

## Lyapunov 차원을 이용한 화자식별 파라미터 추정

### Estimation of Speaker Recognition Parameter using Lyapunov Dimension

유 병 옥\*, 김 창 석\*  
(Byong-Wook Yoo\*, Chang-Seok Kim\*)

#### 요 약

본 논문에서는 음성을 비선형 결정론적 발생메카니즘에서 발생하는 불규칙한 신호인 카오스로 보고 상관차원과 Lyapunov 차원을 구함으로써 음성화자식별 파라미터와 음성인식파라미터에 대한 성능을 평가하였다. Takens의 매립정리를 이용하여 스트레인지 어트랙터를 구성할 때 AR모델의 파워스펙트럼으로부터 주요주기를 구함으로써 정확한 상관차원과 Lyapunov 차원을 추정하였다. 어트랙터 궤도의 특징을 나타내는 상관차원과 Lyapunov 차원을 가지고 음성인식과 화자인식의 특징파라미터로서의 효용성을 고찰하였다. 그 결과, 음성인식보다는 화자식별의 특징파라미터로 타당하였으며 화자식별 특징파라미터로서는 상관차원보다는 Lyapunov 차원이 높은 화자식별 인식율을 얻을 수 있음을 알았다.

#### ABSTRACT

This paper has appraised ability of speaker recognition and speech recognition using correlation dimension and Lyapunov dimension. In this method, speech was regarded the chaos that the random signal is appeared in deterministic raising system. We deduced exact correlation dimension and Lyapunov dimension with searching important orbit from AR model power spectrum when reconstruct strange attractor using Taken's embedding theory. We considered a usefulness of speech recognition and speaker recognition using correlation dimension and Lyapunov dimension that characterized reconstruction attractor. As a result of consideration, which were of use more the speaker recognition than speech recognition, and in case of speaker recognition using Lyapunov dimension were gain much recognition rate more than speaker recognitions using correlation dimension

#### I. 서 론

음성의 발생메카니즘은 비선형 동적 시스템이므로 저주파수의 성분 특성이 잘 나타나는 선형예측계수만으로는 음성의 특징파라미터로 충분하지 못하므로 음성 신호 처리 알고리즘의 개선과 부가적 특징파라미터가 필요하며, 최근에 음성인식과 화자식별에 적절한 파라미터의 추출에 관한 연구와 비선형인 음성의 발생 메카니즘의 새로운 모델링 접근에 관한 연구가 활발히 진행되었다.<sup>1)</sup>

한편 과거의 자연현상을 이해하여 이를 이용하고자 하는 카오스(chaos) 연구가 많이 연구되고 있다. 결정론적 법칙에 따라 서로 비선형적인 결합으로 인하여 불규칙적이고, 예측불가능하게 보이는 현상이 발생될 때, 그 현상을 카오스라고 한다.<sup>(7-14)</sup>

불규칙 시계열 신호에 대해서도 카오스 분석을 통하여 신호의 복잡한 변화 속에 숨겨져 있는 동적 정보를 파악

함으로써, 신호를 정량화하여 공학에 활용하려는 연구가 활발히 진행중이다. 음성신호의 발생 메카니즘은 선형 시스템이 아니라 매우 복잡하게 비선형인 결정론적 동적 시스템이므로 정량화된 동적정보의 특징을 나타내는 카오스에 대한 차원을 음성인식에 적용하고자하는 연구보고도 있다.<sup>(15)</sup> 음성은 개인 마다 서로 상이한 발생습관과 성도의 구조차이에 따라 주파수변동과 시간의 변동을 나타내므로 음성인식은 서로 다른 음성을 변별하고, 개인간의 주파수와 시간의 변동을 흡수하는 문제가 중요하다. 그러나 카오스에 대한 차원은 신호의 궤적에 대한 결정론적 규칙 즉, 발생규칙을 찾고 그것에 대한 정량화된 척도를 뜻하므로 개개인의 다른 음성 발생 메카니즘에 대해 개인간의 성도구조의 차이, 즉 동적시스템에 대한 규칙정도의 차이를 알아내는데 적합하므로 개인간의 변동을 흡수하는 음성인식 보다는 화자인식 파라미터로 유용한 것으로 생각된다.

음성신호가 카오스 신호이던 음성은 결정론적 규칙에 의하여 지배된다. 이러한 규칙은 카오스 운동을 특징 지우는 스트레인지 어트랙터(strange attractor)를 찾음으로

\*명지대학교 전자공학과  
접수일자: 1997년 2월 24일

써 가능하다. 정확한 스트레인지 어트랙터를 구성하기 위해 AR모델에 의해 구해진 파워 스펙트럼으로부터 음성 신호의 주요 주기를 찾고, 스트레인지 어트랙터를 구성함으로써 정확한 상관차원과 Lyapunov 차원의 값이 추정되도록 하였다. 구성된 어트랙터 궤도의 동적 특성을 특징지우는 상관차원과 Lyapunov 차원이 개인차가 심한 음성신호에 대해 개인의 역학계의 특징을 매우 잘 나타내어 줄 것이라는 점에 착안하여 상관차원과 Lyapunov 차원을 추정하고, 평가하여 음성의 화자식별의 특징파라미터로서의 타당성을 고찰하고자 한다.

## II. 상관차원과 Lyapunov 차원

카오스의 발생 시스템은 시계열 데이터로서 관측할 경우 그 동작은 복잡하고 초기값에 대해 매우 민감하게 변하는 불안정한 궤도를 나타낸다. 이러한 불안정한 궤도의 정량적 특징을 알아보기 위해서는 데이터를 공간에 재구성한 어트랙터를 관측하여야 하며, 어트랙터를 구성하기 위해서 Takens의 매립정리를 이용한다.

음성시계열을  $x_1, x_2, \dots, x_N$ 이라할 때, Takens의 매립정리에 의하여 식 (1)과 같이 벡터시계열을 구성할 수 있다.<sup>(1-4)</sup>

$$\begin{aligned} X_1 &= (x_1, x_{1+\tau}, \dots, x_{1+(m-1)\tau}) \\ X_2 &= (x_2, x_{2+\tau}, \dots, x_{2+(m-1)\tau}) \\ &\vdots \\ X_t &= (x_t, x_{t+\tau}, \dots, x_{t+(m-1)\tau}) \end{aligned} \quad (1)$$

여기서  $t = N - (m-1)\tau$ 이다. 위 식에서  $\tau$ 는 지연시간으로서 주된 주기값보다 작은 값으로 설정하여야 하며, 어트랙터의 구성시  $\tau$ 가 너무 적으면 재구성상태의 공간내의 데이터는 극단적으로 상관성이 너무 크게 된다.

$\tau$ 의 최적 결정방법은  $\tau$ 를 관측된 시계열 데이터  $X_t$ 의 주요한 주기의 수분의 1로 선택하는 방법과  $X_t$ 의 자기상관함수  $R$ 이 최초로 0이 되는 시각, 즉 주요 주기의 1/2이 되는 시각으로 선택하는 방법, 그리고  $X_t$ 의 상호정보량  $I$ 가 최초로 극소치를 취하는 시각을 선택하는 방법등이 사용되고 있다. 그러나 음성신호가 관측시계열 데이터인 경우 주요 주파수 성분이 존재하며 또한, 자음과 모음의 주파수 차이는 매우 크게 존재하므로 위와 같은 세가지  $\tau$ 의 결정방법으로 어트랙터를 구성하면, 부적절한 어트랙터가 구성되고 따라서 부적절한 카오스 차원의 추정이 이뤄진다. 따라서 자음과 모음이 섞여 있는 단어의 경우는 자음과 모음에 대한 최소 2개 이상의 주된 주기의 고려가 필요하다. 본 연구에서는 모음만을 음성신호로 하였으나 모음에도 여러 주파수 성분을 갖을 수 있다고 가정하여 음성신호에 대해 AR모델에 의해 파워스펙트럼을 추정후, 가장 진폭이 큰 주파수 성분을 그 음성신호의 주요 주기로 하여 지연시간  $\tau$ 를 결정하였다.

표 1은 관측된 음성 시계열 데이터 각 5개 모음에 대해  $X_t$ 의 자기상관계수가 최초로 0이 되는 시각  $\tau'$ 와 파워스펙트럼을 추정하여 진폭이 가장 큰 2개의 주파수 성분에 대한 주기를  $\tau_1$ 과  $\tau_2$ 로 하여 비교한 것이다. 표를 보면  $\tau'$ 의 값이 모음 '아'에 대해 4로서 주요주기는 샘플수  $\times$  샘플주기 = 4 msec 로 추정되며 주파수 분석하여 결정한  $\tau_1$ 과  $\tau_2$ 와 큰 차이가 있다. 따라서 음성신호에 대한 카오스 분석시 정확한 어트랙터를 구성함에 있어서 지연시간은 자기상관계수로 구하는 것 보다 음성 성분에 대한 주파수 분석후 주된 주기의 추정이 올바른 지연시간 추정이라고 사료된다.

표 1. 지연시간  $\tau$ 과  $\tau'$ 의 추정 비교.

지연시간 \ 모음별		아	애	이	오	우
		$\tau'$	2	3	2	2
$\tau$	$\tau_1$	8	10	10	8	8
	$\tau_2$	7	8	8	7	7

### 1. 상관 차원

상관차원은 어트랙터 궤도상에서 재구성된 데이터와 데이터 사이의 상관관계를 계산하여 구한 변화율이다. 상관관계란 재구성된 데이터와 데이터 사이의 거리, 즉, 임의의 한 데이터값을 중심으로 일정한 반경 안에 존재하는 데이터의 수를 계산하여 구할 수 있다. 이러한 상관관계의 계산으로 일정한 기울기의 변화율을 얻을 수 있으며 이 일정한 기울기의 변화율은 신호의 카오스적 특징, 즉 규칙을 나타내게 된다. 또한 신호의 발생 시스템의 카오스적 특징을 나타내기도 한다.  $m$ 차원 공간에 있어서 재구성된 어트랙터상의 한 점을  $X_i = R^m$ 이라 하면 상관적분  $C^m(\epsilon)$ 은 다음과 같이 정의된다.<sup>(2-6)</sup>

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

$$H(\epsilon - |X_i - X_j|) = \begin{cases} 1, & (\epsilon - |X_i - X_j|) \geq 0 \\ 0, & (\epsilon - |X_i - X_j|) < 0 \end{cases} \quad (3)$$

어트랙터상의 한 점  $X_i (i=1, 2, \dots, N)$ 을 중심으로 나머지  $(N-1)$ 개의 점에 대해 반경  $\epsilon$ 의  $m$ 차원 구내로 들어가는 점  $X_j$ 를 계산하고, 모든  $X_i$ 에 대해서 반복 계산한다. 위 식 (2)에서 상관적분이 임의의  $\epsilon$ 의 한 영역에서 선행적으로 비례관계가 존재한다면 어트랙터상의 선행 비례 범위 위  $\epsilon_1 < \epsilon < \epsilon_2$ 에서 자기상사성을 갖는다고 할 수 있다.

$$C^m(\epsilon) \propto \epsilon^{l(m)} \quad (4)$$

$$Dc = \lim_{\epsilon \rightarrow \infty} \frac{\ln C^m(\epsilon)}{l(m)} \quad (5)$$

식(4)에서 공간차원  $m$ 을 증가시키므로써 식(5)와 같이  $l(m)$ 이 포화되어 점근하는 값  $D_c$ 를 상관차원이라 하고, 이는 공간차원  $m$ 이 증가하여 포화 상태가 되는  $l(m)$  값이 각 신호의 발생 시스템에 따라 다른 특징이 될 수 있다.<sup>(20)</sup>

2. Lyapunov 차원<sup>(2-6, 17-19)</sup>

상관차원은 상관분석과 상관지수와의 관계에서, 기율기를 측정하기 위해서는 직선부분의 결정에 따른 오차와 대단히 많은 데이터수가 요구되는 등의 단점을 가지고 있다. 음성은 카오스적 특성을 갖는 비선형 동적 시스템에서 나오는 불규칙하게 보이는 현상이라 할 수 있으므로 음성 발생 시스템의 특징은 스트레인지 어트랙터 궤도의 기하학적 구조와 질서에 대한 정량적 기준이 되는 Lyapunov 차원으로 추정할 수 있다.

시계열상에서의  $n$  차원 이산신호 발생 시스템은 다음 식으로 정의된다.

$$x_{t+1} = F(x_t, \mu), x_t \in R^n \tag{6}$$

여기서  $x_t$ 는 시간  $t$ 에서의 상태이며,  $\mu$ 는 파라미터 벡터이고,  $F$ 는  $n$  차원 사상이다.  $n$  차원 공간에 있어서  $\delta x_0$ 를 중심으로 서로 직교하는 단위 벡터의 조합을  $u_1(t), u_2(t), \dots, u_n(t)$ 로 하고  $DF(x_t)$ 를 Jacobian행렬이라 하면 각 벡터의 변화는 다음과 같다.

$$e_i(t+1) = DF(x_t) u_i(t) \tag{7}$$

여기서 식(6)는  $DF(x_t)$ 의 가장 적은 고유치에 대응하는 고유 벡터 방향으로 값이 무한히 작아지기 때문에 이것을 피하기 위해서 식(6)을 직교화한  $e'_i(t+1)$ 를 구하여 식(9)과 같이 새로운 정규직교계  $u_i(t+1)$ , ( $i=1, 2, \dots, n$ )로 변환한다.

$$e'_i(t+1) = e_i(t+1) - \sum_{j=1}^{i-1} \langle e_i(t+1), u_j(t+1) \rangle u_j(t+1) \tag{8}$$

$$u_i(t+1) = \frac{e'_i(t+1)}{|e'_i(t+1)|} \tag{9}$$

여기서  $\langle \cdot, \cdot \rangle$ 는 내적이며 각 벡터  $u_i(t+1)$ , ( $i=1, 2, \dots, n$ )을 식(8), (9)에 의해 다시 사상하는 과정을 반복함으로써 Lyapunov 스펙트럼  $\lambda_i$  ( $i=1, 2, \dots, n$ )를 다음 식과 같이 추정할 수 있다.

$$\lambda_i = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=0}^{N-1} \log |e'_i(t)| \tag{10}$$

식(10)에서 구해진  $\lambda_i$ 에서 각각의 Lyapunov 지수를  $i$  순으로 나열한 것이 Lyapunov 스펙트럼이며, 음이 아닌

최대의 정수를 갖는 순서를  $k$ 로 할 때 Lyapunov 차원  $D_L$ 은 다음 식으로 구해진다.

$$D_L = k + \frac{\sum_{i=1}^k \lambda_i}{|\lambda_{k+1}|} \tag{11}$$

III. 음성 데이터의 Lyapunov 차원 추정

실험에 의해 얻어진 음성 시계열 데이터에서는  $DF(x_t)$ 를 직접 알 수 없으므로 실험 데이터에서 미소변위 벡터를 부여할 수는 없지만, 재구성 상태공간내 어트랙터상의 어느 한점에 주목하면, 그 근방에는 다른 점이 존재한다. 따라서 이 점과 그 근방점을 선택하여 미소변위벡터로 간주할 수도 있으며, 이들 점들이 어느 정도 변화했는지를 평가하여 Jacobian 행렬  $F$ 를 추정하고, Lyapunov 스펙트럼을 구한다. Takens 매립정리로 구성된 어트랙터 궤도상의 한 점  $X_i$ 에서 본 미소반경  $\epsilon$ 로 하는 구내에서  $M$ 개의 점  $X_{ki}$ 에 대한 변위벡터  $y_i$ 는

$$y_i = X_{ki} - X_i, y_i \in R^m \tag{12}$$

이고, 미소변위벡터  $\delta x_t$ 를  $y_i$ 로 생각 할 수 있다. 시간이  $s$ 만큼 지난 후, 미소반경  $\epsilon$ 로 하는 구의 중심  $X_i$ 는  $X_{i+s}$ 로, 각 상태  $X_{ki}$ 는  $X_{ki+s}$ 로 각각 변화하므로 시간  $t+s$ 에서의 변위벡터  $z_i$ 는 다음 식과 같다.

$$z_i = X_{ki+s} - X_{i+s}, z_i \in R^m \tag{13}$$

미소반경  $\epsilon$ 로 하는 구의 반경과 시간  $s$ 가 충분히 작다면 식(12), (13)의  $y_i$ 와  $z_i$ 는 선형관계가 성립될 수 있으며, 임의의 행렬  $G_i$ 를 가지고서 다음 식과 같이 나타낼 수가 있다.

$$z_i = G_i y_i \tag{14}$$

행렬  $G_i$ 는 식(7)의 Jacobian 행렬에 근사적이라고 생각할 수 있으므로 다음 식의 거리  $s$ 를 최소로 하는 최소자승법으로 결정할 수 있다.

$$s = \sum_{i=1}^M |z_i - G_i y_i|^2 \tag{15}$$

행렬  $G_i$ 의 제  $k_1$  성분을  $g_{k_1}$ 라고 하면 식(15)의 거리  $s$ 에 대한 각각의  $g_{k_1}$ 에 대한 극소조건,  $\frac{\partial S}{\partial g_{k_1}} = 0$ 에서 다음 식을 얻을 수 있다.

$$G_i V = C \tag{16}$$

$V, C$ 는  $m \times m$ 의 행렬로서 분산·공분산 행렬이며, 행렬

G<sub>i</sub>를 구하여 DF로 대신하여 음성 데이터의 Lyapunov 차원을 구한다.

IV. 실험 및 결과

음성데이터 수집은 주변잡음이 존재하는 실험실에서 잡음을 최소화 한 상태로 20대 남녀 화자 각각 5명씩으로 구성 하여 1명당 10회씩 단어를 발성하여 발성된 음성에 대해 샘플링 주파수를 10 KHz로 하여 A/D 변환하여 음성 시료를 수집하였다.

본 연구에서 사용된 음성 시료는 상관차원과 Lyapunov 차원을 화자식별에 대한 파라미터로써 평가를 하기 위해 모음만을 고려하였다.

3명의 화자가 모음5개 아, 에, 이, 오, 우를 10회씩 발성한 150개의 음성시료 중 1회 발성한 음성을 표준패턴으로 나머지 음성을 시험패턴으로 실험하였다.

그림 1은 모음 '아'에 대한 각 화자의 어트랙터이다. 그림을 보면 동일한 모음이라도 어트랙터 궤도형태는 각 화자에 대해 공통적인 형태를 보이지 않고 있다.

그것은 같은 모음을 발성해도 개개인간의 발성습관의 차이와 주파수 정보의 차이 즉, 각 화자의 비선형적인 발성기관의 차이에 원인이 있다. 따라서 그림 1의 음성 시계열 신호의 어트랙터 모양은 모음 '아'의 특징도 나타내고 있으나 각 화자의 음원의 특징도 나타내고 있다고 생각할 수 있다.

음성인식의 문제는 각 화자의 모음 '아'에 대한 발성시간과 주파수 차이의 특징 흡수의 문제이므로 어트랙터의 궤도 모양의 특징을 정량적으로 나타내는 상관차원, Lyapunov 차원을 음성인식의 파라미터로 사용하기에는 부적합하다. 그러나 상관차원과 Lyapunov 차원이 주어진 한 음성 또는 여러 음성에 대해서는 화자의 특징을 궤도의 형태에 대한 특징으로 나타내므로 화자식별의 파라미터로 사용하는 것이 가능할 것으로 생각된다. 따라서 본 연구에서는 각기 다른 형태의 어트랙터가 각 화자의 비선형적인 발성기관의 차이에서 오는 것으로 보고, 그

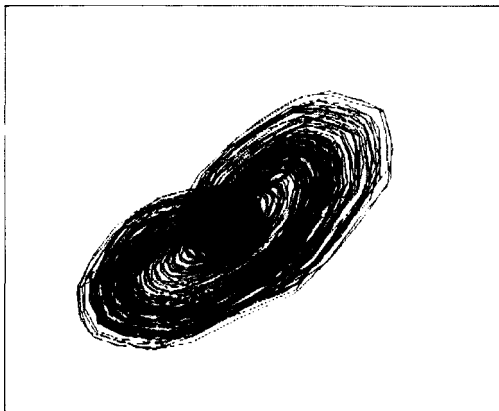


그림 1. 화자 A에 대한 '아'의 어트랙터

차이가 화자간의 특징을 포함하고 있다고 가정하여 어트랙터 궤도의 특징을 정량적으로 평가하여 주는 상관차원과 Lyapunov 차원을 구해 화자식별을 행하였다.

1. 상관차원에 의한 화자식별

상관차원은 식 (5)를 이용하여 구할 수 있으며, 식 (5)에서의 상관적분 C<sup>m</sup>(ε)은 임의의 X<sub>i</sub>와 X<sub>i+1</sub>의 유클리안 거리에 대해 수렴될 때까지 반경 ε을 미소량으로 증가시키면서 각각에 따른 식(2)로부터 구할 수 있다. 상관차원과 상관적분 간의 관계를 고찰하기 위해 차원 m을 증가시키면서 l(m)을 구해본 결과 m=5 이상에서 모든 모음이 일정한 점근값을 나타냈다. m의 증가에 따라 l(m)이 포화되어 점근하는 값 Dc를 구할 수 있었으며 각각의 모음에 대해서도 구하였다. 점근값 Dc가 상관차원이며 각 화자별 모음별 결과를 표 2에 나타내었다.

표 2. 상관차원의 화자별 평균값

화자	상관차원의 평균값					
	아	에	이	오	우	분포
A	4.629	4.760	4.604	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
분포	4.2~4.6	4.4~4.7	4.3~4.6	4.2~4.7	4.1~4.8	

표 2에서 화자의 구별없이 모음에 따라 상관차원의 분포를 나타낸 아래의 분포를 살펴보면, 「아」에 대한 상관차원의 값이 4.2~4.6으로 「에」의 4.4~4.7과 겹치는 분포를 갖으며, 다른 모음에 대해서도 모음간의 식별이 곤란한 것을 알 수 있으며 음성인식의 파라미터로서 부적합한 파라미터로 보인다. 화자 간의 차이는 존재하는가를 알아보기 위해 오른쪽의 분포를 살펴보면, A화자의 상관차원의 분포는 4.6~4.8, B화자의 상관차원의 분포는 4.2~4.5로서 화자 간 차이를 보이고 있으므로 화자식별이 더 용이함을 알 수 있다.

구해진 상관차원값에 대해 1회 발성한 A화자의 5개 모음을 표준패턴으로하여 각 A, B, C화자의 2, 3, 4, 5회 발성한 모음을 시험패턴으로 하여 다음 식 (17)과 같이 정의한 상관차원에 대한 식별오차를 구하였다.

식별오차율

$$= \frac{1}{4} \sum_{i=0}^{i=4} \left| \frac{\text{표준패턴의 차원} - \text{시험패턴의 차원}}{\text{표준패턴의 차원}} \times 100 \right| \quad (17)$$

여기서 i는 음성 발성 수이다. 식 (17)로 평균 식별오차를 구하여 표 3에 나타내었다. 표 3를 보면 A화자 5개 모음에 대한 식별오차율은 평균 6.5%로서 A화자의 인식율이 매우 높음을 알 수 있으며 화자식별 파라미터로 상관차원값을 이용하는 것에 대해 타당하다는 것을 알 수 있다. 한편

표 3. 표준패턴 A화자에 대한 식별오차율

표준패턴 \ 화자	A					B					C				
	아	에	이	오	우	아	에	이	오	우	아	에	이	오	우
아	5.10	28.82	15.91	17.30	30.08	15.99	30.36	40.00	18.30	23.33	9.13	13.55	37.67	35.26	27.51
에	27.23	4.50	21.52	18.99	25.77	26.91	11.90	22.17	16.45	30.13	8.88	25.82	24.81	17.68	26.59
이	17.84	33.00	5.98	27.27	26.53	24.68	17.28	17.16	16.36	20.18	26.59	18.37	7.98	40.04	26.53
오	34.43	23.66	18.50	8.53	11.44	27.37	26.88	19.60	14.95	14.74	30.94	26.11	14.24	15.60	19.56
우	22.02	36.63	41.54	16.92	8.39	26.57	37.08	16.49	18.28	13.40	26.42	18.48	29.64	27.64	23.40
평균	6.50					14.68					16.38				

상관차원에 대한 음성인식 파라미터로의 타당성을 알아보기 위해 1회 발성된 모음을 표준패턴으로 각 화자별 모음에 대한 평균식별 오차율을 구하여 표 4에 나타내었다.

표 4. 각 모음에 대한 평균 식별오차율

화자 \ 모음	아	에	이	오	우
A	15.97	25.76	20.33	17.50	18.09
B	17.31	18.78	27.35	19.28	19.00
C	12.23	13.93	13.57	18.18	16.29
평균	15.17	19.46	20.42	18.32	17.79

표 4에서 보는 것과 같이 각 모음에 대한 평균 식별 오차율이 매우 크다. 그것은 표 2과 같이 각 모음의 상관차원의 값이 서로 겹치는 분포를 가지므로 모음간 식별이 곤란하고, 카오스에 대한 차원이 어트랙터궤도 형태의 흡수이기 보다는 특징을 적도로 나타내어 궤도형태를 분류하므로 표 4의 평균 식별오차가 크게 나타나는 것으로 보이며, 카오스의 차원은 음성인식보다는 화자식별에 더 유용한 파라미터임을 알 수 있다. 표 3에서 식별오차율이 가장 적은 것을 각 화자마다 1회 발성된 모음을 표준패턴으로 하여 식별오차율이 가장 적은 것을 인식하도록 실험을 행한 결과, 표 5와 같은 화자 인식율을 얻을 수 있었다.

표 5. 상관차원에 의한 화자 인식율

화자 \ 인식율	인식율 (%)
A	92.6
B	95.5
C	87.4
평균	91.8

표 5를 보면 평균 91.8%의 높은 인식율을 얻을 수 있었다. 그러나 화자수가 증가하면 상관 차원은 많은 데이터 처리 양이 요구되고 자기상사성을 갖기 위한 선형적 비례범위가 불안정해질 수 있다. 따라서 화자식별에 제한

적이라고 할 수 있으므로 어트랙터의 정확한 차원 추정 이 가능한 Lyapunov 차원을 구하여 보았다.

2. Lyapunov 차원에 의한 화자식별

식 (9)에 의한 Lyapunov 지수를 추정하기 위해 전 절에서 설명한 Takens의 매립정리를 적용하고 AR모델에 의해 음성 주요주기  $\tau$ 를 추정한 다음 식 (15)의  $G_1$ 를 DF로 하여 식 (9)를 계산하여 Lyapunov 지수를 구하였다. 구해진 Lyapunov 지수 중 큰 순서대로 바꿔 가며 음이 아닌 최대 정수를 k로 하여 식 (11)을 이용하여 Lyapunov 차원을 추정하였으며 표 6에 화자별로 Lyapunov 차원값을 추정한 결과를 각 발성횟수에 대해 평균하여 나타내었다.

표 6. Lyapunov 차원의 화자별 평균값

화자 \ 모음	Lyapunov의 평균					분포
	아	에	이	오	우	
A	2.523	2.592	2.404	2.090	2.444	2.0~2.5
B	2.756	2.901	2.699	2.507	2.873	2.7~2.9
C	2.602	2.774	2.479	2.415	2.619	2.4~2.7
분포	2.5~2.7	2.5~2.9	2.4~2.7	2.0~2.5	2.4~2.8	

표 6은 공간차원 m이 3일 때 각 화자별 모음에 따른 Lyapunov 차원의 평균값이며 Lyapunov 차원값이 음성 인식의 파라미터로서 유용한가를 알아보기 위해 화자의 구별없이 모음간의 차이를 고찰할 때 Lyapunov 차원값의 분포로는 각 모음을 식별하기가 곤란하다. 이로 부터 Lyapunov 차원값으로도 음성인식의 파라미터로는 부적당하다는 것을 알 수 있다. 화자인식의 파라미터로서의 효용성을 검토해보면, A화자의 모음 「아」의 Lyapunov 차원값과 B화자의 모음 「아」의 Lyapunov 차원값 2.523과 2.756으로 차이가 있음 알 수 있고, 다른 모음에 대해서도 역시 화자간의 차이는 알 수 있다. 이는 Lyapunov 차원값이 개인간 차와 발성된 모음간 차를 나타내는 것으로 생각되며, Lyapunov 차원은 화자식별의 파라미터로서 유용하다고 할 수 있다. 값의 분포형태에 대해 단어 군과 인식 화자대상이 커지면 인식이 어려울 것으로 생각되지만 시간에 따라 차원의 값이 변화하는 량을 개인

마다 식별오차율을 추정함으로써 카오스 차원의 단점을 보완할 수 있다고 사료된다.

따라서 음성인식과 화자식별의 유용성을 알아보기 위해 A 화자의 각 모음중 1회 발생한 음성을 표준패턴으로 정하고 나머지 각 음성에 대해 Lyapunov 차원을 구하여 화자별로 인식된평균 식별오차율을 구하여 표 7에 나타내었다.

표 7을 보면 평균 4.70%로 표 3에서의 상관차원보다는 식별오차율이 적음을 알 수 있으며 표 6에서도 상관차원의 표 1의 화자별 평균값보다 분포가 화자별로 뚜렷한 특징을 갖고 있음을 고찰할 수 있다. 구해진 화자식별오차율에 대해 최소 식별율을 보이는 화자의 모음을 인식하도록 실험하였다.

표 7. Lyapunov 차원에 대한 평균 식별오차율

화자 표준패턴	A	B	C
아	3.33	24.90	13.90
에	3.84	17.37	24.06
이	4.78	20.28	12.31
오	6.56	12.42	15.55
우	5.01	26.88	18.53
평균	4.70	20.35	16.87

표 8. Lyapunov 차원에 의한 화자 인식율

화자	인식율 (%)
A	97.7
B	97.0
C	97.0
평균	97.2

각 화자에 대해 Lyapunov 차원값에 의한 화자 식별을 행한 결과를 표 8에 나타내었다. 97.2% 인식율을 보이고 있는 것으로 보아 Lyapunov 차원값이 화자식별에 매우 우수한 파라미터로 생각되어 진다. 한편 상관차원과 화자인식 결과와 비교하기 위해 표 9에 나타내었다.

표 9를 보면 Lyapunov 차원을 사용한 화자인식율이 상관차원을 사용한 화자인식율보다 향상된 인식율을 나타

표 9. 상관차원과 Lyapunov 차원값을 사용한 화자인식율의 비교

화자	인식율	
	상관차원	Lyapunov 차원
A	92.6	97.7
B	95.5	97.0
C	87.4	97.0
Average	91.8	97.2

냄을 알 수 있다. 이는 Lyapunov 차원이 어트랙터의 특징을 더욱 잘 나타내며, 또한 발생시스템 자체의 특징을 잘 나타낸다고 할 수 있다.

상관차원과 Lyapunov 차원은 화자식별에 매우 효과적인 파라미터이며 음성인식에도 여러 가지 부수적인 파라미터로 활용될 수 있을 것으로 사료된다.

## V. 결 론

본 논문에서는 음성신호에 비선형적 분석방법인 카오스 차원 중 상관차원과 Lyapunov 차원을 추정하여 화자인식방법을 제안하여 실행한 결과 다음과 같은 결과를 얻었다.

첫째, 음성데이터에 대해 어트랙터를 구성한 결과 음성신호는 카오스 신호임을 알 수 있었고, 재구성된 스트레인지 어트랙터로부터 음성신호처리의 특징파라미터로 사용하기 위해 어트랙터의 궤도적 특징을 정량적으로 나타내주는 상관차원과 Lyapunov 차원을 추정하여 음성인식이나 화자인식의 특징파라미터로서의 효용성을 고찰한 결과 음성인식보다는 화자인식의 특징파라미터로서 타당함을 확인할 수 있었다.

둘째, 상관차원과 Lyapunov 차원을 가지고 화자인식 실험을 해본 결과 상관차원보다는 Lyapunov 차원이 우수하였으며 높은 인식율을 얻을 수 있었다.

## 참 고 문 헌

1. Sadaoki Furui, Advances in Speech Signal Processing. 1992.
2. Thomas S. Parker, Leon O. Chua, "Chaos: A Tutorial for Engineers", Proceeding of the IEEE, Vol. 75, No. 8, pp. 982-1008, 1987.
3. Alan Wolf, "Determining Lyapunov Exponents from a time series", Physica 16D, pp. 285-317, 1985.
4. D.S. KIM, "Chaotic Analysis of Speech Signals", 한·일·중 카오스 공동 세미나 논문집, pp. 124-129, 1994.
5. 合原一幸, 脳とカオス, 東京電氣大學, 1992.
6. J. P. Eckmann, S. O. Kanmhorst, D. Ruelle and S. Ciliberto, "Lyapunov exponents from a time series", Phys. Rev. A, Vol. 6, pp. 4971-4979, 1986.
7. T. Y. Li, J. A. Yorke, "Period three implies chaos", Amer. Math. Monthly, Vol. 82, pp. 985-992, 1975.
8. R. may., "Simple mathematical models with very complicated dynamics", Nature, Vol. 261, pp. 459-467, 1976.
9. P. Grassberger, I. Procaccia, "Measuring strangeness of strange attractors", Physica 9D, pp. 189-208, 1983.
10. M. Sano, Y. Sawada, "Measurement of the Lyapunov spectrum from a chaotic time series", Phys. Rev. Lett., Vol. 55, pp. 1082-1085, 1985.
11. A. M. Rraser., L. Swinney, "Independent coordinates for strange attractors from mutual information", Phys. Rev. A, Vol. 33, pp. 1134-1140, 1986.

12. F. Takens, "Detecting Strange Attractors in Turbulence, in Dynamical Systems and Turbulence", Lecture Notes in Mathematics, Vol. 898, pp. 366-381, Springer, 1981.
13. B. Maltraison, P. Atten, P. Brege and M. Dubois, "Dimension of strange attractors", an experimental de Physique Lettres, Vol. 44, No. 22, pp. 897-902, 1983.
14. J. Theiler, "Spurious dimension from correlation algorithms applied to limited time-series data", Phys. Rev. A, Vol. 34, No. 3, pp. 2427-2432, 1986.
15. J. P. Eckmann and D. Ruelle, "Ergodic theory of chaos and strange attractors", Rev. Mod. Phys., Vol. 57, No. 3, pp. 617-656, 1985.
16. I. Shimada, T. Nagashima, "A Numerical Approach to Ergodic Problem of Dissipative Dynamical Systems", Prog. Theor. Phys., Vol. 61, No. 6, pp. 1665-1616, 1979.
17. J. D. Farmer, E. Ott and J. A. Yorke, "The dimension of chaotic attractors", Physical, Vol. 7, pp. 153-180, 1983.
18. P. Frederickson, J. L. Kaplan, E. D. Yorke and J. A. Yorke, "The Liapunov dimension of strange attractors", j. Diff. Equations, Vol. 49, pp. 185-207, 1983.
19. 김자훈, 유병욱, 김창석, "Lyapunov 차원을 이용한 화자인식", 제9회 신호처리합동학술대회 논문집, Vol. 9, pp. 385-388, 1996.

#### ▲유 병 욱(Byong-Wook Yoo)



1986년 2월: 명지대학교 공과대학 전  
자공학과 졸업(공학사)  
1988년 2월: 명지대학교 대학원 전자  
공학과 졸업(공학석사)  
1992년~현재: 명지대학교 대학원 전  
자공학과 박사과정  
※주관심분야: 음성인식, 정보보호,  
카오스, EEG해석

#### ▲김 창 석(Chang-Seok Kim)

제15권 4호 참조

1985년 3월~현재: 명지대학교 공과대학 전자공학과 교수