

## 직교인자의 동적특성을 이용한 화자인식

正會員 裴 澈 秀\*

### Speaker Recognition Using Dynamic Time Variation of Orthogonal Parameters

Chul Soo Bai *Regular Member*

#### 要 約

음성신호의 분석으로부터 유도되는 직교인자는 화자의 개인성을 많이 포함하고 있으므로, 최근 많은 연구자들이 이것을 이용한 통계적 처리방법으로 화자인식을 수행하여 좋은 화자인식율을 얻고 있다. 그러나 이러한 방법들은 아직 음성의 발성속도나 시간적 동특성으로 인해서 발생하는 문제점을 갖고 있다.

따라서 본 논문에서는 이러한 문제점을 해결하기 위해서 음성분석의 한 방법인 Karhunen-Loeve 직교 변환에 의해서 추출한 직교인자를 화자인식에 이용하는 방법에 DTW법을 결합하는 두가지 기법을 제안하였다. 첫째는 직교인자를 특징벡터로 하여 DTW법을 적용하고, 둘째는 직교인자를 최적경로에 이용하는 기법이다. 이들 두 기법에 의한 화자인식 결과와 직교인자의 통계적 처리에 의한 종래의 화자인식방법의 결과를 비교하였다. 사용된 직교인자는 음성신호에서 선형예측계수와 부분자기상관계수를 각각 추출하여 위의 화자인식방법에 각각 적용하였다. 이를 실험한 결과, 선형예측계수로부터 얻은 직교인자를 최적경로를 이용한 기법에 적용하는 경우 88.6%의 가장 높은 인식율을 얻었다.

#### ABSTRACT

Recently, many researchers have found that the speaker recognition rate is high when they perform the speaker recognition using statistical processing method of orthogonal parameters, which are derived from the analysis of speech signal and contain much of the speaker's identity. This method, however, has problems caused by vocalization speed or time varying feature of speed.

Thus, to solve these problems, this paper proposes two methods of speaker recognition which combine DTW algorithm with the method using orthogonal parameters extracted from Karhunen-Loève Transform(KLT) method which applies orthogonal parameters as feature vector to DTW algorithm and the other is the method which applies orthogonal parameters to the optimal path. In addition, we compare speaker recognition rate obtained from the proposed two method with that from the conventional method of statistical process of orthogonal parameters. Orthogonal parameters used in this paper are derived from both linear prediction coefficients and partial correlation coefficients of speech signal.

\*關東大學校 電子通信科  
Dept. of Elec. Comm. Kwan Dong Univ.  
論文番號: 92-99 (接受1992. 5. 25)

As a result of each experiment, it is shown that a high recognition rate of 88.6% is obtained when orthogonal parameters from linear prediction coefficients are applied to the path method.

## 1. 서 론

화자인식이란 음성의 특징을 토대로 사람들을 구별해내는 작업을 말한다. 1945년 미국의 Bell 연구소에서 관찰에 의한 화자인식을 위하여 성문(voice print)을 자동적으로 추출하는 sound spectrograph가 발명되면서 부터 보다 폭 넓은 자동화자인식에 대한 연구가 시작되었다.

1963년 Pruzansky [1]는 17차 필터군을 이용하여 화자인식을 하였으며, Furui [2]는 12차 PARCOR 계수와 Pitch를 이용하여 화자인식을 하였다. 또한 1976년 Samber [3]는 각각 LPC, PARCOR 및 Log Area Ratio의 계수로 부터 통계적으로 추출한 직교인자들의 기리를 구하여 화자인식의 실험을 수행하여 직교인자가 음운정보보다 개인성을 더 많이 포함하고 있음을 확인하였으며, 1977년 Markel [4]은 평균값을 이용하여 긴 기간 동안의 화자인식을 시작하였다.

음성의 개인성 정보인 음향적 관련량에 대해서는 아직 충분히 해명되고 있지는 않지만 음성신호에 개인성 정보가 생기는 것에 대해서는 크게 두가지 요인으로 분석되고 있다. 음성생성에 주요한 영향을 미치는 음성기관의 해부학적 구조의 차이로 인한 선천적인 것과 언어의 습득과정에서 생기는 발생습관인 음성기관을 움직이는 모양의 차이로 인한 후천적인 것이 있다. 선천적인 요인은 성도의 형태가 화자에 따라 다르기 때문에 성도 전체의 공진특성과 음원의 pitch에 개인성이 나타난다. 그리고 후천적인 요인은 발생습관의 차이 이므로 주로 시간적 변화에 따라 개인차를 나타낸다. 그 대표적인 예로는 인터내이션(internation), 액센트(accent) 및 파열음에 있어서 기적음의 유무와 그 길이 등이 차이가 있다. 이러한 요인을 유출하여 화자를 인식하려는 연구분야는 생각에 의한 방법과 시각에 의한 방법, 기계적 수단에 의한 방법이 있다. 1970년대에 들어서면서 LSI와 음성정보 처리기술이 진보되면서 컴퓨터에 의한 화자인식 연구가 더욱 활발히 연구되고 있다. [47]

이러한 직교인자를 이용하여 화자인식을 수행한 연구는 1976년 Sambur[8], 1981년 Bonger [9]등에

의해 이루어 졌으며, 이들은 음성 전 구간에 대한 직교인자의 평균치와 고유치(통계적 특성)를 이용한 통계적 처리방법으로 화자인식을 수행하여 90% 이상의 인식율을 얻었다. 그러나 음성 신호의 프레임 단위로 나타나는 직교인자 벡터의 차이퀀스는 발생 속도나 발생시간에 따라 불규칙하게 나타나는 특성(동특성)이 있으므로 통계적인 처리 방법만으로는 음성의 동특성을 놓칠수가 있으며, 같은 화자일지라도 음성을 발생하는 기간이 다르게 되면 음성신호도 부터의 개인성 정보가 변화하게 되어 정확한 화자인식이 어려워진다.[10-12]

이러한 문제를 해결하기 위하여 본 연구에서는 개인성 정보를 효과적으로 추출할 수 있는 직교인자와 동시시간특성을 이용할 수 있는 DTW 법을 화자인식에 적용하기 위하여 두가지 기법을 제안하였다. 첫째 기법은 직교인자를 특징벡터로하여 DTW 법에 적용하는 것이며, 둘째 기법은 DTW 법을 수행하여 얻은 최적경로에 직교인자를 적용하는 것이다. 여기서 직교인자는 선형예측분석에 의해 얻은 선형예측계수와 부분자기상관계수로 부터 얻을 수 있으므로 각각의 직교인자를 구하여 각 기법마다 두가지 직교인자를 적용하여 비교하였으며, 이들의 타당성을 확인하기 위하여 종래의 통계적 처리방법과 비교 검토하였다.

## II. 직교인자와 DTW에 의한 화자인식

본 논문에서는 직교인자와 DTW 알고리즘을 결합시켜 화자인식을 수행하는 기법을 제안하고자 한다. 이 기법을 제안하려고 하는 것은 만일 직교인자를 통계적으로만 처리한다면 음성신호의 동특성에 존재하는 개인성 정보를 충분히 이용하지 못하기 때문이다.

직교인자와 DTW 알고리즘을 결합하는 대에는 다음과 같은 두가지기법을 사용하고자 한다. 첫째는 음성신호의 직교인자를 특징벡터로 하여 DTW 알고리즘에 따라 기리를 측정하는 것이며, 둘째는 선형예측계수의 Loglikelihood ratio를 구한 후 이 최적경로에 기리를 측정하는 것이다.

이러한 기법들을 동특성을 이용하지 않은 화자인식 방법과 비교, 평가하기 위하여 먼저종래에 사용되

어왔던 직교인자의 통계적 처리방법을 제시하고 다음에 동특성을 이용하는 두가지 기법을 제안하고자 한다.

1. 직교인자의 통계적 처리방법

이 방법은 일반적으로 직교인자를 통계적으로 처리하기 위하여 Mahalanobis 거리를 사용한다. Mahalanobis 거리는 시험데이터의 평균계수벡터  $(X)_T$  와 표준데이터의 평균계수벡터  $((X)_R$  사이의 거리로 주어진다. 이러한 거리는 다음과 같이 정의된다.

$$D1 = [(X)_T - (X)_R] (C_X)_R^{-1} [(X)_T - (X)_R]^T \quad (1)$$

여기서  $(\cdot)_T$ 는 시험패턴  $(\cdot)_R$ 는 표준패턴이고  $X$ 는 전프레임수  $NF$ 에 대한 직교인자의 평균이며  $C_X$ 는 선형예측계수 또는 부분자기상관계수  $X$ 의 공분산이다. 또한 직교인자의 공분산열은 고유치가 되므로 다음과 같이 표시할 수 있다.

$$D1 = \sum_{k=1}^p \frac{[(\phi_k)_T - (\phi_k)_R]^2}{(\lambda_k)_R} \quad (2)$$

$$\phi_k = \frac{1}{NF} \sum_{j=1}^{NF} \phi_{kj} \quad (3)$$

여기서  $\phi_k$ 는  $k$ 차 평균직교인자,  $\lambda_k$ 는  $k$ 차 고유치,  $NF$ 는 총프레임의 수이다. 그림 1은 이 기법을 수행하는 구성도이다.

음성신호가 입력되면 먼저 선형예측계수분석하고 추출된 계수들을 직교변환하여 식(3)과 같이 평균직교인자를 구하며, 식(2)의 거리  $D1$ 이 최소인 것을 선택하여 화자를 인식한다.

2. 직교인자를 특징벡터로 한 기법

직교인자는 화자의 개인성 정보를 많이 포함하고 있으므로 Mahalanobis 거리 측정과 같은 통계적 처리방법에 의한 화자인식방법으로 쓰여왔다. 그러나 화자의 발성속도가 변화함에 따라 시간축의 특성이 비선형적으로 나타나기 때문에 각화자의 개인성 정보를 시간축에 따라 정규화하여 이용하는 것은 화자를 정확히 추정하는데 중요하다. 이 기법에 의해 화자인식을 수행하는 구성도를 그림 2에 나타내었으며, 직교인자벡터 사이의 거리는 유클리디안 거리측정방

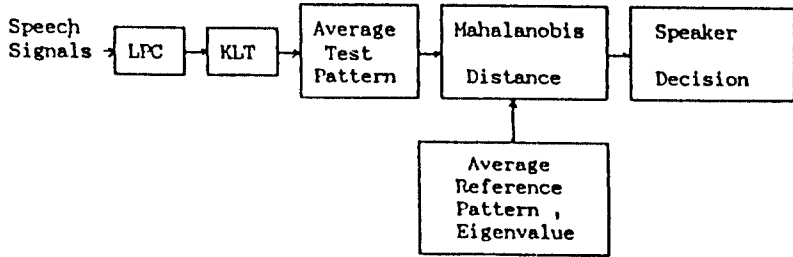


그림 1. 직교인자를 통계적으로 처리한 방법의 구성도  
Fig. 1. Block diagram of the statistical processing method of orthogonal parameters

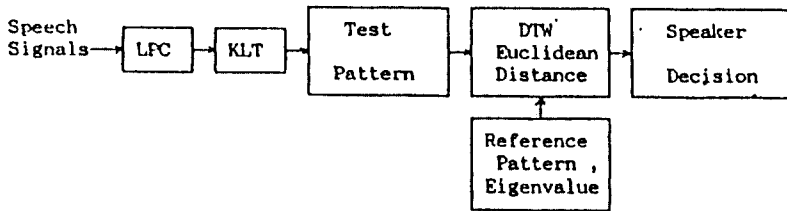


그림 2. 직교인자를 특징벡터로 한 기법의 구성도  
Fig. 2. Block diagram of the technique applying orthogonal parameters as feature vector

법으로 다음과 같다.

$$d(n, m) = \sum_{k=1}^p |(\phi_{kn})_T - (\phi_{km})_R| \quad (4)$$

여기서  $(\cdot)_T$ 는 시험패턴  $(\cdot)_R$ 는 표준패턴이며  $k$ 는 차수이고  $m$ 과  $n$ 은 각 패턴의 프레임을 나타낸다. 식(4)의 거리를 DTW에 사용하며, DTW알고리즘의 기울기는 1이고 형태는 대칭이다. 그림 3은 이 DTW 알고리즘의 경로이며, DTW의 알고리즘의 초기치는 다음식과 같다.

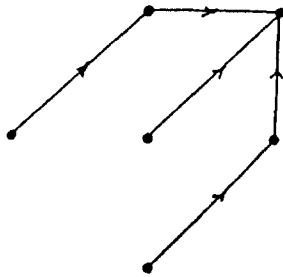


그림 3. 기울기(p)가 1이고 대칭인 DTW경로  
Fig. 3. Symmetric DTW path when slope is 1

(초기치의 설정)

$$\begin{aligned} D(1, 1) &= 2d(1, 1) \\ D(1, m) &= D(1, m-1) + d(1, m) : 2 \leq m \leq r \\ D(n, 1) &= D(n-1, 1) + d(n, 1) : 2 \leq n \leq r \\ D(1, m) &= \infty : m > r \\ D(n, 1) &= \infty : n > r \end{aligned}$$

여기서  $r$ 은 adjustment window이다.

또한 그림 3과 같이 주어진 DTW의 경로에 따라 누적되는 거리는 다음과 같다.

(누적거리)

$$D(n, m) = \begin{cases} D(n-1, m-2) + 2d(n, m-1) + d(n, m) \\ D(n-1, m-1) + 2d(n, m) \\ D(n-2, m-1) + 2d(n-1, m) + d(n, m) \end{cases} \quad (6)$$

또한 시험패턴  $n$ 축의 프레임 수를  $N$ , 표준패턴  $m$ 축의 프레임수를  $M$ 이라하면, 최종적인 시간축정규

화 거리  $D2$ 는 다음과 같다.

$$D2 = \frac{D(n, m)}{N+M} \quad (7)$$

### 3. 최적경로를 이용한 기법

본절에서는 같은 음성으로 발음한 음성신호에 대하여 화자인식을 하려는 것으로 비선형적으로 발생된 음성신호에 대하여 먼저 Itakura의 log likelihood ratio 방법으로 DTW를 수행하여 최적경로를 구하고 이 경로에 KLT에 의해 구해진 직교인자를 적용시켜서 화자의 동특성을 충분히 고려하여 인식하려는 것이다. 이기법으로 화자인식을 수행하는 기본 구성도는 그림 4와 같다.

여기에서 첫번째 처리과정인 LPC에서는 입력된 음성신호를 선형예측분석하여 선형예측계수와 부분 자기상관계수를 구한다. 그리고 점선으로 나타난 처리과정은 log likelihood ratio를 이용하는 DTW 알고리즘으로 각 화자에 해당하는 최적경로를 구하는 것이다. 이때 이용된 log likelihood ratio는 다음과 같다.

$$d(n, m) = \ln \left| \frac{am^T V_n am}{an^T V_n an} \right|$$

여기서  $am$ 은 표준 선형예측계수벡터이고  $V_n$ 는 시험자기상관계수 벡터이고  $an$ 은 시험 선형예측계수 벡터이다.

또한, 실선으로 나타난 처리과정은 KLT 방법으로 직교인자를 구하여 시험패턴을 구성한 후 최적경로의 warping 함수에 따라 각 화자의 직교인자로 구성된 표준패턴과의 거리를 구하여 화자를 인식하는 것이다. 이때 거리를 구하는 식은 다음과 같다.

$$D3 = \sum_{k=1}^p |(\phi_{kn})_T - \phi_{km})_R| \quad (9)$$

여기서  $n$ 과  $m$ 은 각각 warping 함수  $F$ 로 결정된 시험패턴과 표준패턴의 mapping 프레임이며 직교인자  $\phi_{km}$ 은  $m$ 번째 프레임의  $k$ 차 직교인자이다.

## III. 실험 및 결과 고찰

### 1. 실험시스템

본 연구에서 사용할 음성신호의 수집과 음성신호

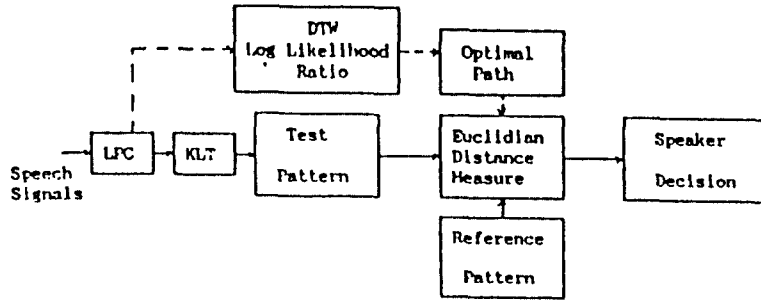


그림 4. 최적경로를 이용한 기법의 구성도  
Fig. 4. Block diagram of the optimal path technique

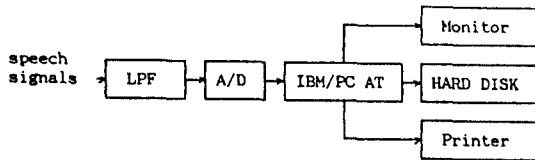


그림 5. 실험시스템의 구성도  
Fig. 5. Block diagram of experiment system

의 분석 및 화자인식을 위하여 다음 그림과 같이 실험시스템을 구성하였다.

화자인식에서 사용할 음성은 방음처리가 되지 않은 실험실에서 다이내믹형 마이크로폰을 사용하여 컴퓨터에 입력하였다. 입력되는 음성신호는 먼저 차단주파수가 4.0[kHz]인 저역통과필터에 통과되고 IBM PC /AT에 부착된 12bit 해상도와 샘플링 주파수가 10[kHz] 인 A/D 콘버터를 거쳐 컴퓨터의 하드 디스크로 저장된다. 이 디지털 처리된 음성신호를 육안으로 확인하기 위하여 비데오모니터에 디스플레이할 수 있도록 하였으며, 메모리용량을 줄이기 위하여 직접파일(direct file) 형식으로 저장할 수 있도록 하였다.

음성신호의 분석 및 화자인식의 실험은 동일한 IBM PC /AT 상에서 수행하였으며 분석 및 인식과정에서 얻은 그래픽 출력은 EPSON LQ-2500 프린터를 사용하였다.

## 2. 전처리 및 표준패턴의 작성

화자의 인식 실험을 수행하기 위해 사용한 문장은 유성음을 되도록 많이 포함하여 유성의 특징을 추출

하기 용이한 “인식의 나무는 생명의 나무가 아니다.” 라는 문장을 사용하였다. 이 문장 음성을 컴퓨터로 처리하기 위한 전처리의 분석조건은 표 1과 같다.

표 1. 분석조건  
Table. 1. Analysis conditions

Sampling frequency	: 10.0[kHz]
Cut-off frequency of LPF	: 4.0[KhZ]
A/D resolution	: 12.0[KhZ]
Analysis window	: 25.6[KhZ]
Overlapping interval	: 12.8[KhZ]
Window function	: Hamming window
LPC order	: 10

이 문장을 7인에게 3초간 발생하도록 하여 표준패턴을 작성하기 위한 음성 데이터를 구성하였고, 인식 실험에 사용할 시험 데이터는 각 사람이 10번씩 이 문장을 발생하도록 하여 메모리에 파일로 수집하였다.

본절에서 사용한 화자인식으로는 종래에 사용되어 온 직교인자의 통계적 처리방법과 본 연구에서 시도하는 직교인자를 특징벡터로한 기법, 최적 경로를 이용한 기법이 있다. 여기서 직교인자는 성형예측계수를 직교변환하여 구할 수 있고 또 부분사기 상관 계수를 직교변환하여 구할 수 있으므로 본 연구에서 사용한 특징부터는 두 가지 직교인자를 모두 사용한다.

각 화자의 문장에 대한 표준 패턴은 다음과 같이 만들었다. 각 화자가 첫기간부터 네번째 기간까지 발생한 문장 10(개)x4(기간) = 40개 중에서 자동집단화방식인 MKM(Modified K-means Method)알고리즘에서 집단의 center를 구하는 minimax center

방식으로 7화자의 7개 표준 패턴을 구성하였다.

minimax center는 각 패턴이 갖는 최대의 거리중에서 최소의 것을 선택한 것으로서 minimax center가  $P_k$ 라 하면 다음식과 같이 표시 할 수 있다.

$$\max_{1 \leq m \leq J} D(P_k, P_m) \geq \min_{1 \leq j \leq J} \max_{1 \leq m \leq J} D(P_j, P_m) \quad (10)$$

여기서  $P_k$ 는 k번째 발성된 음성신호의 선형예측계수로 구성된 패턴이며  $D(\cdot)$ 는 log likelihood ratio를 특징벡터의 거리로 하는 DTW 알고리즘을 이용한 시간축 정규화 거리이다.

### 3. 화자인식 결과

각 화자가 발성한 문장으로부터 추출한 선형예측계수를 가지고 평균직교인자를 얻어 Mahalanobis 거리  $D_{1L}$  (선형예측계수를 직교변환한 경우 "L"을 하위 첨자로 함)을 구하므로써 화자인식 실험을 수행하였다. 그 결과 표 2와 같은 결과를 얻었으며 인식율은 82.9%이었다.

또한 부분자기상관계수로 부터 평균직교인자를 얻어 Mahalanobis거리  $D_{1P}$  (부분자기상관계수를 직교변환하는 경우 "P"를 하위첨자로 함)를 구하므로써 화자인식 실험을 수행한 결과, 표 3과 같은 결과를 얻었으며 인식율은 68.6% 이었다.

직교인자를 특징벡터로 한 기법에서도 먼저 선형예측계수로부터 얻은 직교인자를 DTW 알고리즘의 특징 벡터로 하여 시간축 정규화 거리  $D_{2L}$ 을 구하므로써 화자인식 실험을 수행하였다. 그 결과 표 4와 같은 결과를 얻었으며 인식율은 85.7%이었다. 또한 부분자기상관계수로부터 얻은 직교인자를 특징벡터로 하여 시간축 정규화 거리  $D_{2P}$ 를 구하므로써 화자인식 실험을 수행하여 표 5와 같은 결과와 전체 인식율 84.3%를 얻었다.

음성신호의 선형예측계수를 log likelihood ratio에 적용하여 DTW알고리즘에 의한 최적경로를 얻은 후 이 경로에 따라 선형예측계수로부터 얻은 직교인자를 특징벡터로 하여 거리  $D_{3L}$ 을 구하여 화자인식 실험을 수행하였다. 그 결과를 표 6에 나타내었으며, 88.6%의 인식율을 얻을 수 있음을 알 수 있었다. 또한 이미 구해 놓은 최적경로에 따라 부분자기상관계수로부터 얻은 직교인자를 특징벡터로 하여 거리  $D_{3P}$ 를 구하므로써 화자인식 실험을 수행한 결과 표7과

같은 결과를 얻었으며 인식율은 85.7% 이었다.

### 4. 인식율의 비교 및 고찰

직교인자의 추출과정에 따른 인식율을 비교하여 보면, 부분자기상관계수로 부터 직교인자를 얻는 경우보다 선형예측계수로 부터 얻은 직교인자를 화자인식에 이용한 경우 통계적인 기법에서 14.3%, 직교인자를 특징벡터로 한 기법에서 2.4%, 최적경로를 이용한 기법에서 2.9%의 높은 인식율을 얻으므로써, 선형예측계수로 부터 직교인자를 얻은 경우 화자인식의 적합한 직교인자를 얻을 수 있었다.

그림 6에는 표 2에서 표7까지 나타낸 각 기법에 따른 인식율을 나타내었다. 여기서 횡축은 화자인식을 위한 각 기법의 종류이며, 종축은 각종류에 따른 인식율을 나타내었다. 이 그림의 인식율을 비교하기 위한 막대그래프로 부터 최적경로를 이용한 기법에서 가장 높은 인식율을 얻을 수 있음을 확인하였다.

이와같이 동적인 특성을 이용한 화자인식 기법에서 보다 높은 인식율을 얻을 수 있었던 것은 음성의 발성속도가 서로 다르므로 생기는 시간의 비선형적인 수축을 제거할 수 있었기 때문으로 생각되며, 특히 최적경로를 이용한 기법에서 가장 좋은 인식율이 나타낸 것은 같은 문장으로 발성된 음성신호에 대하여 화자인식을 했기 때문에 언어적 정보가 많은 선형예측계수의 log likelihood ratio에 의한 DTW의 최적경로가 개인정보를 많이 포함한 직교인자를 정합시키기에 적합했기 때문이라 사료된다.

표 2. 선형예측계수의 직교인자를 통계적 처리한 인식결과  
Table. 2. Recognition results using statistical processing method of orthogonal parameters from linear predictive coefficients

	LCY	JGH	KWG	PYC	IBG	CSB	SBS
LCY	8				6		
JGH		10					
KWG			10				
PYC				10		4	
IBG	2				4		
CSB						6	
SBS							10

표 3. 부분자기상관계수의 직교인자를 통계적 처리한 인식 결과

Table. 3. Recognition results using statistical processing method of orthogonal parameters from PARCOR coefficients

	LCY	JGH	KWG	PYC	IBG	CSB	SBS
LCY	5				6		
JGH	3	10					
KWG			9				
PYC			1	10		10	
IBG	2				4		
CSB							
SBS							10

표 4. 선형예측계수의 직교인자를 특징벡터로 한 인식결과

Table. 4. Recognition results of the technique applying orthogonal parameters from linear predictive coefficients

	LCY	JGH	KWG	PYC	IBG	CSB	SBS
LCY	6				3		
JGH		10					
KWG			10		3		
PYC				10			
IBG					4		
CSB						10	
SBS	4						10

표 5. 부분자기상관계수의 직교인자를 특징벡터로 한 인식 결과

Table. 5. Recognition results of the technique applying orthogonal parameters from PARCOR coefficients

	LCY	JGH	KWG	PYC	IBG	CSB	SBS
LCY					1		
JGH		10					
KWG			10				
PYC				10			
IBG					9		
CSB						10	
SBS	10						10

표 6. 선형예측계수의 직교인자를 최적경로를 이용한 기법에 적용한 인식결과

Table. 6. Recognition results of the technique using an optimal path applying orthogonal parameters from linear predictive coefficients

	LCY	JGH	KWG	PYC	IBG	CSB	SBS
LCY	2						
JGH		10					
KWG			10				
PYC				10			
IBG	3				10		
CSB						10	
SBS	5						10

표 7. 부분자기상관계수의 직교인자를 최적경로를 이용한 기법에 적용한 인식결과

Table. 7. Recognition results of the technique using an optimal path applying orthogonal parameters from PARCOR coefficients

	LCY	JGH	KWG	PYC	IBG	CSB	SBS
LCY							
JGH		10					
KWG			10				
PYC				10			
IBG					10		
CSB						10	
SBS	10						10

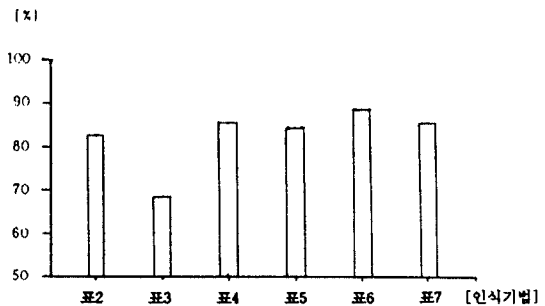


그림 6. 인식율의 비교

Fig. 6 Comparison of recognition rates

### IV. 결 론

본 연구에서는 직교인자를 이용한 화자인식방법에 화자의 발성에 따른 시간축 동특성을 고려하기 위하여 DTW 알고리즘에 의한 최적경로를 이용하는 화자인식방법을 제안하였다. 이 방법의 우수성을 확인하기 위하여 직교인자를 통계적으로 처리하는 종래의 방법과 비교 검토한 결과 다음과 같은 결론을 얻었다.

첫째, 직교인자를 DTW 알고리즘에 의한 최적경로에 이용한 기법으로 화자인식을 수행하여 직교인자를 통계적으로 처리하는 화자인식 방법보다 높은 인식율을 얻었다.

둘째, 직교인자의 추출과정에 따른 인식율을 비교하여 보면 부분자기상관계수로 부터 직교인자를 추출한 경우보다 선형예측계수로 부터 직교인자를 얻은 경우 화자인식에 적합한 직교인자를 얻을 수 있음을 알 수 있었다.

본 실험에서는 5개월동안 7인의 화자가 발성한 동일한 문장의 음성데이터를 수집하여 실험을 행하였으나 좀더 많은 화자를 통해서 장기간에 걸친 더욱 많은 음성 데이터를 수집하여 표준패턴을 구성하면 높은 신뢰도를 갖는 화자인식기법으로 정착될 수 있을 것으로 기대되며, 본 연구의 방법이 다른 문장의 음성 데이터에 대한 화자인식 실험에도 성과가 있을 것으로 사료된다.

### 参 考 文 獻

1. Douglas O'Shaughnessy, "Speaker recognition," IEEE ASSP Magazine, pp. 4-17, Oct. 1986
2. A.E.Rosenburg, "Automatic speaker verification: A review," Proc. IEEE Vol. 64, pp. 475-487, Apr. 1976.
3. Bishnu S.Atal, "Automatic recognition of speakers from their voice" Proc. IEEE, Vol. 64, pp. 460-474, Apr. 1976.
4. 新美康永, 音聲認識, 共立出版
5. Reich A., Moll K., Curtis J., "Effects of selected vocal disguises upon spectrographic speaker identification," J.Acoust. Soc.Amer., Vol. 60, pp.919-925, 1976.

6. J.W.SRayleigh, The Theory of Sound, 2nd Ed., Dover Publications Inc., New York, 1945.
7. S. Pruzansky, "Pattern-matching procedure for automatic talker recognition," J.Acoust. Soc. Amer., Vol.35, 1963.
8. M.R.SAMBUR, "Speaker recognition using orthogonal linear prediction," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP 24, pp.403-409, Aug. 1976.
9. R.E. BOGNER, "On talker verification via orthogonal parameters," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-29, pp. 1 12, Feb. 1981.
10. H.Sakoe, S.Chiba, "Dynamic programming optimization for spoken word recognition," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-26, pp.43-49, Feb. 1978.
11. S.FURUI, "Comparison of speaker recognition method using statistical features and dynamic features," IEEE TransAcoust., Speech, Signal processing, Vol. ASSP 29, pp 342-350, Jun. 1981.
12. J.d.MARKEL, et al., "Long-term feature averaging for speaker recognition," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP25, pp.421-428, Aug.1977.



**裴 澈 秀 (Chul Soo Bai) 正會員**  
 1950년 3월 14일생  
 1979년 2월 : 명지대학교 전자공학과 졸업  
 1982년 2월 : 명지대학교 대학원 전자공학과 석사  
 1988년 8월 : 명지대학교대학원 전자공학과 박사  
 1981년 3월~1990년 11월 : 관동대학교 전자공학과 부교수  
 1990년 12월~현재 : 관동대학교 전자통신과 부교수  
 1989년 11월~현재 : 한국통신학회 강원지부장