

화자인식을 위한 퍼지-상관차원과 퍼지-리아프노프차원의 평가

The Evaluation of the Fuzzy-Chaos Dimension and the Fuzzy-Lyapunov Dimension

유 병 옥* · 박 현 숙** · 김 창 석***

(Byongwook-Yoo · Hyunsook-Park · Changseok-Kim)

ABSTRACT

In this paper, we propose two kinds of chaos dimensions, the fuzzy correlation and fuzzy Lyapunov dimensions, for speaker recognition. The proposal is based on the point that chaos enables us to analyze the non-linear information contained in individual's speech signal and to obtain superior discrimination capability.

We confirm that the proposed fuzzy chaos dimensions play an important role in enhancing speaker recognition ratio, by absorbing the variations of the reference and test pattern attractors.

In order to evaluate the proposed fuzzy chaos dimensions, we suggest speaker recognition using the proposed dimensions. In other words, we investigate the validity of the speaker recognition parameters, by estimating the recognition error according to the discrimination error of an individual speaker from the reference pattern.

Keywords : speaker recognition, fuzzy chaos dimensions, discrimination error

1. 서 론

화자인식(speaker recognition)이란 음성파에 포함된 개인차를 이용하여 화자를 알아내는 것을 말한다. 이러한 화자인식 시스템은 예금 잔고의 전화문의 등과 같은 개인 정보의 검색, 각종 기밀 정보의 검색, 금고, 공장, 군사 시설 등 중요 기밀 지역에서의 출입통제를 위한 신분 확인 수단으로 사용할 수 있고 서명, 도장, 신용카드, 신분 증명서 등과 같은 각종 증서를 대신할 수 있으며 범죄 수사에서 범인 식별에 이용되는 등 그 응용분야가 넓다.

화자인식 시스템에서 많이 사용되는 특징 파라미터로서는 에너지 또는 진폭, 평균 피치(pitch)주파수, 장시간 평균 스펙트럼, 선형 예측 파라미터의 평균값과 주변분산, 캡스트럼

* (주)데이터씨큐어 정보기술연구소장

** 동아방송대학 방송기술과 조교수

*** 명지대학교 전자정보통신공학부

(cepstrum)의 시간 변화를 많이 사용한다. 그러나 이러한 해석방법은 선형시스템에 기초한 해석이므로 부가적인 인식기의 개선 문제가 뒤따르고 근원적 해결 방법이라고 할 수 없다. 화자인식의 문제는 다음과 같이 몇 가지가 있다.

첫째 음운적 정보에 대한 음향학적 물리량에 대해서는 상당히 밝혀져 있지만 개인적 특징을 나타내는 정보에 대해서는 연구가 미흡하므로 개인성 정보에 대한 폭넓은 연구가 요구된다. 둘째, 동일화자에 의한 패턴변동 차이를 최소화할 수 있는 특징으로 파라미터가 개개인의 음성차이를 최대로 변별하고 동일화자의 음성차이를 최소화할 수 있는 특징 파라미터 추출이 중요하다. 셋째, 음성 모방자를 얼마만큼 정확히 구분하는가 하는 것이다. 넷째로는 시간에 따른 음성 파형의 특징 변화를 어떻게 일반화할 수 있는가 하는 문제로 음성은 동일화자라도 시간 변화에 따라 차이가 있으므로 등록된 화자를 언제나 정확하게 인식하기 위해서는 특징 파라미터의 특성은 시간이 경과해도 변화가 적어야 한다.^[1-3]

한편, 최근에 과거의 자연현상을 이해하고 이를 이용하고자 하는 카오스(chaos) 현상에 대한 연구가 진행되고 있으며, 시계열신호에 대해서도 카오스차원을 추정하여 신호의 복잡한 변화 속에 숨겨져 있는 비선형 동적 정보를 파악함으로써 신호를 정량화 하여 시계열신호의 패턴인식과 분석에 활용하려는 연구들이 시도되고 있다. 대표적인 카오스의 차원에는 프랙탈 차원으로 알려진 상관차원과 리아프노프차원(Lyapunov dimension)이 있다. 이러한 카오스차원은 비선형으로 결합된 결정론적 동적시스템의 특징을 나타내는 것이므로 음성 또는 뇌파와 같은 불규칙신호의 동적특징을 분류하는 식별력이 우수한 파라미터이다. 그러나 카오스차원을 화자인식에 적용하기 위해서는 어트랙터의 동적특징을 분류하는 식별도 우수하고 궤도의 표준패턴과 시험패턴간의 흡수력이 커야 한다. 즉, 표준패턴에 대해서 식별력이 크며, 동일패턴에 대해서는 그 변동흡수특성이 우수하여야 한다.^[2-5]

한편 퍼지이론은 대사에 대한 애매함의 정도를 소속정도 즉, 유사도로 표현하여 주는 것이다. 따라서 뛰어난 변별력의 특징을 가지고 있으며 궤도의 정량화 척도라 할 수 있는 카오스차원에 카오스 어트랙터궤도에 퍼지관계를 설정하여 준다면 음성표준패턴 상호간의 변별특징도 양호하고, 표준패턴에 속하는 시험패턴의 궤도특징을 유사도로 그 변동을 흡수할 수 있다. 따라서 퍼지화된 카오스차원은 화자인식과 인식한 화자에 대한 음성인식에 새로운 파라미터로 활용할 수 있다.

그림 1에서와 같이 표준패턴과 시험패턴 어트랙터간의 기하학적 구조에 대한 변동량을 흡수해야 하지만 스트레인지 어트랙터는 그 크기와 위상구조가 초기값에 크게 영향을 받으므로 동일화자의 동일 음성이라도 어트랙터의 기하학적크기와 형태가 크게 달라질 수 있어서 단순한 기하학적 형태에 대한 퍼지유사도의 적용으로 패턴변동량을 흡수하는 것은 곤란하다.

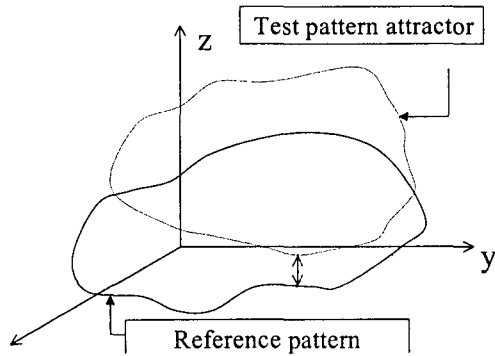


그림 1. 스트레인지 어트랙터의 변동량출수

본 연구에서는 비선형 발생메커니즘으로부터 발생되는 음성신호에 대해서 시계열상의 비선형 특징을 가지고 있는 카오스 어트랙터를 구성하고, 궤도특성을 정량적으로 평가하는 분별력 강한 상관차원과 리아프노프차원에 대해 퍼지멤버쉽함수를 적용한 퍼지상관차원과 퍼지리아프노프차원을 제안하고 두 차원에 대해 화자인식파라미터로서의 성능을 검토한다.

2. 퍼지상관차원의 제안

본 연구에서는 음성신호의 길이를 등분할 법을 사용하여 선형화한 후 어트랙터를 재구성하였으며, 상관적분의 변화율^[4-10]을 퍼지집합으로 설정하기 위해서 그림 2와 같은 멤버쉽함수 $\mu_A(x)$ 를 설정하였다. 그림 2에서 X_{ii} 은 다수개의 표준패턴에 대한 공통적인 물리적 특성을 가지는 파라미터이고, d_k 는 패턴의 변동량으로서 주로 분산값을 많이 사용한다. 패턴변동이 흡수될 수 있도록 변동폭 $X_{ii}-d_k$ 와 $X_{ii}+d_k$ 를 결정한다. 임의의 재구성된 10개의 동일한 음성 어트랙터의 구성점 X_i 를 평균하여 $X_0, X_1, \dots, X_{ii}, \dots, X_{iK}$ (K 은 데이터 수)를 작성한 다음 표준패턴으로 사용하였으며 그림 2에서 멤버쉽함수를 구성하기 위한 평균값 X_{ii} 로 사용하였다.

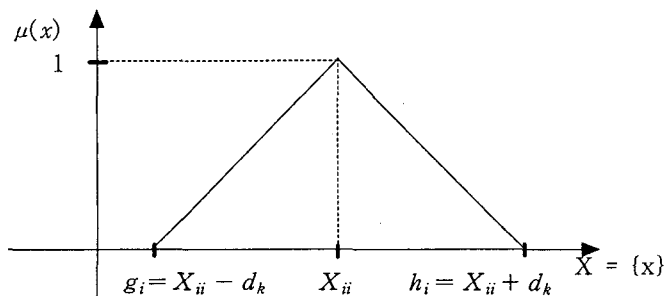


그림 2. 삼각형 퍼지 유사도 함수

그림 3과 같이 각각의 거리 $|X_j - X_{ii}|$ 를 계산하여 i 에 대한 각 평균 d_i 를 구하여 그림 2의 패턴변동을 흡수하기 위한 멤버쉽함수의 구성요소로 사용한다. 따라서 본 연구에서 사용된 멤버쉽함수 $\mu_A(X_i)$ 는

$$\mu_A(X_i) = \begin{cases} \frac{1}{g_i - X_{ii}} (X_{ii} - X_j) + 1, & X_j \leq X_{ii} \\ \frac{-1}{h_i - X_j} (X_{ii} - X_j) + 1, & X_j > X_{ii} \end{cases} \quad (1)$$

여기서 $g_i = X_{ii} - d_k$, $h_i = X_{ii} + d_k$ 이다.

그림 3은 시각 x_t 에서 한 표준패턴과 시험패턴을 동일한 m차원공간에 구성한 것이다. 기존의 상관차원 계산에 있어서는 소속 내는 1, 소속 외는 0으로 계산하였으나 표준패턴의 각각의 거리 $|X_j - X_{ii}|$ 를 평균한 d_k 를 이용하여 X_{ii} 에 가까운 정도를 멤버쉽함수에 의해 퍼지수로 계산함으로써 퍼지상관적분 (fuzzed correlation integral)을 계산한다.

본 연구에서 제안한 퍼지상관적분 $FC^m(d_k)$ 은 다음 식과 같다.

$$FC^m(d_k) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \mu_A(X_i) \quad (2)$$

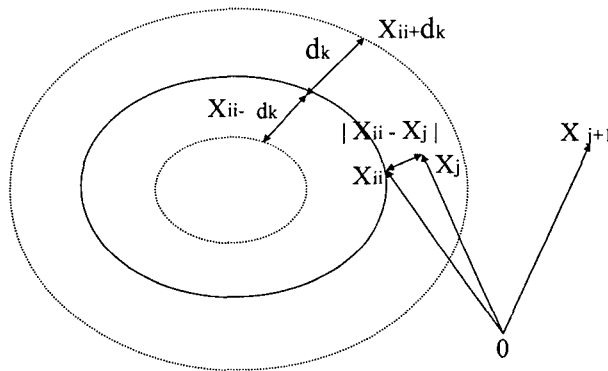


그림 3. X_{ii} 에 대한 X_j 의 퍼지유사도 추정

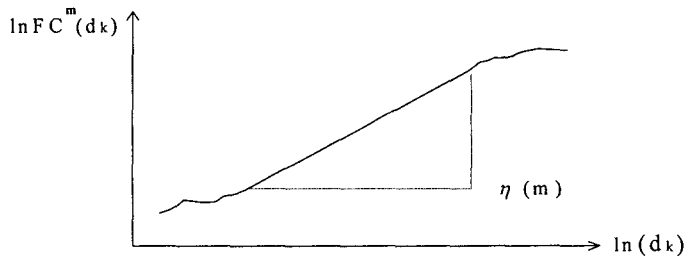


그림 4. 퍼지상관적분과 거리평균 d_k 와의 관계

그림 3에서와 같이 m 차원 공간에 재구성한 어트랙터상의 한 점 $X_{ii}(i=1,2,\dots,n)$ 를 중심으로 나머지 $(N-1)$ 개의 점에 대해 유사도를 계산하여 변화량을 측정함으로써 규칙성을 찾는다. 또한 d_k 를 최소값에서 최대값으로 배열한 다음 그 사이에서 퍼지상관적분 $FC^m(d_k)$ 가 식 (3)과 같이 선형적 비례관계가 존재한다면 어트랙터 비례범위 $d_1 \leq d_i \leq d_2$ 에서 미소구와 스트레인지 어트랙터는 자기상사성을 가지므로 유사정도의 추정이 가능하며, 규칙성을 찾을 수 있고, 시험패턴과 표준패턴과의 변동분흡수도 가능하다.

$$FC^m(d_k) \propto d_k^{\eta(m)} \tag{3}$$

그림 4는 유사도 $\mu_A(X_i)$ 와 퍼지상관적분 $FC^m(d_k)$ 의 선형적 비례관계를 나타낸 것으로서 $\eta(m)$ 은 $\ln d_k$ 와 $\ln FC^m(d_k)$ 의 일정한 기울기이며, 식 (4)과 같다. 식 (5)과 같이 공간차원 m 을 증가시키면 $\eta(m)$ 이 포화됨으로서 그림 5와 같이 접근하는 값 FD_c 가 추정되며, 이것을 퍼지상관차원이라 하고 식 (5)과 같이 정의한다.

$$\eta(m) = \lim_{d_k \rightarrow d_{max}} \frac{\ln FC^m(d_k)}{\ln(d_k)} \tag{4}$$

$$FD_c = \lim_{m \rightarrow \infty} \eta(m) = constant \tag{5}$$

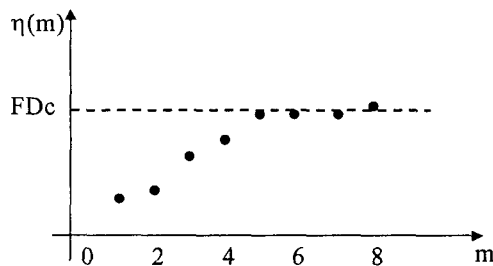


그림 5. 퍼지상관차원의 추정

3. 퍼지 Lyapunov 차원의 제안

Lyapunov 차원은 스트레인지 어트랙터상의 어느 한 점과 그 주위의 점에 대해 시간에 따른 변화정도를 미소벡터의 변화로 추정한다. 즉, 궤도불안정성을 각 방향의 변화율로 추정하므로 다수 개의 어트랙터에 대해 기하학적구조의 질서에 대한 정도 차이를 정밀히 분류하는 성능이 우수하다.^[4-11,17] 그러나 수직(수평) 방향에 대한 지수적 확대(축소)벡터인 Lyapunov 스펙트럼 λ_i 가 단위벡터방향의 변화만을 고려하여 추정하므로 표준패턴 어트랙터와 동일한 시험패턴 어트랙터에 대해서는 각 점들의 표준패턴 어트랙터에 대한 변화량을 흡수하지 못하며, 동일화자라 해도 표준패턴 어트랙터와 다른 시험패턴 어트랙터 간에 유사정도를 추정하지 못하므로 어트랙터 패턴 인식에 대한 성능이 저하된다.

따라서 Lyapunov 지수의 추정에 퍼지유사도를 적용하여 Lyapunov 차원을 퍼지화함으로써 패턴인식하는 것이 타당하다. 그러나 Lyapunov 차원추정에서는 표준패턴에 대해서 동일화자 동일 패턴의 변동량 흡수가 안되므로 그림 6과 같이 표준패턴의 어트랙터상의 한 점을 중심으로 단위구간 내에 존재하는 미소벡터들의 변화정도를 흡수하기 위해 퍼지유사도를 이용한다. 따라서 어트랙터 미소점들의 벡터계산과정에서 표준패턴과 시험패턴간의 벡터 변화량을 퍼지유사도로서 추정하여 패턴인식에 적합한 퍼지 Lyapunov 차원을 제안한다.

그림 6에서와 같이 표준패턴 어트랙터궤도상의 한 점을 X_{ii} 라고 하고, 이 점을 중심으로 이 구에 들어가는 시험패턴의 어트랙터상의 다른 점 $M(j=1, 2, \dots, M)$ 개의 X_j 를 설정한다. 이때 X_{ii} 에서 구내의 M 개의 점 X_j 에 대해 퍼지 Lyapunov 차원을 추정하기 위해서 변위벡터 y_j 에 대한 퍼지유사도 $\mu_y(X_j)$ 를 다음과 같이 정의한다.

$$\mu_y(X_j) = \begin{cases} \frac{1}{h_y - X_{ii}} (X_j - X_{ii}) + 1, & X_j \leq X_{ii} \\ \frac{-1}{g_y - X_{ii}} (X_j - X_{ii}) + 1, & X_j > X_{ii} \end{cases} \quad (6)$$

여기서 $h_y = X_{ii} - \Delta_1$, $g_y = X_{ii} + \Delta_1$ 이다.

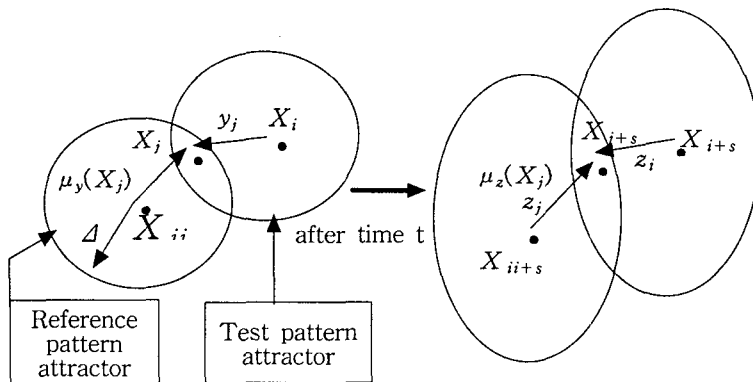


그림 6. 퍼지유사도 $\mu_y(X_j)$ 와 $\mu_z(X_j)$ 의 추정

그림 6에서 시간이 s만큼 지난 후, 미소반경 Δ_2 로 하는 구의 중심은 X_{ii} 에서 X_{ii+s} 로, 각 상태 X_j 는 X_{j+s} 로 변화하므로 시간 t+s에서 z_j 에 대한 퍼지유사도 $\mu_z(X_j)$ 는 다음과 같이 정의한다.

$$\mu_z(X_j) = \begin{cases} \frac{1}{h_z - X_{ii}} (X_{j+s} - X_{ii}) + 1, & X_{j+s} \leq X_{ii} \\ \frac{-1}{g_z - X_{ii}} (X_{j+s} - X_{ii}) + 1, & X_{j+s} > X_{ii} \end{cases} \quad (7)$$

여기서 $h_z = X_{ii} - \Delta_2$, $g_z = X_{ii} + \Delta_2$ 이다.

미소반경 Δ_2 로 하는 구의 반경과 시간 s가 충분히 작다면 행렬 $\overline{H}(X_j)$ 는 퍼지유사도 $\mu_y(X_j)$ 와 $\mu_z(X_j)$ 사이에 선형관계가 성립되므로 퍼지화된 벡터행렬은 다음과 같이 표현할 수 있다.

$$\mu_z(X_j) = F\overline{H}(X_j) \mu_y(X_j) \quad (8)$$

행렬 $F\overline{H}(X_j)$ 는 Jacobian 행렬에 근사적이므로 다음 식의 거리오차 S를 최소로 하는 최소자승법으로 결정할 수 있다.

$$S = \sum_{j=1}^M [\mu_z(X_j) - F\overline{H}(X_j) \mu_y(X_j)] \quad (9)$$

행렬 $F\overline{H}(X_j)$ 의 제 kl 성분을 $F\overline{H}_{kl}$ 이라고 하면 식 (6)의 $\mu_y(X_j)$ 와 $\mu_z(X_j)$ 에 대한 거리오차 S에 대해 각각의 $F\overline{H}_{kl}$ 에 대한 최소조건은 $\frac{\partial S}{\partial F\overline{H}_{kl}} = 0$ 이므로 다음 식이 성립된다.

$$F\overline{H}(X_j) V = C \quad (10)$$

$F\overline{H}(X_j)$ 를 추정하고 행렬 $F\overline{H}(X_j)$ 를 Lyapunov 차원에서의 $\overline{G}(X_j)$ 로 적용하여 퍼지화된 Lyapunov 스펙트럼을 구할 수 있다.

Lyapunov 차원추정에서는 패턴변동량을 흡수하지 못하므로 표준패턴과 동일한 시험패턴간의 변동량을 흡수하기 위해서 퍼지유사도 $\mu_y(X_j)$ 와 $\mu_z(X_j)$ 의 선형비례관계에 있는 행렬벡터 $F\overline{H}(X_j)$ 를 적용한다. 또한 Lyapunov 차원에서와 같이 연산을 감소시키기 위해서 서로 직교하는 단위벡터를 적용하면 다음 식을 얻을 수 있다.

$$Fe(j+1) = F\overline{H}(X_j) u(i) \quad (11)$$

퍼지화된 벡터행렬 $F\overline{H}(X_j)$ 는 $\mathbf{u}(i)$ 의 벡터방향으로만 추정되는 퍼지화된 $F\mathbf{e}(j+1)$ 을 구할 수 있다.

$F\mathbf{e}(j+1)$ 를 직교화하면

$$F\hat{\mathbf{e}}(j+1) = F\mathbf{e}(j+1) - \sum_{i=1}^j \langle F\mathbf{e}(j+1), \mathbf{u}(i+1) \rangle \mathbf{u}(i+1) \quad (12)$$

이고, 이것을 반복 계산하여 다음의 정규직교계 $\mathbf{u}(j+1)$, ($j=1, 2, \dots, m$)를 구한다.

$$\mathbf{u}(i+1) = \frac{F\hat{\mathbf{e}}(j+1)}{|F\hat{\mathbf{e}}(j+1)|} \quad (13)$$

$\mathbf{u}(i+1)$, ($i=1, 2, \dots, m$)을 식 (12)에 의해 다시 사상하는 과정을 되풀이하여 얻어진 $F\hat{\mathbf{e}}(i)$ 의 계열을 Lyapunov 지수의 추정식에 대입하면 퍼지 Lyapunov 지수 γ_i ($i=1, 2, \dots, m$)를 추정하는 식은 다음과 같다.

$$\gamma_i = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \log |F\hat{\mathbf{e}}(i)| \quad (14)$$

식 (14)에서 구해진 γ_i 는 표준패턴과 시험패턴에 대한 Lyapunov 지수의 변동분에 대한 것을 의미하므로 본 논문에서 제안하고자 하는 퍼지 Lyapunov 차원의 추정식은 식 (14)에서 구해진 퍼지 Lyapunov 지수를 i 순으로 나열하여 $\gamma_{k+1} < 0$ 이 만족되는 γ 가 있을 때 다음 식과 같다.

$$FD_L = k + \frac{\sum_{i=1}^k \gamma_i}{|\gamma_{k+1}|} \quad (15)$$

4. 화자인식평가 방법

화자인식과정은 화자를 식별하여 등록된 음성이 발생한 음성과 유사정도를 판정함으로써 인식을 결정한다. 그러므로 발생한 음성의 음향적 특성과 화자간의 특징이 구분되어야 하고 동일화자의 동일음성에 대해서는 특징 변동이 흡수되어야 한다. 본 논문에서는 화자인식 파라미터를 평가하기 위해서 화자식별과 패턴변별 및 변동흡수에 대해 다른 화자, 다른 모음과 동일화자, 동일모음과의 차를 구하였다. 그 차가 클수록 우수한 화자인식 파라미터이다. 즉, 모든 화자가 패턴별로 발생한 음성에 대한 차원과 등록된 표준패턴차원과의 차를 구

하여 식 (16)과 같은 식별오차율에 의하여 화자를 인식하는 평가방법을 제안하였다. 이 식별 오차율은 다른 화자의 동일한 패턴에 대해서 그 차가 크면 클수록 우수한 화자인식파라미터이며, 동일화자이지만 다른 모음패턴일 때에도 역시 차가 크면 클수록 우수하다. 동일화자, 동일모음일 때에는 그 차가 적을수록 우수한 파라미터이다. 평가를 위해 화자별로 각 모음 패턴마다 10회씩 발생한 시험패턴에 대해 차원을 구한 후 화자에 따라 등록된 표준패턴의 차원과 차를 구하여 10회 발생된 모음에 따라 평균하여 식별오차를 구한다.^[10]

$$\delta = \frac{1}{P} \sum_{\pi=1}^P \left| \frac{\text{표준패턴의차원} - \text{시험패턴의차원}}{\text{표준패턴의차원}} \right| \times 100 \quad (16)$$

여기서 P는 음성발성수이다. 구해진 식별오차는 오차율에 대한 값의 크기를 고찰함으로써 제안된 퍼지상관차원과 퍼지리아프노프차원의 화자식별 파라미터로 평가하였다. 또한 표준패턴과 동일한 시험패턴의 식별오차를 δ_1 로 하고, δ_1 과 가장 근사값인 식별오차를 δ_2 로 하여 다음 식 (17)과 같이 차를 구함으로써 그 차가 크면 화자인식 파라미터로서 우수하도록 오인식이 일어날 수 있는 평가기준으로 사용하였다.

$$R_s = \delta_2 - \delta_1 \quad (17)$$

화자인식은 화자별 모음마다 임의의 시험패턴을 입력하여 제안된 퍼지화된 상관차원과 퍼지리아프노프차원을 구하고, 등록된 각 화자의 1개로 구성되는 모음표준패턴과의 차를 구한 후 그 차에서 최소값을 나타내는 것을 인식화자의 발생모음으로 화자인식하였다.

5. 퍼지상관차원에 의한 화자인식

서로 다른 스트레인지 어트랙터 패턴 사이의 특징을 정량적으로 나타낼 수 있도록 동일 패턴 사이에 존재하는 변동은 흡수될 수 있도록 퍼지 상관적분 $FC^m(d_k)$ 를 임의의 X_i 와 X_{i+1} 의 유클리안 거리에 대해 거리 d_k 을 미소량으로 증가시키면서 멤버쉽함수를 식 (1)에 의하여 계산하고 식 (2)에 의해 퍼지상관적분을 추정한다.

식 (16)에 의해 화자 'A'에 대해 각 A, B, C 화자의 10회 발생한 모음에 대한 평균 식별 오차를 구하여 표 1에 나타내었다. 표 1에서 A화자 5개 모음에 대한 식별오차율은 전체 평균값이 5.11 %로써 다른 화자에 대한 식별오차와는 오인식의 차가 최소 9.78 %에서 최대 12.49 %이다. 이로부터 표준패턴과 동일한 모음에 대해서도 화자식별이 가능하며 또한 동일한 화자에서도 모음의 인식이 가능함을 알 수 있으며 화자식별 파라미터로서의 퍼지상관 차원값이 우수하다는 것을 알 수 있다. 한편 퍼지상관차원에 대한 화자식별과 모음인식에 대한 파라미터로의 타당성을 알아보기 위해 임의의 화자에 대한 5개의 모음 표준패턴을 정하고 식 (16)에 따라 식별오차를 구하여 모음인식을 하였다. 각 모음마다 구해진 식별오차

에서 표준패턴 모음과 동일한 모음의 식별오차를 $\delta 1$ 으로 하고 S1과 가장 근사값인 식별 오차를 $\delta 2$ 로 하여 식 (17)과 같이 인식거리 R_s 를 표 2에 나타내었다.

표 2는 표 1에서 A화자의 '아'의 4.15 %에 대해 다른 화자 C의 '아'에 대한 식별오차율 10.09 %와 최소거리차 5.94 %를 구하여 나타낸 것이다. 다른 화자의 경우에서도 식별오차율의 거리가 매우 크므로 퍼지상관차원은 화자인식에 타당한 파라미터임을 알 수 있다. 표 1에서 식별오차가 가장 적은 것을 각 화자마다 각 모음을 표준패턴으로 하여 인식하도록 실험을 행한 결과, 표 3과 같은 화자 인식률을 얻을 수 있었다. 표 3은 각 화자마다 각 모음을 시험음성으로 하여 화자인식 실험을 행한 결과를 나타낸 것이다. 표 3을 보면 화자A '아'의 퍼지상관차원의 인식률은 96.7 %를 나타내었다. 이는 상관적분은 임의의 반경 r 을 기준으로 그 변화량의 정도를 추정하지를 않으나 퍼지상관적분은 그 변화량의 정도를 퍼지 유사도로써 추정하여 차원을 추정하기 때문에 인식률이 높은 것으로 나타낸 것이다. 다른 경우의 패턴에 대한 인식률에 있어서도 전체 평균인식률은 95.34 %로써 매우 높은 것으로 나타났다.

표 1. A화자의 표준패턴에 대한 식별오차율

Reference pattern Speaker		/a/	/e/	/i/	/o/	/u/	Average
		A	/a/	4.15	25.67	14.85	
/e/	28.24		3.66	20.72	19.35	25.53	
/i/	18.91		32.69	4.76	25.81	23.62	
/o/	30.59		21.36	20.84	6.99	10.47	
/u/	23.12		35.47	25.77	13.50	5.97	
B	/a/	16.24	29.24	30.83	22.22	26.02	15.58
	/e/	25.10	12.16	26.01	25.09	29.11	
	/i/	26.93	24.54	20.06	25.97	25.01	
	/o/	26.94	26.36	27.46	14.12	23.55	
C	/u/	27.69	37.30	18.34	21.81	15.32	17.60
	/a/	10.09	20.74	29.28	34.95	33.33	
	/e/	20.95	24.24	29.58	30.73	34.66	
	/i/	24.13	28.35	10.93	34.83	30.78	
	/o/	27.11	24.81	22.12	15.75	28.17	
D	/u/	29.37	24.60	26.16	25.86	26.99	16.82
	/a/	17.51	22.39	24.15	26.72	27.48	
	/e/	21.49	20.21	20.01	24.76	23.72	
	/i/	21.60	24.32	15.77	24.06	21.84	
	/o/	21.42	27.48	25.28	16.67	24.09	
E	/u/	13.35	12.49	17.95	20.02	13.93	14.89
	/a/	15.76	19.66	12.26	15.71	15.85	
	/e/	20.80	12.58	24.38	23.58	21.09	
	/i/	15.25	17.28	13.66	20.21	21.42	
	/o/	17.61	18.41	23.26	20.92	19.71	
	/u/	21.15	23.32	26.70	21.87	11.53	

이는 퍼지상관차원이 음성시계열로 구성되는 서로 다른 표준 패턴사이에서 어트랙터 궤도에 대해 변별력이 강하고, 각 표준 패턴인 어트랙터와 동일 음성 어트랙터인 시험패턴에 대해서는 그 시간에 따른 궤도 변동분을 흡수 할 수 있으므로 높은 인식률이 나타난 것이다. 따라서 퍼지상관차원은 화자인식 파라미터로서 활용이 가능한 것으로 볼 수 있다.

표 2. 퍼지상관차원에 대한 평균식별오차율의 거리

표준패턴 \ 화자	아	에	이	오	우
A	5.94	8.92	6.17	7.13	5.56
B	8.08	10.49	9.10	10.21	12.10
C	9.67	8.88	11.19	6.53	6.42
D	5.76	5.67	7.74	6.94	6.33
E	8.27	7.56	13.57	14.82	15.53
평균	7.5	8.3	9.6	9.1	9.2

표 3. 퍼지상관차원에 의한 화자인식률

표준패턴 \ 화자	퍼지상관차원				
	A	B	C	D	E
아/a/	96.7	93.3	100	86.7	93.3
에/e/	96.7	96.7	96.7	90.0	93.3
이/i/	96.7	93.3	96.7	96.7	96.7
오/o/	93.3	96.7	93.3	93.3	96.7
우/u/	93.3	96.7	90.0	96.7	96.7
평균	95.34	95.34	95.34	92.68	95.34

6. 퍼지 Lyapunov 차원에 의한 화자인식

Lyapunov 차원은 시간에 따른 어트랙터 궤적의 각 미소벡터방향에 따른 변화를 정밀하게 추정하므로 다수개의 시험패턴 어트랙터간의 변별특성은 우수하지만 동일 화자에 의해 발생된 서로 다른 음성에 대해서도 어트랙터 궤도 변화에 민감하다. 따라서 이러한 Lyapunov 차원에 퍼지유사도를 적용하여 변별 특성과 동일 패턴에 대한 변동흡수를 해결한 퍼지 Lyapunov 차원으로 화자인식률 향상을 시도하였다. 시각 t에서 표준패턴 어트랙터에 대해 한 점 X_t 를 중심으로 일정 반경의 미소구안에 존재하는 시험패턴 어트랙터의 M개의 점을 계산하고, X_t 의 t+s후에 변화를 퍼지유사도를 이용하여 시험패턴의 어트랙터의 궤적변화를 정밀히 추정하며 그 변동을 표준패턴에 대해 가중치로 흡수함으로써 퍼지 Lyapunov 차원을 추정하였다.

식 (14)에 의한 퍼지 Lyapunov 지수를 추정하기 위해 식 (10)의 $\overline{FH}(X_i)$ 를 식 (11)으로부터 계산하고 식 (14)을 계산하여 퍼지 Lyapunov 지수를 구하였다. m은 5로 하고, 구해진 퍼지 Lyapunov 지수 중 큰 순서대로 바꿔 가면서 음(-)이 아닌 최대정수를 k로 하여 퍼지 Lyapunov 차원을 추정하였다.

화자인식의 유용성을 알아보기 위해서 A화자의 각 모음 중 1회 발생한 음성을 표준패턴으로 설정하고, 나머지 각 음성에 대해 퍼지 Lyapunov 차원을 구하여 화자별로 인식된 식별오차율의 평균값 δ 를 식 (16)으로 구하여 표 4에 나타내었다.

표 4를 고찰하면 평균 4.07 %로 식별오차율이 낮음을 알 수 있다. 각 화자에 대해 퍼지 Lyapunov 차원에 의한 화자인식을 수행한 결과를 표 5에 나타내었다. 전체평균 97 % 인식율은 퍼지 Lyapunov 차원이 화자인식에 매우 우수한 파라미터임을 보이고 있다.

표 5를 고찰하면 퍼지 Lyapunov 차원을 사용한 화자인식률이 퍼지 상관차원을 사용한 화자인식률보다 향상된 인식률을 나타냄을 알 수 있다. 이는 퍼지 Lyapunov 차원이 어트랙터의 특징을 더욱 잘 나타내며, 음성시계열로 구성되는 서로 다른 표준 패턴 사이에서 어트랙터 궤도에 대해 변별력이 강하고, 각 표준패턴인 어트랙터와 동일한 음성 어트랙터인 시험패턴에 대해서는 그 시간에 따른 궤도변동분을 흡수할 수 있으므로 퍼지 상관차원보다 높은 인식률이 나타난 것이다. 따라서 퍼지 Lyapunov 차원은 패턴인식 파라미터로서 우수하다. 서로 다른 음성패턴간은 분별시키고, 동일 패턴에 대해서는 패턴변동분을 흡수시키는 패턴인식에 적합한 카오스차원임을 확인할 수 있다. 또한 음성 발생시스템 자체의 특징을 잘 나타낸다고 할 수 있다. 따라서 퍼지 Lyapunov 차원은 화자인식에 매우 효과적인 파라미터이다.

표 4. A화자의 표준패턴에 대한 식별오차율 δ

Speaker \ Reference pattern		「a」	「e」	「i」	「o」	「u」	Average
		「a」	「e」	「i」	「o」	「u」	
A	「a」	4.22	29.14	26.83	23.59	21.05	4.07
	「e」	24.13	3.04	27.45	20.37	23.76	
	「i」	20.14	20.17	3.96	21.85	25.59	
	「o」	24.84	25.27	24.34	4.16	20.91	
	「u」	26.13	20.84	26.59	19.68	4.96	
B	「a」	20.27	32.21	39.32	31.77	28.99	20.29
	「e」	31.86	19.82	34.23	32.01	35.62	
	「i」	36.06	27.32	24.27	26.52	34.40	
	「o」	35.34	31.63	28.22	19.59	33.60	
	「u」	33.95	35.77	29.36	27.42	17.53	
C	「a」	15.58	31.64	33.51	36.42	31.83	20.01
	「e」	28.56	26.68	39.76	35.87	29.45	
	「i」	27.61	32.00	16.13	37.34	28.91	
	「o」	35.65	26.61	37.11	18.56	31.54	
	「u」	32.09	29.69	36.12	30.77	23.11	

Speaker \ Reference pattern		Reference pattern					Average
		「a」	「e」	「i」	「o」	「u」	
D	「a」	26.89	29.04	31.70	29.56	27.21	22.37
	「e」	29.54	25.66	32.35	35.90	26.42	
	「i」	28.15	31.37	20.12	36.20	28.23	
	「o」	27.28	30.49	33.33	18.18	30.92	
	「u」	31.71	29.32	34.85	34.92	21.01	
E	「a」	28.04	24.46	23.12	26.43	30.03	23.33
	「e」	28.28	20.61	27.45	26.60	31.38	
	「i」	27.13	26.33	19.89	28.74	27.45	
	「o」	25.46	30.55	30.52	24.67	29.98	
	「u」	26.59	29.31	27.20	29.71	23.44	

표 5. 퍼지 Lyapunov 차원에 의한 화자인식률(%)

Speaker \ Reference Pattern	A	B	C	D	E
「a」	98.0	96.0	98.0	98.0	98.0
「e」	96.0	94.0	96.0	98.0	96.0
「i」	98.0	96.0	98.0	96.0	98.0
「o」	96.0	98.0	96.0	96.0	98.0
「u」	96.0	98.0	96.0	98.0	98.0
Average	96.8	96.4	96.8	97.2	97.6

7. 퍼지화된 카오스차원의 화자인식 고찰

퍼지상관차원과 퍼지 Lyapunov 차원에 대해, 화자식별력과 인식된 화자 내에서의 음성 인식을 통한 화자인식 성능을 평가한다. 표준패턴과 동일한 화자에서 다른 시험패턴에 대한 인식력, 그리고 표준패턴과 다른 화자의 동일한 시험패턴에 대해서 화자들간의 식별력을 비교 검토하기 위해서 표 1, 4의 식별오차율로부터 화자별로 평균하여 그림 7과 그림 8에 나타냈다. A화자의 시험패턴에 대해서는 표준패턴과 동일한 시험패턴으로 추정된 식별오차율은 화자인식된 값이므로 평균계산과정에서 제외하였다.

그림 7과 그림 8을 고찰하면 화자 B, C, D, E에서 A화자의 표준패턴에 대해 식별오차율이 높게 나타내고 있으므로 보아 화자식별파라미터로서 퍼지상관차원과 퍼지 Lyapunov 차원은 우수하다. 퍼지상관차원에서 A화자의 식별오차율은 22.8 %로써 D, E 화자에 비하여 높게 나타난 것은 화자인식된 동일한 음성시험패턴을 제외하고 동일한 화자의 다른 음성시험패턴에 대한 화자별 평균값을 나타낸 것이기 때문이다.

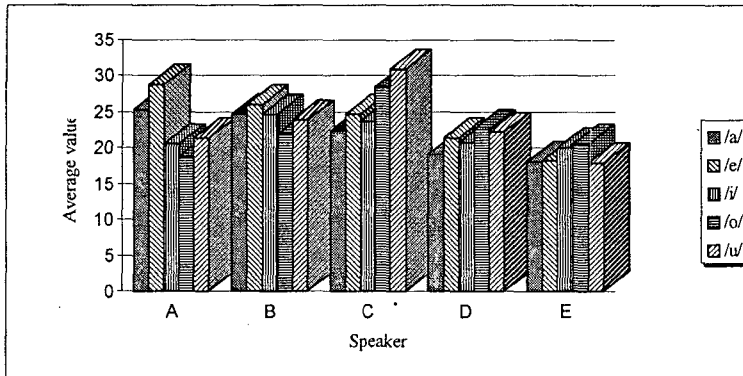


그림 7. 퍼지상관차원에 대한 화자별 식별오차율 δ 의 평균값

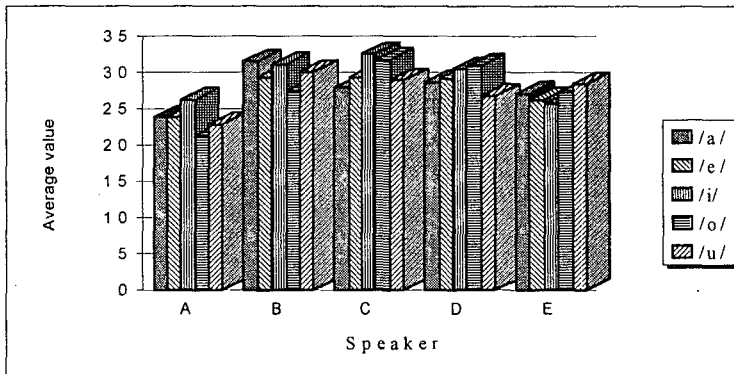


그림 8. 퍼지 Lyapunov 차원에 대한 화자별 식별오차율 δ 의 평균값

그러나 퍼지 Lyapunov에서는 A화자의 화자별 평균값이 퍼지상관 차원과 근사하게 나타났으며, 화자 B, C, D, E에서 화자별 식별오차율 평균값이 A화자보다 높은 것은 퍼지상관차원보다 퍼지 Lyapunov 차원이 화자식별력이 우수하다는 것을 알 수 있다.

퍼지상관차원보다 퍼지 Lyapunov 차원이 식별력이 우수한 것으로 나타난 것은 퍼지 Lyapunov 차원이 시간에 따른 어트랙터의 미소변화에 대한 변화를 추정하기 때문이다. 표 1, 4로부터 표준패턴별로 평균값을 구하여 그림 9에 나타내었다. 그림 9에서도 퍼지상관차원보다 퍼지 Lyapunov 차원의 화자식별력이 우수한 것으로 나타났다.

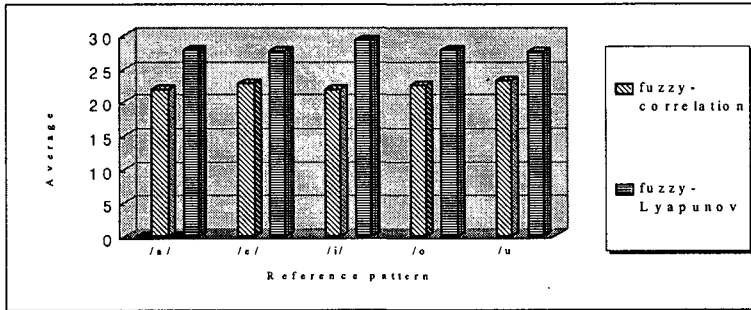


그림 9. 차원에 대한 표준패턴 식별오차율 δ 의 평균값

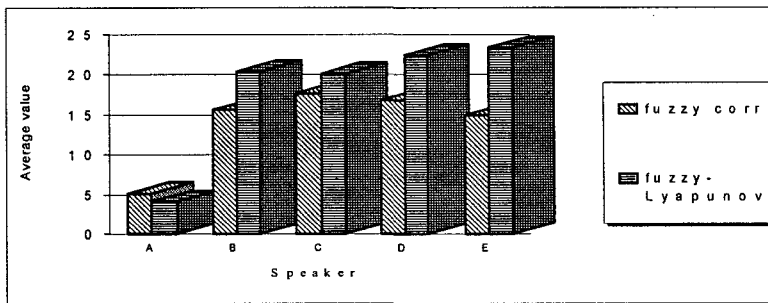


그림 10. 표준패턴과 동일한 시험패턴에 대한 화자별 식별오차율

퍼지상관차원과 퍼지 Lyapunov의 표준패턴과 동일한 시험패턴에 대한 음성인식력을 비교하기 위하여 각 화자의 표준패턴과 동일한 시험패턴에 대해 화자별로 식별오차율 평균하여 그림 10에 나타냈다. 그림과 같이 시험패턴이 A화자일 때 표준패턴과 동일한 시험패턴의 식별오차율은 퍼지 리아프노프차원이 퍼지 상관차원보다 우수한 결과를 얻었다.

퍼지 Lyapunov 차원은 화자식별력 뿐만 아니라 동일 화자 내에서도 동일한 시험패턴을 인식하는 성능이 우수함을 확인하였다.

8. 결론

카오스차원에 대해 화자인식향상을 위한 퍼지상관차원과 퍼지 Lyapunov 차원의 추정방법을 제안하였다. 화자인식 파라미터로서의 퍼지카오스차원을 평가하는 화자인식 평가방법을 제안하였다. 이상과 같이 여러 가지 제안한 방법을 화자인식에 적용하여 평가한 결과 다음과 같은 결론을 얻었다.

퍼지카오스차원을 적용한 화자식별은 표준패턴과 다른 시험패턴에 대해서 퍼지상관차원보다 퍼지 Lyapunov 차원이 우수하였고 동일한 시험패턴에 대해서도 퍼지 Lyapunov 차원이 퍼지상관 차원보다 우수하였다.

제안한 각 차원을 적용하여 화자인식 실험을 수행한 결과 인식률은 퍼지상관차원 94.6 %,

퍼지 Lyapunov 차원 97.0 %의 인식률을 나타냄으로써 퍼지 Lyapunov 차원이 가장 우수하였으며, 제안된 카오스차원이 화자인식 파라미터로서 적합함을 확인하였다.

참 고 문 헌

- [1] Douglas O'shaughnessy. 1986. "Speaker recognition". *IEEE ASSP Magazine*. Vol.74, 4-17.
- [2] Sadaok., Furui. 1992. *Advances in Speech Signal Processing*.
- [3] S. Saito, K. Nakata. 1985. *Fundmantals of Speech Processing*, Academic Press.
- [4] 合原一辛. 1992. *腦どカオス*. 東京電氣大學.
- [5] Thomas S. Parker, Leon O. Chua . 1987. "Chaos:A Tutorial for Engineers". *Proceeding of the IEEE* . Vol. 75, No. 8, 982 -1008.
- [6] Alan Wolf . 1985. ."Determining Lyapunov Exponants from a time series". *Physica 16D*, 285-317.
- [7] B.Malraison, P. Atten, P. Brege and M. Dubois. 1983. "Dimension of strange attractors". *an experimental de Physique Letters*. Vol. 44, No. 22, 897-902.
- [8] J. D. Farmer, E. Ott and J. A. Yorke. 1983 ."The dimension of chatic attractors". *Physical*, Vol.7, 153-180.
- [9] P. Eckmann, S. O. Kanmphorst, D. Ruelle and S. Ciliberto. 1986. "Liapunov exponent from a time series". *Phys. Rev. A*, Vol. 6, 4971-4979.
- [10] P. Frederickson, J. L. Kaplan, E. D. Yorke and J. A. Yorke. 1983. "The Lyapunov dimension of stringe attractors". *J. of Diff. Equations*, Vol. 49, 185-207.
- [11] 유병욱, 김창석. 1987. "Lyapunov 차원을 이용한 화자식별 파라미터추정". *한국음향학회지 제16권 제4호*.
- [12] Bart Kosko. 1991. *Neural Ntworks and Fuzzy Systems*. Prentice Hall.
- [13] 寺野,菅野. 1987. *ファジィ システム入門*. オーム社.
- [14] Earl Cox. 1994. *The Fuzzy systems Hand book*. AP.
- [15] Fujimoto, T. Nakatani and M. Yoneyama. 1989. "Speaker independent word recognition using fuzzy pattern mathing". *Fuzzy Sets and System*. Vol. 32, 181-191.
- [16] Alan Wolf. 1985. "Determining Lyapunov Exponants from a time series". *Physica 16D*, 285-317.
- [17] D. S. KIM. 1984. Chaotic Analysis of Speech Signals". *한·일·중 카오스 공동 세미나 논문집*. 124-129.
- [18] 최승호, 최갑석. 1993. "다구간 등분할법과 퍼지추론을 이용한 단어인식". *한국음향학회지*. Vol. 12-4, 47-56.

접수일자: 2000. 8. 2.

게재결정: 2000. 9. 3.

▲ 유 병 옥

서울 강남구 청담동 45-4

(주)데이터씨큐어 정보기술연구소

Tel : +82-2-549-2540 (O), Fax : +82-2-549-2509

e-mail: bwyoo@TeleSecure.co.kr

▲ 김 창 석

경기도 용인시 남동 산 38-2

명지대학교 전자정보통신공학부 (449-728)

Tel : +82-335-337-2300 (O), Fax : +82-335-330-6374

e-mail: kimcs@myongji.ac.kr

▲ 박 현 숙

경기도 안성시 삼죽면 632-18

동아방송대학 방송기술과 (456-880)

Tel : +82-334-670-6700 (O)