

다층신경망 기반 화자증명 시스템에서 학습 데이터 감축을 통한 화자등록속도 향상방법

A Method on the Improvement of Speaker Enrolling Speed for a Multilayer Perceptron Based Speaker Verification System through Reducing Learning Data

이 백 영*, 이 태 승*, 황 병 원*
(Baeg-Yeong Rhee*, Tae-Seung Lee*, Byong-Won Hwang*)

*한국항공대학교 항공전자공학과
(접수일자: 2002년 2월 22일; 채택일자: 2002년 8월 12일)

다층 신경망 (MLP: multilayer perceptron)은 기존의 패턴인식 방법에 비해 몇 가지 이점을 제공하지만 학습에 비교적 많은 시간을 요구한다. 이 점은 화자증명 시스템의 인식방법으로서 다층 신경망을 사용할 경우 등록시간이 길어지는 문제를 발생시킨다. 본 논문에서는 기존의 시스템에서 채택한 화자군집 방법을 응용하여 다층 신경망 학습에 필요한 배경화자 수를 줄임으로써 화자등록 시간을 단축하는 방법을 제안하고, 지속음을 인식단위로 하는 다층 신경망 화자증명 시스템에 이 방법을 적용한 실험결과를 통해 그 효과를 확인한다.

핵심용어: 다층신경망, 화자증명, 화자군집, 음성인식, 패턴인식

투고분야: 음향 신호처리 분야 (1.1, 1.7), 음성처리 분야 (2.5)

While the multilayer perceptron (MLP) provides several advantages against the existing pattern recognition methods, it requires relatively long time in learning. This results in prolonging speaker enrollment time with a speaker verification system that uses the MLP as a classifier. This paper proposes a method that shortens the enrollment time through adopting the cohort speakers method used in the existing parametric systems and reducing the number of background speakers required to learn the MLP, and confirms the effect of the method by showing the result of an experiment that applies the method to a continuous- and MLP-based speaker verification system.

Keywords: Multilayer perceptron, Speaker verification, Cohort speakers, Speech recognition, Pattern recognition

ASK subject classification: Acoustic signal processing (1.1, 1.7), Speech signal processing (2.5)

I. 서론

생체적 특징으로서 음성을 이용하여 음성을 발성한 화자를 구별하는 처리를 화자인식이라고 한다. 화자인식은 시스템에 여러 화자를 등록하고 이들 중 현재 음성과 일

치하는 화자를 선택하는 화자식별과 시스템에 미리 등록해 둔 특정 화자의 신원을 선택한 뒤 현재 음성이 그 화자와 일치하는지 여부를 판별하는 화자증명으로 나뉜다. 이 중 화자증명이 불특정 다수를 대상으로 신원을 확인하고, 이러한 화자증명 모듈을 다수 결합함으로써 화자식별 시스템을 구성할 수 있기 때문에 화자증명 시스템의 연구 비중이 더 높다[1].

특정화자의 음성을 학습하고 입력된 음성을 이렇게 학

책임저자: 이태승 (thestaff@hitel.net)
412-791 경기도 고양시 덕양구 화전동 200-1
한국항공대학교 전자관 324호
(전화: 019-208-4667; 팩스: 02-3159-9986)

습된 음성과 비교하여 신원을 판별하기 위해 다양한 인식 방법이 사용된다. 이들 중 다층 신경망 (MLP: multilayer perceptron)은 다음과 같은 이점을 갖고 있어 화자인식 외에도 다양한 인식문제에서 사용되고 있으며, 화자식별과 화자증명에 대한 응용도 시도되고 있다[2-5].

- 논파라메트릭 (nonparametric) 방식이기 때문에 문제에서 가정해야 하는 하부확률분포가 필요없다.
- 학습되는 각 모델 사이의 차이를 최대한 구별하는 분별학습능력이 있기 때문에 인식오류 가능성을 최소화한다.
- 학습모델별로 +1, 0 (또는 -1)의 학습목표치를 사용할 때 선형식별분석 (LDA: linear discriminant analysis)과 유사한 특징공간 변환능력을 갖는다.

화자증명을 위해서는 신원점수를 측정할 의뢰화자와 비교를 위한 배경화자가 필요하다. 의뢰화자는 다시 실제화자와 사칭화자로 나뉘는데, 실제화자를 사칭화자로 잘못 인식하는 비율인 오인거부율 (False Reject Rate)과 그 반대의 비율인 오인수락률 (False Accept Rate)을 기준으로 화자증명 시스템의 인식률을 측정한다. 화자증명 인식률을 높이기 위해서는 의뢰화자를 되도록 특성이 유사한 배경화자와 비교하여 엄밀한 판별이 이뤄지도록 해야 한다. 그러나 의뢰화자의 특성을 미리 알 수 없기 때문에 가능한 한 많은 배경화자를 준비하여 어떤 화자가 신원확인을 의뢰하더라도 정확한 판별이 이뤄질 수 있도록 한다[6]. 그러나 배경화자의 규모가 늘어날수록 화자증명의 인식률은 향상되지만 처리에 필요한 계산량도 함께 증가한다. 이 점은 화자의 등록과 신원확인을 실시간으로 처리해야 하는 화자증명 시스템에서 반드시 해결해야 할 문제이다.

이에 대해 파라메트릭 방식의 기존 화자증명 시스템에서는 화자등록시 등록화자와 유사한 제한된 수의 배경화자를 선택한 뒤 의뢰화자의 신원판별시 이 군집의 화자정보를 이용하여 의뢰화자의 신원점수를 계산하는 방법을 도입하였다[6-7]. 이 시스템은 신원확인 처리에 집중되는 계산을 화자등록 단계로 분산함으로써 신원확인에 필요한 시간을 단축한다.

파라메트릭 방식의 시스템에 비해 다층 신경망을 이용하는 화자증명 시스템에서는 의뢰화자가 지명하는 신원의 화자와 배경화자의 영역차이를 등록단계에서 학습해 두고 의뢰화자가 어느 영역에 속하는지 판별함으로써 중

명처리를 수행한다. 이 경우 신원확인을 위한 계산은 신속히 이뤄지는 반면 화자를 등록하기 위한 시간, 즉 학습 시간은 상대적으로 길다[3]. 화자등록 후 곧바로 화자증명 서비스를 제공하기 위해서는 이러한 학습시간을 최대한 줄일 필요가 있다.

본 논문에서는 기존의 시스템에서 사용하는 화자군집 방법을 도입하여 다층 신경망의 화자학습에 필요한 데이터량을 줄임으로써 다층 신경망 화자증명 시스템의 화자등록시간을 단축하는 방법을 제안한다. 이 방법을 설명하기 위해 먼저 2절에서 파라메트릭 방식 시스템에서 인식을 향상과 증명처리속도 증대를 위해 사용하는 화자평준화 방법과 화자군집 방법을 소개한 뒤, 3절에서 다층 신경망을 사용하는 시스템에서의 등록시간 문제를 제기한다. 4절에서는 이러한 특성을 감안하여 화자군집을 적용하는 방법을 제안하고, 5절에서 이 방법을 실제 시스템에 적용하여 성능을 시험한 뒤, 6절에서 본 논문의 성과를 정리한다.

II. 화자평준화 및 화자군집 화자증명

Higgins 등[8]은 신뢰할만한 신원점수를 얻기 위해 의뢰화자의 유사도 점수를 의뢰화자와 가장 유사한 배경화자의 점수로 평준화하는 방법을 다음과 같이 제안하였다.

$$L(\mathbf{X}) = \frac{p(\mathbf{X} | S_C)}{\max_{S_{BG}, S_{BG} \neq S_C} p(\mathbf{X} | S_{BG})} \quad (1)$$

여기서 \mathbf{X} 는 s_C 는 의뢰화자의 음성열이고, 의뢰화자가 주장하는 신원을, S_{BG} 는 의뢰화자에 가장 근접한 배경화자를 나타낸다.

그러나 이 점수는 (a) 의뢰화자와 가장 근접한 배경화자를 찾기 위해 모든 배경화자를 탐색하므로 계산량이 배경화자의 규모와 비례하고, (b) 배경화자 집단에서 의뢰화자의 위치에 따라 (즉, 의뢰화자 주위의 배경화자 분포밀도에 따라) 식 (1)의 분모항의 값이 크게 달라지므로 안정적인 점수를 얻기 힘들다는 단점을 갖는다.

이 문제를 해결하기 위해 화자를 시스템에 등록할 때 일정한 수의 근접 배경화자를 선택한 뒤, 의뢰화자 점수를 계산할 때 이들의 유사도 점수를 합한 값을 평준화 항으로 사용하는 화자군집 방법이 제안되었다.

$$L(\mathbf{X}) = \frac{p(\mathbf{X} | S_C)}{\sum_{S_{Cohort}, S_{Cohort} \neq S_C} p(\mathbf{X} | S_{Cohort})} \quad (2)$$

여기서 S_{Cohort} 는 등록화자와 근접한 화자군집을 뜻한다. S_{Cohort} 을 이루는 배경화자의 수가 많을수록 증명점수의 신뢰성이 높아진다.

S_{Cohort} 의 배경화자 수가 무한대라고 가정한다면 Bayes 법칙에 따른 의뢰화자의 주장화자에 대한 사후확률에서 화자들의 사전확률이 고정됨에 따라 식 (2)를 사후확률로 볼 수 있다.

$$\begin{aligned} P(S_C | \mathbf{X}) &= \frac{p(\mathbf{X} | S_C)}{\sum_{S} p(\mathbf{X} | S)} \approx L(\mathbf{X}) \\ &= \frac{p(\mathbf{X} | S_C)}{\sum_{S_{Cohort}} p(\mathbf{X} | S_{Cohort})} \quad (3) \end{aligned}$$

그러나 화자군집을 이루는 배경화자의 수가 일정한 수준 이상이면 인식을 향상에는 거의 기여하지 못하면서 계산량은 선형적으로 증가하게 된다. 따라서 실험을 통해 만족할만한 인식을 수준에서 가장 적은 배경화자를 선택하는 것이 유리하다.

처리속도의 측면에서 봤을 때 화자군집 방법은 식 (1)의 신원증명에 요구되는 처리시간을 단축하기 위해 근접 배경화자 탐색처리를 화자등록 단계로 분산시킨 것으로 간주할 수 있다. 화자증명 시스템에서 계산이 발생하는 두 단계인 등록과 증명 중에서 등록단계가 상대적으로 증명단계보다 실시간에 대한 엄격도가 낮으므로 이러한 계산량 분산은 전반적인 실시간 성능을 높이는 효과를 갖는다.

III. 화자증명을 위한 다층 신경망의 등록시간 문제

Gish[9]는 다층 신경망이 충분한 학습용량을 가졌다면 다수모형을 분류하는 패턴인식에서 각 모델의 사후확률을 학습할 수 있다는 사실을 증명했다. 이에 따라 2개의 학습모델로서 등록화자와 충분한 수의 배경화자가 주어 진다면 다층 신경망의 학습은 식 (3)으로 표현되는 화자군집 방법을 근사화한다.

그러나 학습해야 하는 배경화자가 늘어날수록 다층 신경망의 학습속도는 길어지고, 이 점은 오류 역전파 (EBP: error backpropagation) 학습 알고리즘을 사용하는 다층 신경망의 학습속도가 다른 인식방법에 비해 상당히 느리

다는 점에 가중되어 등록시간이 길어지는 요인이 된다. 전술한 바와 같이 화자증명 시스템에서는 등록단계의 실시간 엄격도가 증명단계보다 여유롭지만, 사용자는 등록 후 곧바로 증명 서비스가 제공되기를 원할 것이므로 등록 시간을 무한정 늘릴 수 없다.

등록속도를 향상시키는 방법은 다층 신경망의 학습속도 자체를 증가시키거나 학습에 필요한 데이터를 줄이는 방법을 고려할 수 있는데, 본 논문에서는 식 (2)에서처럼 근접화자로 이루어지는 화자군집을 선택하여 학습에 필요한 데이터를 줄이는 방법을 제안한다.

IV. 제안방법

다층 신경망 학습에 필요한 배경화자의 수를 줄이기 위해 식 (2)와 같이 등록하는 화자에 인접한 제한된 수의 배경화자를 선택하되, 화자증명 시스템의 인식률을 저하시키지 않는 수준에서 최소수를 선택한다. 이 방법을 위해서 등록화자가 각 배경화자와 얼마나 근접하는지를 평가하는 다층 신경망과 이 다층 신경망을 통해 선택된 배경화자를 이용하여 등록화자의 특성을 학습하는 다층 신경망이 필요하다. 본 논문에서는 전자를 MLP-I, 후자를 MLP-II라고 한다.

MLP-I은 배경화자의 수와 동일한 개수의 출력뉴런을 가지며, 사전에 배경화자의 데이터를 이용하여 각각을 분류할 수 있도록 학습된다. 이 때 은닉뉴런이 많으면 배경화자의 근접도 평가만으로도 많은 시간이 소요될 수 있으므로, 각 배경화자 데이터를 80% 이상의 정확도로 분류할 수 있을 정도로만 은닉뉴런 개수를 선택한다. MLP-II는 MLP-I을 통해 선택된 배경화자의 데이터를 이용하여 등록화자를 학습한다. 단, 다층 신경망의 경우 화자군집 방법과 달리 학습모델을 등록화자와 배경화자의 2개로 설정하므로 MLP-I으로 선택된 배경화자들의 패턴을 한 모델로 통합한다.

MLP-I에서 선택된 배경화자의 수는 실험을 통해 결정한다. 이 수는 등록화자의 분포범위를 선택된 배경화자들이 둘러쌀 수 있는 수준 이상이어야 하며, 인식률의 하락을 방지하기 위해서는 배경화자가 등록화자를 밀도 있게 둘러싸야 한다. 배경화자의 밀도에 따른 등록화자의 학습양상을 그림 1에서 설명하고 있다. 이 그림에서 점선은 등록화자와 배경화자의 영역을 결정하는 경계선을 나타낸다. 그림 2는 MLP-I과 MLP-II의 동작원리를 설명한다.

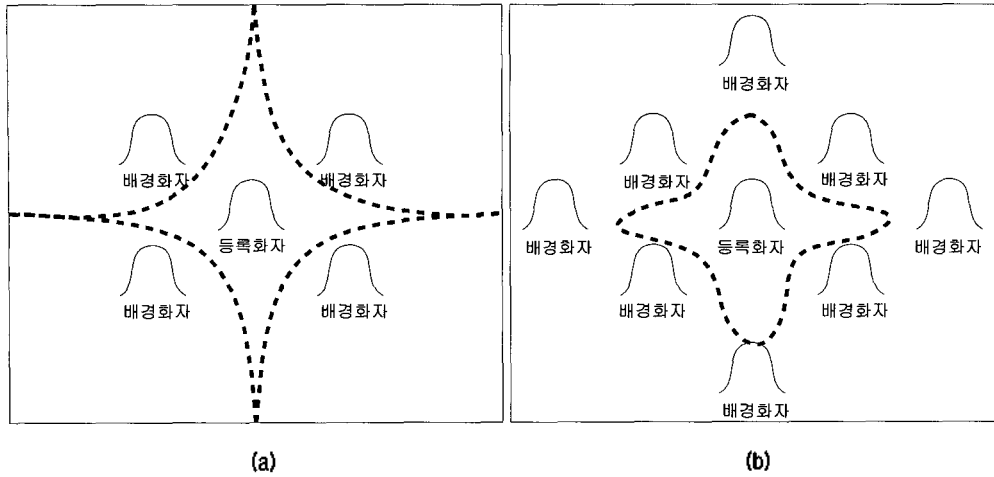


그림 1. (a) 저밀도 배경화자와 (b) 고밀도 배경화자의 경우에서 다층 신경망의 등록화자학습
 Fig. 1. MLP learning of enrolling speaker for (a) the sparse background speakers and (b) the dense background speakers.

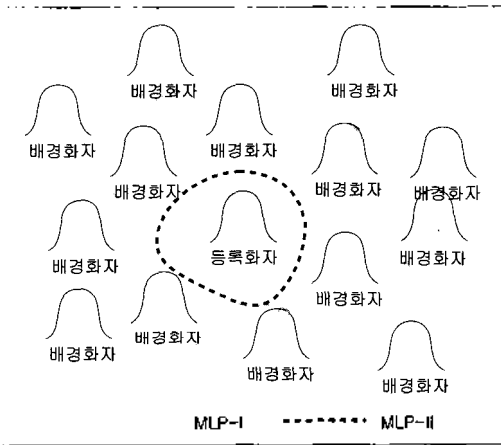


그림 2. MLP-I의 배경화자군집 선택과 MLP-II의 등록화자학습
 Fig. 2. MLP-I used to select background speakers and MLP-II used to learn enrolling speaker with background speakers.

V. 실험

5.1. 데이터베이스

이 실험에 사용할 데이터는 한국인 남녀 40명의 4연숫자 발성을 녹음한 것이다. 여기서 4연숫자라 함은 아라비아 숫자 0~9에 해당하는 /goN/, /il/, /i/, /sam/, /sa/, /o/, /yug/, /cil/, /pal/, /gu/ 음을 연속해서 4자리를 발성한 것을 의미한다. 각 화자가 총 35개의 서로 다른 숫자 음 배열을 4회씩 발성하였고, 발성은 16 kHz 주기의 16 bit 크기로 녹음되었다. 4회 발성 중 3회를 각 화자의 등록 음성으로 사용하고 나머지 1회를 증명시험 음성으로 사용한다.

등록화자 학습시 필요한 배경화자로는 위의 40명 외의

남녀 29명을 사용한다.

5.2. 화자증명 시스템

본 연구에서 구현한 화자증명 시스템은 입력음성에서 고립단어를 추출하고, 이 고립단어에서 한국어 지속음 (/a/, /e/, /o/, /u/, /i/, /l/, 비음)을 인식한 다음, 각 지속음별로 MLP-I과 MLP-II를 이용하여 화자를 학습하고 증명점수를 계산한다. 이 시스템에서 이뤄지는 처리를 그림 3에서 보여주며, 각 처리의 설명은 아래와 같다.

(1) 음성분석 및 특징추출

- 16 bit 16 kHz로 샘플링된 등록화자의 입력음성을 20 ms 오버랩시킨 30 ms 길이의 프레임으로 나눈다.
- 각 프레임에 대해 16차 Mel 간격 필터뱅크 (filter bank)[10]를 추출하여 고립단어 및 지속음 검출에 사용한다. 필터뱅크 계수는 전체 스펙트럼 포락에 미치는 성량의 영향을 제거하기 위해 1 kHz까지의 계수를 평균하여 이 값을 모든 계수에서 차감한다. 그리고 다층 신경망의 효과적인 학습을 위해 다시 모든 계수의 평균이 0이 되도록 조정한다.
- 각 프레임에 대해 50차의 0~3 kHz 대역 균등간격 Mel 필터뱅크를 추출하여 화자증명에 사용한다. 이 음성특징은 2차 포먼트 (formant)에 더 많은 화자정보가 집중된다는 연구결과[11]에 의한 것이다. 필터뱅크 계수는 전체 스펙트럼 포락에 미치는 성량의 영향을 제거하기 위해 1 kHz까지의 계수를 평균하여 이 값을 모든 계수에서 차감한다. 그리고 MLP의 효과적인 학습을 위해 다시 모든 계수의 평균이 0이 되도록 조정한다.

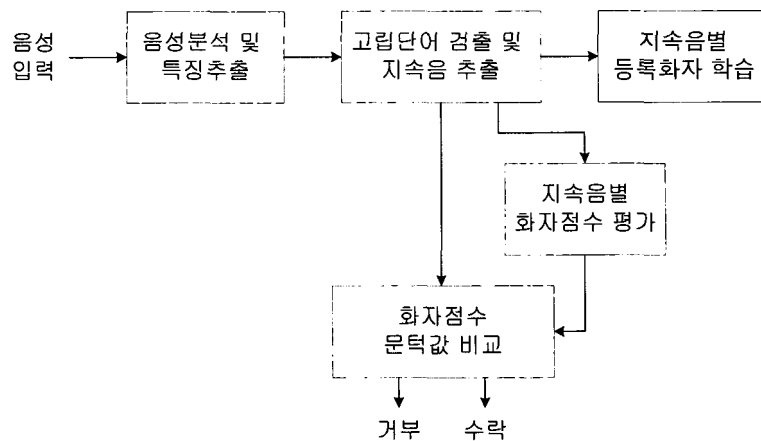


그림 3. 화자증명 시스템 처리 흐름도
Fig. 3. Process flow of the speaker verification system.

(2) 고립단어 검출 및 지속음 검출

- 각 지속음과 목음을 화자독립 방식으로 검출하도록 학습된 다층 신경망을 사용하여 고립단어와 고립단어 내의 지속음을 검출한다.

(3) 지속음별 등록화자 학습

- 지속음별로 전체 고립단어의 각 지속음을 MLP-I에 입력한 뒤, 출력뉴런의 수치를 평균하고, 이 평균치가 높은 순서로 n명의 배경화자를 선택한다.
- 지속음별로 선택된 n명의 배경화자 데이터를 이용하여 MLP-II에 등록화자를 학습시킨다.
- 지속음별로 MLP-II 학습에 사용되는 패턴수는 등록화자 당 10개씩이다.

(4) 지속음별 화자점수 평가

- 지속음별로 전체 고립단어의 각 지속음을 MLP-I에 입력한 뒤, 출력뉴런의 수치를 평균하고, 이 평균치가 높은 순서로 n명의 배경화자를 선택한다.
- 선택된 배경화자 가운데 (3)의 배경화자가 1명 이상 포함된 모든 지속음에 대해 MLP-II의 출력치를 평균한다.
- (3)의 배경화자가 1명 이상 포함된 지속음이 전무한 경우 의뢰화자를 거부한다.

(5) 등록어 및 화자점수 문턱값 비교

- (4)의 평균치와 사전 설정한 문턱값을 비교하여 최종적인 거부/수락을 결정한다.

이 화자증명 시스템에서는 지속음을 화자인식단위로 사용하기 때문에 하부확률분포가 단봉분포 (mono-modal)의

형태를 띤다[12]. 따라서 MLP-I과 MLP-II는 모두 1개의 은닉계층이 포함된 2층 구조만으로 충분하다[13-14]. 이 중 MLP-II는 학습하는 모델이 2개뿐이므로 1개의 출력뉴런과 2개의 은닉뉴런이면 충분히 이들을 학습할 수 있다.

5.3. 실험결과 및 분석

실험화자 40명을 한 명씩 차례로 등록화자와 실제화자로 사용하고 이를 제외한 나머지 39명을 사칭화자로 사용한다. 결과적으로 화자당 35회의 실제화자 시도와 1,560회의 사칭화자 시도를 평가하게 되고, 실험 전체로 봤을 때는 1,400회의 실제화자 시도와 54,600회의 사칭화자 시도를 평가하게 된다.

실험은 AMD 1.4G Hz급 컴퓨터에서 실시하였으며 그 결과를 그림 4에 정리하였다. 그림에서 오류율은 동일오류율 (Equal Error Rate)을 의미하고, 학습패턴수는 1명의 화자를 등록하기 위해 학습된 총 패턴수를 나타내며, 학습시간은 이 패턴들을 학습하는데 걸린 실제시간을 가리킨다. 오류율, 학습패턴수, 학습시간은 전체 증명시도에 대한 평균치이다.

이 결과를 보면 화자군집 방법을 적용했을 때 오류율이 상승하지 않는 군집내 최소 화자수는 26명이고 0.19% 상승시 화자수는 14명이며, 각각의 경우에서 학습시간은 4.5%, 35.4% 단축되었다. 또한 군집내 화자수가 배경화자 수와 같을 때 (즉, 학습데이터 감축 효과가 없을 때) 화자군집 방법을 적용하지 않은 경우보다 5% 정도의 학습시간 지연이 관측되었는데, 이는 MLP-I의 적용에 따른 계산량의 증대 때문이다.

군집내 화자수가 14명 이하일 때 인식률이 급격히 하락하는 것을 볼 수 있는데, 이것은 등록화자와 동일한 성별

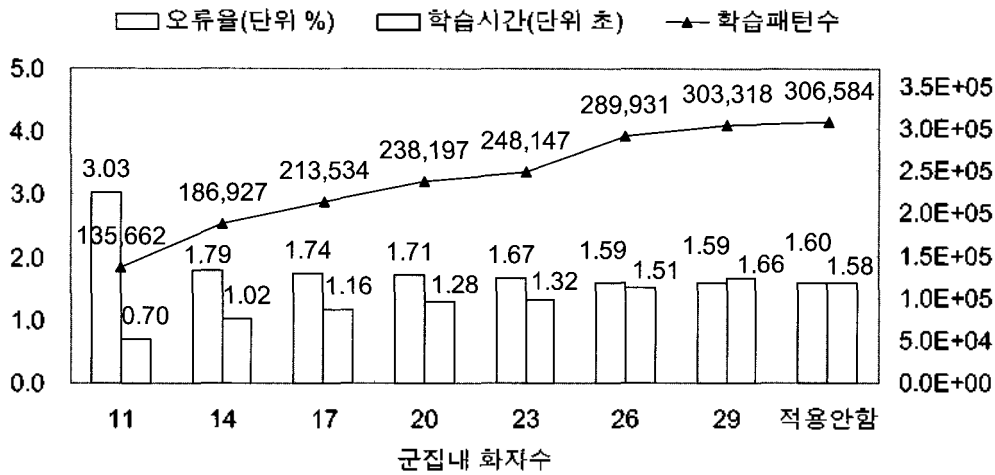


그림 4. 실험결과
Fig. 4. Experiment results.

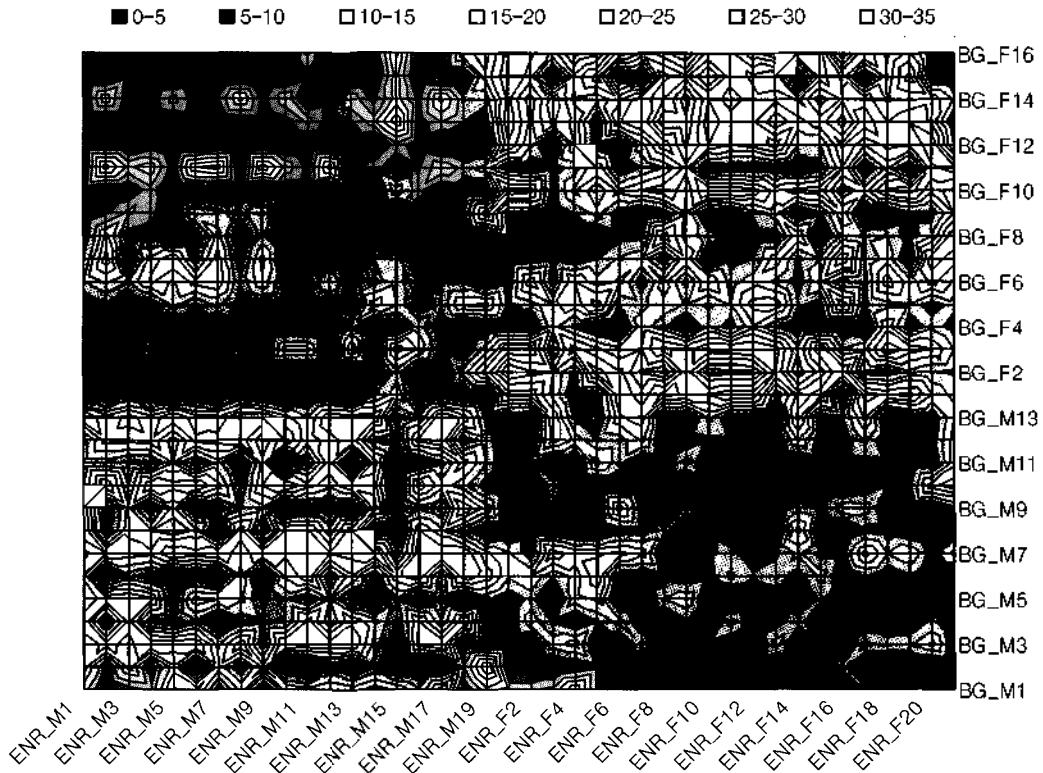


그림 5. 등록화자에 따른 배경화자의 선택분포
Fig. 5. The distribution of background speakers selected as to an enrolling speakers.

의 배경화자 수가 군집내에서 실질적으로 줄어들어 따른 결과이다. 그림 5는 등록화자에 따른 배경화자의 선택분포를 보여준다. 이 그림은 군집내 화자수를 11명으로 제한했을 때 35개 숫자열에 대해 등록화자별 배경화자의 선택수를 누적하여 나타낸 것으로, 가로축이 등록화자이고 세로축이 배경화자이다. 화자의 성별은 숫자 앞의 문자가 M이면 남성이고 F이면 여성을 뜻하고, 색이 흰색으

로 갈수록 해당 배경화자가 더 많이 선택되었음을 의미한다. 그림은 등록화자의 성별에 따라 확연히 같은 성별의 배경화자가 선택됨을 보여준다. 따라서 군집내의 화자수가 최소한 등록화자와 같은 성별의 배경화자 수보다 클 경우에는 인식률의 큰 하락없이 등록속도를 줄일 수 있다는 사실을 예측할 수 있다.

VI. 결론

다층 신경망이 다른 패턴인식 방법에 비해 제공하는 몇 가지 이점에도 불구하고, 다층 신경망의 상대적으로 느린 학습속도는 다층 신경망을 인식방법으로 사용하는 화자증명 시스템에서 화자등록시간이 길어지게 만드는 원인이 된다. 이에 본 논문에서는 화자군집 방법을 다층 신경망에 도입하여 화자등록시 필요한 배경화자의 수를 줄임으로써 학습시간 단축을 시도하였다. 지속음 단위로 화자특성을 구분하는 화자증명 시스템에서 이 방법을 실험한 결과 인식을 하락이 미미한 수준에서 등록속도를 향상시킬 수 있음을 확인하였다. 차후 이 방법과 다층 신경망의 학습속도를 향상시키는 방법을 결합하면 더 큰 효과가 있을 것으로 기대된다.

참고 문헌

1. S. Furui, "An Overview of Speaker Recognition Technology," *Automatic Speech and Speaker Recognition*, Kluwer Academic Publishers, 31-56, 1996.
2. N. Morgan and H. Bourlard, "Hybrid Connectionist Models for Continuous Speech Recognition," *Automatic Speech and Speaker Recognition*, Kluwer Academic Publishers, 259-283, 1996.
3. S. Haykin, *Neural Networks*, Prentice Hall, 1999.
4. Y. Bennani and P. Gallinari, "A modular connectionist architecture for text-independent talker identification," *Proceedings of the International Joint Conference on Neural Networks*, 2, 857-860, 1991.
5. N. Fakotakis and J. Sirigos, "A high performance text independent speaker recognition system based on vowel Spotting and Neural Nets," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2, 661-664, 1996.
6. A. E. Rosenberg and S. Parthasarathy, "Speaker background

- models for connected digit password speaker verification," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 1, 81-84, 1996.
7. T. Matsui and S. Furui, "Likelihood normalization for speaker verification using a phoneme- and speaker-independent model," *Speech Communication*, 17, 109-116, Aug 1995.
8. A. L. Higgins et al., "Speaker verification using randomized phrase prompting," *Digital Signal Processing*, 1, 89-106, 1991.
9. H. Gish, "A probabilistic approach to the understanding and training of neural network classifiers," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 3, 1361-1364, 1990.
10. C. Becchetti and L. P. Ricotti, *Speech Recognition*, John Wiley & Sons, 1999.
11. P. Cristea and Z. Valsan, "New cepstrum frequency scale for neural network speaker verification," *Proceedings of the IEEE International Conference on Electronics, Circuits and Systems*, 3, 1573-1576, 1999.
12. M. Savic and J. Sorensen, "Phoneme based speaker verification," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2, 165-168, San Francisco, USA, 1992.
13. R. P. Lippmann, "An Introduction to computing with neural nets," *IEEE Acoustics, Speech, and Signal Processing Magazine*, 4, 4-22, Apr 1987.
14. D. P. Delacretaz and J. Hennebert, "Text-prompted speaker verification experiments with phoneme specific MLPs," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2, 777-780, 1998.

저자 약력

● 이 백 영 (Baeg-Yeong Rhee)

한국음향학회지 제21권 제5호 참조

● 이 태 승 (Tae-Seung Lee)

한국음향학회지 제21권 제5호 참조

● 황 병 원 (Byong-Won Hwang)

한국음향학회지 제21권 제5호 참조