

論文99-36S-1-12

화자인식을 위한 퍼지상관차원 제안

(A Proposition of the Fuzzy Correlation Dimension for Speaker Recognition)

兪炳旭*, 金昌錫**, 朴賢淑***

(Byongwook-Yoo, Changseok-Kim, and Hyun Sook-Park)

요 약

본 논문은 음성신호가 카오스 신호임을 확인하고 화자인식 파라미터로 사용하기 위해 상관차원을 분석하였다. 화자식별과 인식 향상을 위하여 개인의 성도특징을 매우 잘 나타내는 음성의 스트레인지 어트랙터를 구성하고 퍼지유사도를 상관차원에 적용하여 퍼지상관차원을 제안하였다. 퍼지상관차원은 어트랙터 구성점들의 상관관계를 퍼지상관적분으로 추정하고 공간차원값에 따라 퍼지상관지수가 일정하게 수렴되는 차원값을 구하여 표준패턴 어트랙터와 시험패턴 어트랙터의 변동을 흡수하였다. 퍼지상관차원에 대해 화자와 표준패턴별로 식별 오차의 평균값에 따른 거리를 추정함으로써 화자인식파라미터의 타당성을 검토하였다.

Abstract

In this paper, we confirmed that a speech signal is a chaos signal, and in order to use it as a speaker recognition parameter, analyzed chaos dimension. In order to raise speaker identification and pattern recognition, by making up the strange attractor involving an individual's vocal tract characteristics very well and applying fuzzy membership function to correlation dimension, we proposed fuzzy correlation dimension. By estimating the correlation of the points making up an attractor as fuzzy correlation integral and obtaining the dimension value where fuzzy correlation exponents are limited according space dimension value, fuzzy correlation dimension absorbed the variation of the reference pattern attractor and test pattern attractor. Concerning fuzzy correlation dimension, by estimating the distance according to the average value of discrimination error per each speaker and reference pattern, investigated the validity of speaker recognition parameter.

I. 서론

* 正會員, (株)데이터씨큐어 情報技術研究所

(DataSecure corp. Institute of Information Technology)

** 正會員, 明知大學校 電子情報通信工學部

(Myongji Univ., Dept. of Electronics, Information and Communication Engineer)

*** 正會員, 東亞放送技術大學, 放送技術科

(Dong Ah Broadcasting College. Dept. of broadcasting technology)

接受日字:1998年8月1日, 수정완료일:1998年12月17日

음성신호 또는 뇌파신호 발생시스템과 같이 매우 복잡하게 비선형으로 결합된 결정론적 동적시스템을 선형시스템으로 해석할 경우 여러문제들이 발생되고 있으며, 음성인식 또는 뇌파신호처리등의 경우 부가적인 파라미터 또는 인식알고리즘의 개선이 필요하기 때문에 처리시스템이 매우 복잡하게 되고 있다. 따라서 이러한 신호발생메커니즘은 비선형시스템으로 해석하는 것이 이상적이다.^[1-2]

최근에 과거의 자연현상을 이해하고 이를 이용하고

자 하는 카오스(chaos)현상에 대한 연구가 진행되고 있으며, 시계열신호에 대해서도 카오스차원을 추정하여 신호의 복잡한 변화속에 숨겨져 있는 비선형 동적 정보를 파악함으로써 신호를 정량화 하여 시계열신호의 패턴인식과 분석에 활용하려는 연구들이 시도되고 있다.^[3-6] 결정론적 규칙에 의하여 지배되는 카오스신호에 대한 분석방법은 신호의 궤적특징의 결정론적 규칙, 즉 카오스차원을 찾는 것이다. 이러한 카오스차원은 비선형으로 결합된 결정론적 동적시스템의 특징을 나타내는 것이므로 음성 또는 뇌파와 같은 불규칙신호의 동적특징을 분류하는 식별력이 우수한 파라미터이다.^[3-9] 그러나 카오스차원을 음성인식 또는 화자인식에 적용하기 위해서는 궤도의 동적특징을 분류하는 식별도 우수하고 궤도의 표준패턴과 시험패턴간의 흡수력이 커야한다. 즉, 표준패턴에 대해서 분별력이 크며, 동일패턴에 대해서는 그 변동흡수특성이 우수하여야 한다.^[1-2,11-15] 카오스차원은 발생규칙정보가 숨겨진 스트레인지 어트랙터에 대해 정량화된 변화량을 척도로서 규칙을 나타내는 것이므로 패턴특징량의 변동을 흡수하는 음성인식보다는 서로 다른 패턴에 대해 표준패턴간의 규칙정도의 차이를 알아내는데 적합한 파라미터로서 인식보다 식별이 더 우수하다.

한편 퍼지이론은 대사에 대한 애매함의 정도를 소속 정도 즉, 유사도로 표현하여 주는 것이다. 따라서 뛰어난 변별력의 특징을 가지고 있으며 궤도의 정량화 척도라 할 수 있는 카오스차원에^[10,11] 카오스 어트랙터 궤도에 퍼지관계를 설정하여 준다면 음성표준패턴 상호간의 변별특징도 양호하고, 표준패턴에 속하는 시험패턴의 궤도특징을 유사도로 그 변동을 흡수할수 있다. 따라서 퍼지 카오스차원은 화자인식과 인식한 화자에 대한 음성인식에 새로운 파라미터로 활용할 수 있다.

본 연구에서는 비선형 발생메커니즘으로부터 발생하는 음성신호에 대해서 시계열상의 비선형특징을 가지고 있는 카오스 어트랙터를 구성하고, 궤도특성을 정량적으로 평가하는 분별력 강한 카오스차원에 대해 퍼지멤버쉽함수를 적용한 퍼지 카오스 차원을 제안하고 화자인식파라미터로서의 성능을 평가한다.

II. 퍼지상관차원의 제안

카오스신호의 특성을 정량화하기 위한 카오스차원 중 일반적으로 많이 이용되는 것이 Grassberger와

Procaccia에 의해 제안된 상관차원이다. 공간차원m에 구성된 스트레인지 어트랙터상에서 어트랙터 궤도상의 한점을 중심으로 일정한 반경내에 존재하는 다른점과의 거리를 추정한 것을 상관적분이라고 한다. 반경의 크기를 변화시키면서 상관적분과 일정한 기울기의 선형관계가 존재할 때 그 변화율을 상관지수라고 하며 상관차원은 이 상관지수에 대해 공간차원 m을 증가시켰을 때 일정하게 수렴되는 추정된 값이며 이 상관차원은 스트레인지 어트랙터의 형태에 대한 특징정보를 표현한다.

m의 공간차원을 R^m로 표현할 때 구성된 어트랙터 상의 한 점을 X_i라고 하면 상관적분 C^m(r)은 다음과 같이 정의된다.^[3-9]

$$C^m(r) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i, j=1}^N U(r - |X_i - X_j|) \quad (1)$$

$$U(r - |X_i - X_j|) = \begin{cases} 1, & (r - |X_i - X_j|) \geq 0 \\ 0, & (r - |X_i - X_j|) < 0 \end{cases} \quad (2)$$

여기서 r은 반경이며, |X_i - X_j|는 유클리드안 거리이고, N은 구성된 어트랙터상의 점의 수이다. 그림 1에서 m차원공간에 구성된 어트랙터상의 한 점 X_i (i = 1, 2, ..., N)를 중심으로 나머지 (N-1)개의 점에 대해서 미소반경 r인 구내에 존재하는 점 X_j(i≠j)를 계산하고, X_i를 중심으로 하여 반복계산하며 미소반경 r을 증가시키면서 구내에 포함되는 점들의 수를 계산하여, 그 변화량을 추정한다. 임의의 m공간차원 벡터를 X_i=(x_{i1}, x_{i2}, ..., x_{im}), X_j=(x_{j1}, x_{j2}, ..., x_{jm})라 하면 상관적분식 (1)에서 거리|X_i - X_j|의계산식은 다음과 같다.

$$|X_i - X_j| = \left\{ \sum_{i=1}^m (X_i - X_j)^2 \right\}^{1/2} \quad (3)$$

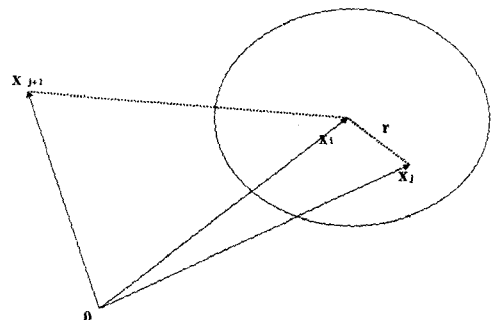


그림 1. 상관적분의 계산.
Fig. 1. Calculation of correlation integral.

이러한 상관적분 식(1)로 부터 미소반경 r 을 증가시키면서 r 과 상관적분 $C^m(r)$ 이 임의의 영역에서 선형적인 비례관계가 존재할 때 일정한 기울기를 상관지수라고 하고 공간차원 m 을 증가시킬 때 일정한 값으로 포화되는 상관지수의 수렴되는 값을 상관차원이라고 한다. 상관차원은 발생시스템의 카오스특징 즉, 궤도형태의 특징을 차원의 정밀한 값으로 나타내므로 궤도형태를 더욱 더 정밀하게 분류한다. 궤도형태의 특징을 정밀하게 구분하는 것은 상관차원이 패턴식별에서는 우수한 파라미터 이지만 동일한 표준패턴사이에서도 그 특징을 구분하는 특성을 나타내므로 패턴인식 파라미터로서는 우수하지 못하다.^[11]

한편 퍼지에 의한 패턴인식방법은 표준패턴과 시험패턴간의 변동량을 흡수하고, 그 정도를 퍼지수로 나타내며^[1-2,11-15] 음성시계열에 대한 카오스 어트랙터에서는 시계열의 동특성이 잘 보존되어 있으므로 각 패턴간의 변별력이 뛰어나다고 할 수 있다. 이러한 카오스 어트랙터에 퍼지패턴매칭방법을 적용하여 유사도를 측정하면 각 카오스 어트랙터에 대한 시간별 궤도 변화의 변동분흡수가 가능하다. 또한 주어진 시계열패턴에 대하여 카오스어트랙터를 구성한 다음 어트랙터에 대한 카오스차원을 계산하는 과정에 퍼지수를 도입하면 각 어트랙터에 대해서 변별력도 우수하고, 패턴인식에 적합한 새로운 파라미터인 퍼지화된 카오스차원의 추출이 가능하다.

본 연구에서는 음성신호의 길이를 등분할 법을 사용하여 선형화 한후 어트랙터를 재구성하였으며, 상관적분의 변화율을 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.

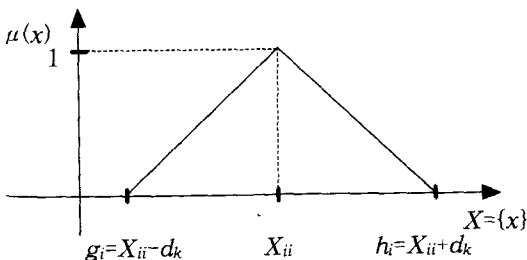


그림 2. 삼각형 퍼지 유사도 함수
Fig. 2. Fuzzy membership function.

패턴변동이 흡수될 수 있도록 변동폭 $X_{ii}-d_k$ 와 $X_{ii}+d_k$ 를 결정한다.^[12-15] 임의의 재구성된 10개의 동일한 음성 어트랙터의 구성점 X_i 를 평균하여 $X_{i1}, X_{i2}, \dots, X_{iK}, \dots, X_{iK}$ (K 은 데이터 수)를 작성한 다음 표준패턴으로 사용하였으며 그림 2에서 멤버쉽함수를 구성하기 위한 평균값 X_{ii} 로 사용하였다. 그림3과 같이 각각의 거리 $|X_j - X_{ii}|$ 를 계산하여 i 에 대한 각 평균 d_i 를 구하여 그림 2의 패턴변동을 흡수하기 위한 멤버쉽함수의 구성요소 D 로 사용한다. 따라서 본 연구에서 사용된 멤버쉽함수 $\mu_A(X_i)$ 는

$$\mu_A(X_i) = \begin{cases} \frac{1}{g_i - X_{ii}} (X_{ii} - X_i) + 1, & X_i \leq X_{ii} \\ \frac{-1}{h_i - X_{ii}} (X_{ii} - X_i) + 1, & X_i > X_{ii} \end{cases} \quad (4)$$

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

그림 3은 시각 x_i 에서 한 표준패턴과 시험패턴을 동일한 m 차원공간에 구성한 것이다. 기존의 상관차원 계산에 있어서는 소속내는 1, 소속외는 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.

$$FC^m(d_k) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i,j=1}^N \mu_A(X_i) \quad (5)$$

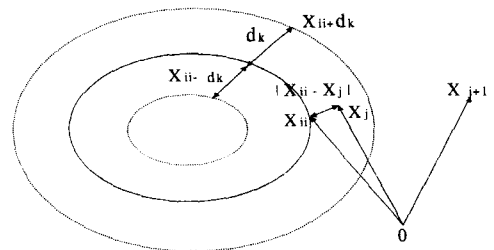


그림 3. X_{ii} 에 대한 X_j 의 퍼지유사도 추정
Fig. 3. Estimation of fuzzy membership grade of measurement X_{ii} .

그림 3에서와 같이 m 차원 공간에 재구성한 어트랙터상의 한점 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 (6)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 시퀀스패턴과 표준패턴과의 변동분흡수도 가능하다.

$$FC^m(d_k) \propto d_k^{\eta(m)} \tag{6}$$

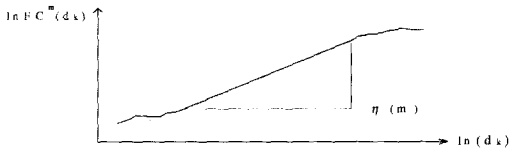


그림 4. 퍼지상관적분과 거리평균 d_k 와의 관계
Fig. 4. Relation of between fuzzy correlation integral and d_k .

그림 4는 유사도 $\mu_A(X_i)$ 와 퍼지상관적분 $FC^m(d_k)$ 의 선형적 비례관계를 나타내낸 것으로서 $\eta(m)$ 은 $\ln d_k$ 와 $\ln FC^m(d_k)$ 의 일정한 기울기이며, 식 (7)과 같다. 식(8)과 같이 공간차원 m 을 증가시키면 $\eta(m)$ 이 포화되므로서 그림5와 같이 접근하는 값 FD_c 가 추정되며, 이것을 퍼지상관차원이라 하고 식(8)과 같이 정의한다.

$$\eta(m) = \lim_{d_k \rightarrow d_{max}} \frac{\ln FC^m(d_k)}{\ln(d_k)} \tag{7}$$

$$FD_c = \lim_{m \rightarrow \infty} \eta(m) = constant \tag{8}$$

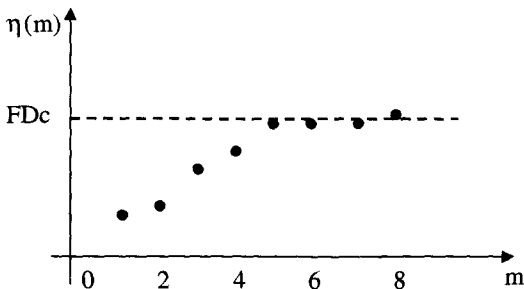


그림 5. 퍼지상관차원의 추정
Fig. 5. Estimation of fuzzy correlation dimension.

III. 카오스 차원에 의한 화자인식

본 연구에서는 음성인식에 대한 퍼지상관차원의 타당성 검토를 위해 사용된 음성은 16bit 양자화된 신호로서 주변잡음이 존재하는 실험실안에서 20대 남녀 화자 5명에 대해 1명당 20회씩 발생하도록 하고, 처음 10회를 표준패턴으로 사용하였으며, 나머지 10회를 시험패턴으로 사용하였다. 이 음성신호는 sonitec사의 audio interface SAIB를 거쳐서 음성신호를 증폭하여 4 [kHz]의 저역통과 필터를 통과시켜 고주파 성분을 제거한 후 샘플링주파수를 10 [kHz]로 하여 A/D 변환한 음성데이터를 수집하였다.

음성데이터에 대한 카오스역학계의 불안정 궤도특징을 알아보기 위해서는 데이터를 Taken 매핑정리에 따라 m 차원 공간의 어트랙터를 구성하여 관측해야 한다. 어트랙터를 구성하기 위해서는 지연시간 τ 를 구해야 하며 τ 가 너무 작게 설정되면 재구성 상태의 공간내의 데이터는 극단적으로 상관성이 너무 크게 되어 최적의 어트랙터를 구성할 수 없으므로 τ 는 주요 주기의 이하로 설정하여야 한다. 한편 음성신호는 자음 또는 모음 각각에도 여러 주파수성분으로 이루어져 있으므로 본 연구에서는 음성신호에 대해 파워 스펙트럼을 구한 후 가장 큰 주파수성분을 주요 주기로한 지연시간 τ 를 결정하여 어트랙터를 구성하였다. 위와 같은 방법으로 구성된 「아」, 「에」, 「이」, 「오」, 「우」의 어트랙터를 그림 5(a)~(e)에 나타내었으며, 음성데이터의 어트랙터는 전형적인 스트레인지 어트랙터형태를 나타내고 있으므로, 음성신호는 카오스신호라는 것을 알 수 있으며, 카오스분석방법에 의해 음성신호를 분석하는 것은 타당성이 있다.

그림 6의 어트랙터간을 고찰하면 형태가 차별이 있음을 확인할 수 있으므로 스트레인지 어트랙터의 특징을 정량적으로 나타내주는 카오스차원은 시계열데이터에 대해 패턴분류 또는 패턴인식의 파라미터로서 사용이 가능하며, 상관차원을 음성인식 파라미터로서 사용이 가능한가를 확인하기 위해서 식 (1)의 상관적분값으로부터 상관차원을 구하여 표1에 나타내었다.

표 1에서 화자의 구별없이 각 패턴에 따른 상관차원의 분포값 D_2 을 보면, 「아」에 대한 상관차원의 값이 4.1 ~ 4.6으로 「에」의 4.4 ~ 4.7과 겹치는 분포를 갖으며, 다른 패턴에 대해서도 패턴간의 식별이 곤란하므로 음성인식의 파라미터로서 적합하지 않다.

그러나 상관차원 분포값 D1를 분석하면 A화자의 상관차원의 분포는 4.6 ~ 4.8, B화자의 상관차원의 분포는 4.2 ~ 4.5로서 화자간의 차이를 나타내고 있으므로 화자식별이 용이함을 알 수 있다. 따라서 상관차원은 음성인식보다는 화자식별에 적합함을 알 수 있다. 본 연구에서는 음성인식은 서로다른 패턴에 대한 식별력과 동일 패턴에 대한 변동흡수의 특성이 존재하여야 하므로 카오스의 궤도특징 분석방법에 있어서 궤도의 형태식별과 궤도형태의 인식의 특성을 갖도록 퍼지개념을 도입하여 퍼지상관차원값을 구하였다.

표 1. 화자별 상관차원의 평균값.

Table 1. Average of correlation dimension with speaker.

표준패턴 화자	/a/(아)	/e/(에)	/i/(이)	/o/(오)	/u/(우)	분포 D1
A	4.629	4.760	4.504	4.698	4.887	4.6~4.8
B	4.264	4.498	4.336	4.202	4.134	4.2~4.5
C	4.503	4.599	4.603	4.453	4.497	4.4~4.6
D	4.183	4.078	4.348	4.126	4.091	4.0~4.3
E	4.565	4.876	4.964	4.935	4.883	4.5~4.8
분포 D2	4.1~4.6	4.4~4.7	4.3~4.6	4.2~4.7	4.1~4.8	

즉, 서로 다른 스트레인지 어트랙터 패턴사이의 특징은 정량적으로 나타낼수 있도록 동일 패턴사이에 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 를 미소량으로 증가시키면서 멤버쉽함수를 식 (4)에 의하여 계산하고 식(5)에 의해 퍼지상관적분을 추정한다. 퍼지상관차원은 흐름도 그림 7에 식 (8)을 적용하여 추정하였다. 표 2는 화자 A에 대한 퍼지상관차원값을 나타낸 것으로서 공간차원 m가 증가함에 따라 퍼지상관차원값이 증가하고, 공간차원 m값이 작을 경우에는 퍼지상관차원값이 추정되지 않으므로 수렴될 때의 공간차원 m값을 구하여 퍼지상관차원값을 추정한다.

그림 8은 표 2에 따라 공간차원 m와 퍼지상관차원의 관계를 그림으로 나타낸 것으로 공간차원 m이 4이상에서 다른 화자에 대해서도 대부분 패턴이 수렴됨을 알 수 있었다.

퍼지상관차원에 대한 화자식별과 음성인식 파라미터로서의 효용성을 검토하기 위해서 등분할법에 의해 작

성된 각 화자별 표준패턴으로 각 모음에 대한 퍼지상관차원의 값을 그림 7에 따라 구하여 표 3에 나타내었다. 표 3은 각 화자별 표준패턴에 대해 각 모음별로 퍼지상관차원을 구하여 평균한 것이다. 표 3에서 화자의 구별 없이 모음에 따라 퍼지상관차원의 분포를 나타낸 아래의 D2분포를 살펴보면, 「아」에 대한 퍼지상관차원의 값이 4.1~4.7 으로 「에」의 4.3~4.8과 겹치는 분포를 갖으며, 다른 모음에 대해서도 모음간의 식별이 곤란한 것을 알 수 있으므로 음성인식의 파라미터로는 사용하기에 부족한 파라미터이다. 그러나 D1의 분포에 있어서는 A화자의 경우 퍼지상관차원의 분포는 4.65 ~ 4.89, B화자의 상관차원의 분포는 4.09 ~ 4.39, 화자 C의 퍼지상관차원값은 4.40 ~ 4.60으로 화자식별이 양호한 것으로 나타났으며 표 1의 상관차원의 화자별 평균값 D1의 분포에 대한 화자식별정도보다 퍼지상관차원이 더욱 화자식별력이 뛰어나다고 할 수 있다.

표 2. 공간차원 m의 변화에 따른 퍼지상관차원

Table 2. Fuzzed correlation dimension of space dimension m.

화자 \ m	1	2	3	4	5	6
/ a/(아)	0.567	1.766	2.165	3.132	3.126	4.367
/ e/(에)	0.774	1.765	2.331	3.301	4.361	4.557
/ i/(이)	0.652	1.743	2.655	3.644	3.683	4.620
/ o/(오)	0.583	1.738	2.900	3.899	3.957	4.795
/ u/(우)	0.573	1.844	2.747	3.766	3.732	4.749

한 화자마다 10회발성한 모음에 대해 등분할법을 이용하여 한 모음에 대한 표준패턴을 작성하였다. 즉 화자별 그리고 모음별의 표준패턴을 구한다음, 화자별 그리고 모음별로 10회의 시험패턴에 대해 퍼지상관차원을 구하고 다음 식(9)과 같이 식별오차율을 정의하여 퍼지상관차원에 대한 평균 식별오차율을 구하였다.

$$\text{식별오차율} = \frac{1}{10} \sum_{i=1}^{10} \left| \frac{\text{표준패턴의차원} - \text{시험패턴의차원}}{\text{표준패턴의차원}} \right| \times 100 \quad (9)$$

여기서 i는 음성 발생 수이다. 식(9)에 의해 화자 'A'에 대해 각 A,B,C화자의 10회 발생한 모음에 대한 평균 식별오차를 구하여 표 4에 나타내었다.

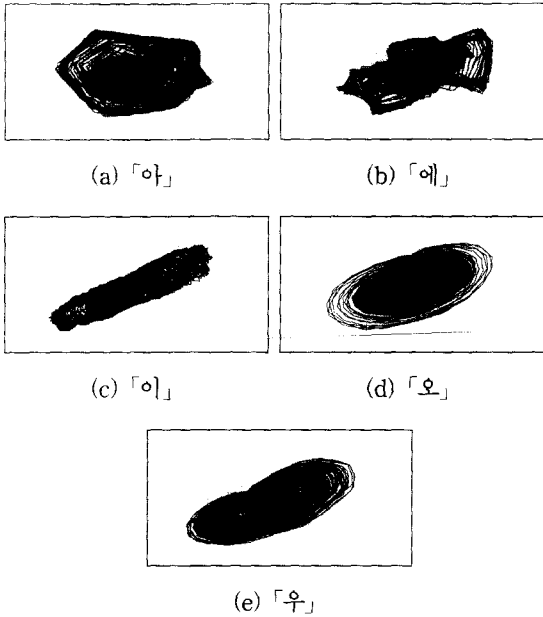


그림 6. A화자의 각 패턴에 따른 스트레인지 어트랙터
 Fig. 6. Strange attractor for each of pattern in speaker A.

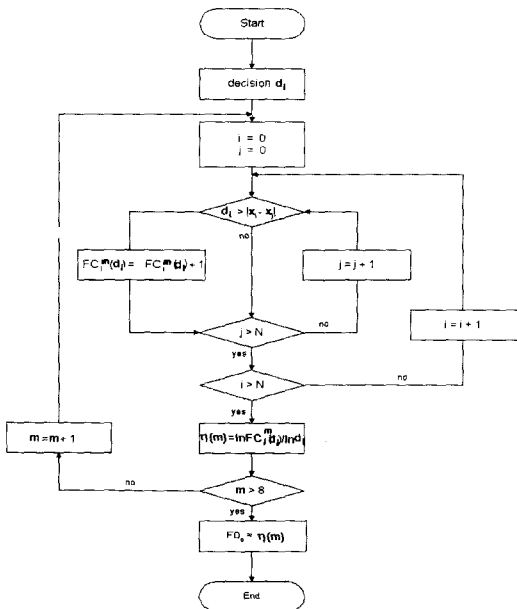


그림 7. 퍼지상관차원의 추정흐름도
 Fig. 7. Flow chart for estimation of fuzzy correlation dimension.

표 4에서 A화자 5개 모음에 대한 식별오차율은 전체 평균 9.99%로서 다른 화자에 대한 식별오차 16.46%와 15.17%와는 평균 식별오차의 거리가 각각 2.05%,

5.18%로서 표준패턴과 동일한 모음에 대해서도 화자 식별이 가능하며 또한 동일한 화자에서도 모음의 인식이 가능함을 알 수 있으며 화자식별 파라미터로서의 퍼지상관차원값이 우수하다는 것을 알 수 있다. 한편 퍼지상관차원에 대한 화자식별과 모음인식에 대한 파라미터로의 타당성을 알아보기 위해 임의의 화자에 대한 5개의 모음 표준패턴을 정하고 식(9)에 따라 식별 오차를 구하여 모음인식을 하였다. 각 모음마다 구해진 식별오차에서 표준패턴 모음과 동일한 모음의 식별 오차를 δ_1 으로 하고 S1과 가장 근사값인 식별오차를 δ_2 로 하여 식(10)과 같이 인식거리 R_s 를 표5에 나타내었다.

$$R_s = \delta_2 - \delta_1 \quad (10)$$

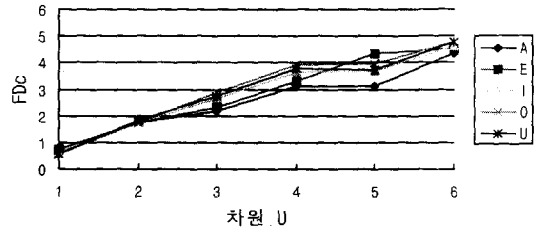


그림 8. 퍼지상관차원에서의 m의 결정
 Fig. 8. Decision of space dimension m in fuzzy correlation dimension.

표 3. 퍼지상관차원의 화자별 평균값
 Table 3. Average value of fuzzed-correlation dimension.

퍼지상관차원 / 화자	/a / (아)	/e / (에)	/i / (이)	/o / (오)	/u / (우)	분포 D1
A	4.700	4.849	4.653	4.750	4.899	4.65~4.89
B	4.157	4.322	4.097	4.293	4.395	4.09~4.39
C	4.401	4.589	4.490	4.572	4.608	4.40~4.60
D	4.054	4.179	4.219	4.153	4.092	4.05~4.21
E	4.855	4.926	4.836	4.905	4.972	4.83~4.97
분포 D2	4.0~4.8	4.1~4.9	4.0~4.8	4.1~4.9	4.0~4.9	

표 5는 표 4에서 A화자의 '아'의 6.77%에 대해 다른 화자 B의 '아'에 대한 식별오차율 18.97%, C화자의 10.33%와 최소 거리인 3.56%을 구하여 나타낸 것이다. 다른 화자의 경우에서도 식별오차율의 거리가 매우 크므로 퍼지상관차원은 화자인식에 타당한 파라미터임을 알 수 있다. 표 4에서 식별오차가 가장 적은 것을 각 화자마다 각 모음을 표준패턴으로 하여 인식

하도록 실험을 행한 결과, 표 6와 같은 화자 인식율을 얻을 수 있었다.

표 6은 각 화자마다 각 모음을 시험음성으로 하여 화자인식 실험을 행한 결과를 나타낸 것이다.

표 6을 보면 화자A 「아」의 상관차원에 의한 인식율은 86.7%이고 퍼지상관차원의 인식율은 96.7%를 나타내었다. 이는 상관적분은 임의의 반경 r을 기준으로 그 변화량의 정도를 추정하지를 않으나 퍼지상관적분은 그 변화량의 정도를 퍼지 유사도로써 추정하여 차원을 추정하기 때문에 인식율이 높은 것으로 나타난 것이다. 다른 경우의 패턴에 대한 인식율에 있어서도 상관차원보다는 퍼지상관차원에 의한 경우가 전체 평균인식율은 95.34 %로서 매우 높은 것으로 나타났다.

표 4. A화자의 표준패턴에 대한 식별오차율
Table 4. δ for reference pattern of speaker A.

Reference pattern Speaker	/a/	/e/	/i/	/o/	/u/	Average	
A	/a/	4.15	25.67	14.85	16.29	25.81	5.11
	/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	
	/u/	27.69	37.30	18.34	21.81	15.32	
C	/a/	10.09	20.74	29.28	34.95	33.33	17.60
	/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	
	/u/	29.37	24.60	26.16	25.86	26.99	
D	/a/	17.51	22.39	24.15	26.72	27.48	16.82
	/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	
	/u/	13.35	12.49	17.95	20.02	13.93	
E	/a/	15.76	19.66	12.26	15.71	15.85	14.89
	/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 인식율이 나타난 것이다. 따라서 퍼지상관차원은 화자인식 파라미터로서

활용이 가능한 것으로 볼 수 있다.

표 5. 퍼지상관차원에 대한 평균식별오차율의 거리

Table 5. Distance of δ for fuzzy correlation dimension

패턴 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 6. 퍼지상관차원에 의한 화자인식율

Table 6. Speaker recognition ratio by fuzzed correlation dimension.

패턴 화자	상관차원				
	A	B	C	D	E
아/a/	86.7	83.3	86.7	76.7	86.7
에/e/	83.3	80.0	83.3	80.0	83.3
이/i/	83.3	83.3	80.0	86.7	83.3
오/o/	83.3	80.0	83.3	86.7	83.3
우/u/	86.7	83.3	83.3	83.3	83.3
평균	84.66	81.98	83.3	82.68	83.98

패턴 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

IV. 결 론

본 연구에서는 음성 어트랙터에 대하여 상관차원 값을 구한 결과, 상관차원값은 패턴인식보다는 화자식별에 적합한 파라미터임을 알 수 있었다.

상관적분에 있어서 시간에 따른 궤도변화의 거리측정시 퍼지유사도를 적용함으로써 퍼지상관차원을 구할 수 있었으며, 퍼지상관차원은 서로 다른 패턴간은 분별시키고, 동일패턴에 대해서는 패턴변동분을 흡수시킴

으로서 음성인식과 화자인식에 적합한 파라미터임을 확인할 수 있었으며 제안된 퍼지상관차원으로 화자인식을 수행한 결과, 퍼지상관차원은 평균 95.34%의 우수한 인식율을 얻을 수 있었다.

참 고 문 헌

- [1] Sadaoki, Furui, Advances in Speech Signal Processing, 1992.
- [2] S. Saito, K. Nakata, Fundamentals of Speech Processing, Academic Press, 1985.
- [3] 合原一幸, 脳どカオス, 東京電氣大學, 1992.
- [4] Thomas S. Parker, Leon O. Chua, "Chaos: A Tutorial for Engineers", Proceeding of the IEEE, vol. 75, no. 8, pp. 982-1008, 1987.
- [5] Alan Wolf, "Determining Lyapunov Exponents from a time series", Physica 16D, pp. 285-317, 1985.
- [6] B. Malraison, P. Atten, P. Brege and M. Dubois, "Dimension of strange attractors", an experimental de Physique Letters, vol. 44, no. 22, pp. 897-902, 1983.
- [7] J. D. Farmer, E. Ott and J. A. Yorke, "The dimension of chaotic attractors", Physical, vol. 7, pp. 153-180, 1983.
- [8] P. Eckmann, S. O. Kanmhorst, D. Ruelle and S. Ciliberto, "Liapunov exponents from a time series", Phys. Rev. A, vol. 6, pp. 4971-4979, 1986.
- [9] P. Frederickson, J. L. Kaplan, E. D. Yorke and J. A. Yorke, "The Lyapunov dimension of strange attractors", j. Diff. Equations, vol. 49, pp. 185-207, 1983.
- [10] D. S. KIM, "Chaotic Analysis of Speech Signals", 한일중 카오스 공동 세미나 논문집, pp. 124-129, 1984
- [11] 유병욱, 김창석, "Lyapunov 차원을 이용한 화자식별 파라미터추정", 한국음향학회지 제 16권 4호, 1987
- [12] Bart Kosko, "Neural Networks and Fuzzy Systems", Prentice Hall, 1991.
- [13] 寺野, 菅野, ファジィ システム入門, オーム社, 1987
- [14] Earl Cox, The Fuzzy systems Hand book, AP, 1994.
- [15] 유병욱, 최승호, "선형예측계수의 주파수 변동에 따른 멤버쉽 함수의 개선", 한국음향학회 음성통신 및 신호처리 워크샵 논문집, 1993

저 자 소 개

俞炳旭(正會員)

1986년 2월 명지대학교 전자공학과 졸업(공학사). 1988년 2월 명지대학교 전자공학과 졸업(공학석사). 1999년 2월 명지대학교 전자공학과 졸업(공학박사). 1997년 ~ 현재 (주)데이터씨큐어 정보기술연구소 주관심분야는 화자인식, 생체신호처리, 정보보호

朴賢淑(正會員)

1990년 2월 전국대학교 대학원 전자공학과 졸업(공학박사). 1992년 ~ 1995년 공업기술원 연구원. 1997년 ~ 현재 동아방송대학 방송기술과 조교수. 주관심분야는 음성인식

金昌錫(正會員)

1985년 2월 명지대학교 대학원 전자공학과 졸업(공학박사). 1985년 3월 ~ 현재 명지대학교 전자공학과 정교수. 1997년 3월 ~ 현재 명지대학교 전자정보통신공학부장. 주관심분야는 음성인식, 화자인식, 생체신호처리