가중치를 부여한 뒤 다시 인식 후보를 선정하여 인식 단어를 판정하는 방법을 제안하였다.

3-1. 다구간 등분합법과 퍼지추론[13-15]

시간변동을 흡수하기 위한 일반적인 방법은 DTW를 사용하는 것이지만 간단히 하기 위해서는 패턴의 비선형적인 관계를 선형적으로 팽창하는 과정에서 정보손실이 발생한다. 이러한 정보손실을 방지하기

위해서 계수를 음성의 음향학적 특성이 유지되도록 발생순서에 따라 등간격으로 다구간 분할하는 다구간 등분할법을 사용한다.

이 방법으로 음성패턴을 4개와 8개 그리고 16개의 각각 등간격으로 분할하며, 첫째, 각 구간의 차수별 평균값을 구하여 표준패턴을 작성하고, 둘째, 멤버쉽함수의 중심값과 변동폭은 표준패턴의 평균값과 분산값으로 각각 대응하도록 구성한다. 이때 멤버쉽함수는 변형된 멤버쉽함수를 사용한다.

따라서, 행렬로 나타낸 멤버쉽함수를 이용한 다구간 등분할된 표준패턴과 미지패턴간의 퍼지추론의 과정은 그림 7과 같다.

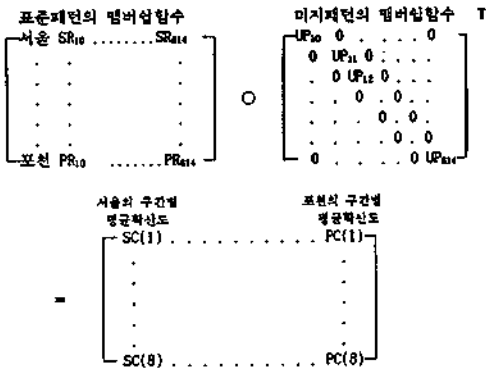


그림 7. 다구간으로 등분할된 표준패턴과 미지패턴간의 퍼지추론.

Fig 7. Fuzzy inference between reference pattern and unknown pattern in the multi-section equi-seg. emtation.

3-2. 판정방법

다구간 등분할된 표준패턴과 미지패턴들은 구간별로 확신도를 구하기 때문에 같은 발생 위치에 동일한 음절이 있을 때 오인식이 발생하게 된다. 이러한 오인식을 개선하기위해, 판정방법은 4단계의 과정을 거친다.

여기에서는 구간수가 8개인 경우에 한정하여 기술하였고, 각 변수명에서 C(i)는 구간별 확신도, W는 단어명 그리고 i는 구간수를 나타낸다.

단계 1: 구간별 가중치 V(i)

구간별 확신도인 C(i)의 최대값을 세 번째

후보까지는 1로, 그 이외는 0으로 부여하며, 그 값들은 V(i)에 저장한다.

$$\begin{aligned}
 V(i) &= \max_1 [C(i)w_1] = 1 \\
 &= \max_2 [C(i)w_2] = 1 \\
 &= \max_3 [C(i)w_3] = 1 \\
 &= 0 \text{ (otherwise)}
 \end{aligned}$$

여기에서, $\max_1 [C(i)w_1] = \max [C(i)w]$
 $\max_2 [C(i)w_2] = \max [C(i)w]$,
 except $\max_1 [C(i)w_1]$
 $\max_3 [C(i)w_3] = \max [C(i)w]$,
 except $\max_1 [C(i)w_1]$ and $\max_2 [C(i)w_2]$

이다.

단계 2: 후보의 선정

구간별 가중치 V(i)의 합을 단어별 Sw로 구하고, Sw의 세 번째 후보까지 R1, R2, R3를 선택한다.

$$\begin{aligned}
 R_1 &= \max [S_w] \\
 R_2 &= \max [S_w \text{ except } R_1] \\
 R_3 &= \max [S_w \text{ except } R_1 \text{ and } R_2]
 \end{aligned}$$

여기에서, Sw는 $\sum_{i=1}^8 V(i)$ 이다.

단계 3: T1과 T2의 분리

세번째 후보까지 선택된 패턴의 구간별 확신도 값을 합하여 1 그룹 T1의 S11, S12, S13과 2 그룹 T2의 S21, S22, S23를 산출한뒤 분리한다.

$$\begin{aligned}
 T_1 &= S_{11} + S_{12} + S_{13} \\
 T_2 &= S_{21} + S_{22} + S_{23}
 \end{aligned}$$

여기에서, $S_{11} = \sum_{i=1}^4 C(i)_{R1}$, $S_{12} = \sum_{i=1}^4 C(i)_{R2}$,

$$S_{13} = \sum_{i=1}^4 C(i)_{R3}, \quad S_{21} = \sum_{i=5}^8 C(i)_{R1},$$

$$S_{22} = \sum_{i=5}^8 C(i)_{R2}, \quad S_{23} = \sum_{i=5}^8 C(i)_{R3} \text{ 이다.}$$

단계 4: 인식단어 Qz의 판정

두 그룹으로 구분된 T₁와 T₂를 이용하여 인식단어를 판정한다.

우선, T₁와 T₂중 최소값을 택한다.

$$M = \min [T_1, T_2]$$

최소값이 1 그룹인 T₁이면 1 그룹을 구성하고 있는 단어 중 최대값인 단어 Q₁을 찾는다. 1 그룹의 최대값인 단어 Q₁이 2 그룹 T₂의 최소값인 단어 Q₂가 아니면 인식된 단어 Qz로 판정한다.

또한, 최소값이 두 번째 그룹인 T₂일 때도 1 그룹인 경우와 동일한 방법으로 판정한다.

if M = T₁

Q₁ = max (S₁₁, S₁₂, S₁₃) and Q₂ ≠ min (S₂₁, S₂₂, S₂₃)
then Q₁ && Q₂ ⇒ Qz이고,

if M = T₂ 일때

Q₂ = max (S₂₁, S₂₂, S₂₃) and Q₁ ≠ min (S₁₁, S₁₂, S₁₃)
then Q₂ && Q₁ ⇒ Qz이다.

if Q₁ = Q₂ 이면 REJECT한다.

IV. 실험 및 고찰

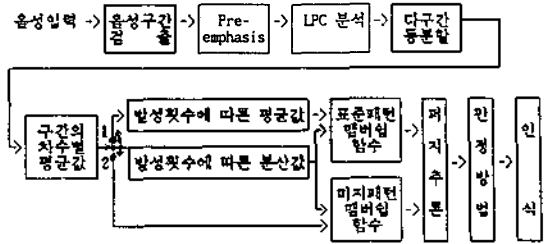
4-1. 음성데이터

본 연구에서 사용된 데이터는 DDD지역명 28개(직할시 이상 지역명 6개, 경기도 지역명 22개)를 20대 이상의 성인남성 2인이 각각 5회씩 발성한 280개이다. 이들 중 표준패턴은 2인이 3회씩 발성한 데이터로부터 다구간 등분할법으로 재작성하여 각 단어당 한개의 패턴을 작성하고, 미지패턴은 나머지 2인이 2회씩 발성한 112개 데이터를 사용한다. 이때, 음성패턴을 작성하기 위한 분석조건은 표 2, 인식구성도는 그림 8과 같다.

표 2. 분석조건

Table 2. Analysis conditions for speech signals.

LPF cut-off Freq	3.4kHz
A/D data	10kHz Sampling, 12-bit
Window Function	Hanning Window and pre-emphasis(A=0.95)
Frame Length	20ms
Frame Period	10ms
Analysis	14-order LPC



1 : Generation reference pattern
2 : Recognition

그림 8. LPC를 이용한 퍼지추론 인식구성도

Fig 8. Block diagram of recognition fuzzy inference using the LPC sequences.

4-2. 결과 및 고찰

본 실험에서는 삼각형 멤버십함수와 변형된 삼각형 멤버십함수로 퍼지추론하여 인식율을 비교하였고, 특히 판정방법을 적용하여 각각 인식율을 비교하였다.

4-2-1. 다구간 등분할법과 멤버십함수에 따른 결과 [13-15]

퍼지추론에 의한 단어인식결과를 표 3에 나타내었고, 4구간과 8구간 그리고 16구간으로 각각 구분하여 멤버십함수의 종류별로 인식율을 구하였다.

표 3.에서보면 삼각형 멤버십함수와 제안된 변형 삼각형 멤버십함수는 유사한 인식결과를 보이고 있으며, 8 구간의 경우가 4 구간과 16 구간보다 인식율이 높은 것은 실험에 사용된 대상단어가 대부분 이음절이기 때문이다.

표 3. 등간격으로 분할된 구간 수와 멤버십함수의 종류에 따른 인식율.

Table 3. Recognition rate according to the multi-sections by equi-segmentation and membership functions.

	[%]		
구간 멤버십함수	4구간	8구간	16구간
삼각형	87.5	92.0	87.5
변형된 삼각형	89.3	92.0	88.4

또한, 삼각형 멤버십함수로 퍼지추론했을때의 인식결과를 표 4.에 나타내었으며, 이로부터 오인식은

표 4. 구간과 삼각형 멤버십함수의 인식결과.

Table 4. Recognition result in case 8 section number with membership function.

표준 미지	서 울	인 천	부 산	대 전	광 주	대 구	수 원	오 산	평 택	안 성	용 인	이 천	여 주	양 평	발 안	김 포	성 남	안 양	원 당	안 산	구 리	문 산	강 화	의 정 부	동 두 천	전 곡	가 평	포 천		
서 울	4																													
인 천		3										1																		
부 산			4																											
대 전				4																										
광 주					4																									
대 구						3																		1						
수 원							4																							
오 산	1							3																						
평 택									3	1																				
안 성										4																				
용 인											4																			
이 천												4																		
여 주													4																	
양 평														4																
발 안															4															
김 포																4														
성 남																	4													
안 양																		4												
원 당																			4											
안 산																				1		3								
구 리																							4							
문 산																								4						
강 화																									4					
의정부																										4				
동두천																											4			
전 곡	1																										3			
가 평																													2	
포 천		1																												3

같은 발성위치에 동일한 음절이 있을 때 발생하는 경우와 음향적 특성이 서로 다른 경우로 분류할 수 있다. 전자는 인천 ↔ 이천, 안산 ↔ 안양, 가평 ↔ 양평, 포천 ↔ 인천인 경우이며, 후자는 대구 → 의정부, 오산 → 서울, 전곡 → 서울, 가평 → 강화, 평택 → 용인으로 오인식되는 경우이다.

따라서 이러한 오인식 문제를 개선하기 위해서 판정방법을 제안하였다.

4-2-2. 판정방법에 따른 결과.

판정방법은 세번째 후보까지의 인식대상으로 고려하는 방법으로 그 예를 표 5.에 나타내었다. 그 과정

표 5. 세번째 후보까지 가중치를 준 예.

Table 5. An example of weights to the third candidate.

표준패턴 구간수	인천	이천	동두천	포천	가평
1	1	1	0	0	0
2	1	1	0	0	0
3	1	0	1	0	0
4	1	0	0	0	0
5	1	0	0	0	0
6	1	1	1	0	0
7	1	0	1	1	0
8	1	0	1	1	0
합	8	3	4	2	0

V. 결 론

본 연구에서는 선형예측계수를 이용하여 퍼지추론 단어인식에 대해 고찰하였다. 주파수 변동을 흡수하기 위해 계수가 차수에 따라 스펙트럼에 미치는 영향과 계수의 변동에 따른 전력 변화율로부터 변형된 삼각형 멤버쉽함수를 제안하여 사용하였으며, 퍼지추론에 판정방법을 제안하여 인식실험 하였다.

그 결과, 제안된 변형 삼각형 멤버쉽함수가 삼각형 멤버쉽함수보다 우수한 인식결과를 보임으로써 계수의 변동에 효과적으로 대응되어 주파수 변동이 흡수됨을 확인하였으며, 가중치를 도입한 판정 방법에 적용함으로써 오인식률이 감소되었다.

DDD 지역명 28개를 대상으로 실험하여 93.8[%]의 인식율을 얻음으로써 사용된 멤버쉽함수와 판정방법이 단어인식에 효과적임을 알 수 있었다.

참 고 문 헌

1. J.D.Markel and A.H.Gray, Linear Prediction of Speech, Springer Verlag, N.Y. 1976.
2. L.R.Rabiner and R.W.Shafer, Digital Processing of Speech Signal, Prentice Hall, 1978.
3. L.R.Rabiner and B.Gold, Theory and Application of Digital Signal Processing, Prentice Hall, 1975.
4. J.Makhoul, "Linear prediction : a tutorial review," Proceeding of IEEE, Vol.63, No.4, pp.561-580, Apr. 1985.
5. S.Saito and K.Nakata, Fundmantals of Speech Processing, Academic Press, 1985.
6. D.K.Burton, "Text-Dependent speaker verification

- using vector quantization source coding," IEEE Trans.Acoust., Speech, Signal Processing, Vol.ASSP-35, No.2, pp.133-143, Feb.1987.
7. J.I.Fujimoto, T.Nakatani and M.Yoneyama, "Speaker-independent word recognition using fuzzy pattern mathing," Fuzzy Sets and Systems,32, pp. 181-191, 1989.
8. B.S.Atal and S.L.Hanauer, "Speech analysis and synthesis by linear prediction of the speech wave," JASA, Vol.50, No.2, 1971.
9. T.Fukabayash and H.Suzuki, "Speech analysis by linear pole-zero model," 信學論, Vol.58-A, NO.5, pp. 270-277, 1975.
10. J.Makhoul, "Spectral linear prediction : properties and applications," IEEE Trans.Acoust., Speech, Signal, Vol.ASSP-23, No.3, pp.282-296, Jun.1975.
11. W.Pedrycz, "Fuzzy set is in pattern recognition : methodology and methods," Pattern Recognition, Vol.23, 1, pp.121-146, 1990.
12. R.De Mori and P.Laface, "Use of fuzzy algorithm for phonemic labeling of continuous speech," IEEE Trans.Acoust., Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No.2, Mar.1980.
13. 최승호, 최갑석, "다구간 등분할법과 퍼지추론을 이용한 단어인식," 음향학회지, 제 12권 4호, 1993.8.
14. 유병욱, 최승호, "선형예측계수의 주파수변동에 따른 멤버쉽함수의 개선," 제 10회 음성통신및 신호처리 워크샵 논문집, 1993.8.
15. D.K Burton, J.E.Shore and J.T.Buk, "Isolated-word speech recognition using multisection vector quantization codebooks," IEEE Trans. Acoust., Speech, Signal: Vol.ASSP-33, No.4, Aug.1985.

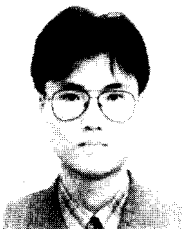
▲ 崔 承 鎬

1955년 8월 24일생

1981년 2월 : 전북대학교 물리학과 (이학사)

1984년 8월 : 명지대학교 대학원 전자공학과(공학석사)

1992년 2월 : 명지대학교 대학원 전자공학과(공학박사)



현재 : 동신대학교 정보통신공학과

▲ 金 亨 根

1958년 4월 12일생

1982년 2월 : 명지대학교 전자공학과 (공학사)

1984년 2월 : 명지대학교 대학원 전자공학과 (공학석사)

1991년 8월 : 명지대학교 대학원 전자공학과 (공학박사)

현재 : 한국방송통신대학 전자계산학과 조교수