한국음향학회지 제20권 제4호 pp. 48~53 (2001)

다층 퍼셉트론에서의 빠른 화자 적응을 위한 선택적 주의 학습

Selective Attentive Learning for Fast Speaker Adaptation in Multilayer Perceptron

김 인 철*, 진 성 일*
(In Cheol Kim*, Sung Il Chien*)

*경북대학교 전자전기공학부

(접수일자: 2001년 3월 24일; 채택일자: 2001년 4월 30일)

본 논문에서는 에러 역전파 알고리듬에 기반한 다층 퍼셉트론의 학습 속도를 개선하기 위해 선택적 주의 학습 방식을 제안한다. 제안된 방식은 학습 과정에서 세 가지 선택적 주의 기준을 적용하여 학습 데이터베이스 내의 일부 데이터만을 입력 패턴으로 사용하거나 주어진 입력 패턴에 대해 신경회로망내의 특정 영역만 선택적으로 학습이 이루어지도록 한다. 이러한 선택적 주의 기준은 다층 퍼셉트론의 출력층에서 계산된 평균 자승 에러와 은닉층의 각 노드에서 획득된 클래스 의존적인 적합도 (relevance)를 이용하여 설정된다. 학습 속도의 개선은 학습 반복 횟수 당 계산량을 줄임으로써 이루어진다. 본 논문에서는 고립 단어 인식시스템에서의 화자 적응 문제에 대해 제안한 선택적 주의 학습 방법을 적용하여 그 유효성을 알아보았다. 실험 결과로부터 제안한 선택적 주의 기법이 학습 속도를 평균 60%이상 개선시킬 수 있음을 확인하였다.

핵심용어: 선택적 주의, 클래스 의존적인 적합도, 에러 역전파 알고리듬, 다층 퍼셉트론, 화자 적응
투고분야: 음성처리 분야 (2.5)

In this paper, selectively attentive learning method has been proposed to improve the learning speed of multilayer Perceptron based on the error backpropagation algorithm. Three attention criterions are introduced to effectively determine which set of input patterns is or which portion of network is attended to for effective learning. Such criterions are based on the mean square error function of the output layer and class-selective relevance of the hidden nodes. The acceleration of learning time is achieved by lowering the computational cost per iteration. Effectiveness of the proposed method is demonstrated in a speaker adaptation task of isolated word recognition system. The experimental results show that the proposed selective attention technique can reduce the learning time more than 60% in an average sense.

Keywords: *Selective attention, Class-selective relevance, Error backpropagation algorithm, Multilayer Perceptron, Speaker adaptation*
ASK subject classification: *Speech signal processing (2.5)*

I. 서 론

다층 퍼셉트론 (multilayer Perceptron)은 교사 학습 (supervised learning)에 기반한 인식 기법으로서 패턴인식, 시계열 함수의 비선형 예측, 그리고 비선형 제어 등의 분야에서 널리 사용되고 있다. 다층 퍼셉트론의 학습은 일반적으로 에러 역전파 (error backpropagation) 학습 알고리듬[1]을 이용하여 수행된다. 에러 역전파 알고리듬은 다층 퍼셉트론 내의 가중치 값들을 반복적으로 수정하

책임저자: 김인철 (kiminc@palgong.knu.ac.kr)
702-701 대구시 북구 산격동 1370
경북대학교 전자전기공학부 대학원
(전화: 053-940-8645; 팩스: 053-950-5505)

여 출력층의 평균 자승 에러 (mean square error) 함수가 최소화 되는 방향으로 학습을 수행한다. 그러나 이러한 학습 방법은 학습 데이터베이스의 규모가 커지거나 인식의 복잡도가 높아질수록 학습 시간이 매우 길어진다는 단점을 가진다. 오프라인 학습 방식을 채택하고 있는 응용분야에서는 느린 학습 속도가 큰 문제가 되지 않으나 온라인 및 재학습 방식에 관련된 분야에서는 그 적용을 어렵게 하는 주요인이 될 수 있다.

본 논문에서는 에러 역전파 알고리듬의 느린 학습 속도를 개선하기 위해 선택적 주의 학습 방법을 제안한다. 제안된 학습 방법은 학습 과정에서 학습 데이터베이스 내의 일부 데이터만을 선별적으로 학습을 위한 입력 패턴으로 사용하거나 주어진 입력 패턴에 대해 다층 퍼셉트론의 일부 영역만 학습이 이루어지도록 한다. 학습 시간 개선을 위한 기존의 연구들은 학습 과정에서 학습률을 적절히 조절하거나[2], 조기 포화 (premature saturation) 현상으로 인한 학습 지연을 방지하기 위해 평균 자승 에러가 아닌 다른 형태의 에러 함수를 사용함으로써[3,4] 학습 반복 (iteration) 횟수를 줄이는데 주안점을 두었다. 그러나 제안한 방식은 이러한 학습 방식과는 달리 한 반복 횟수 당 학습에 소요되는 계산량을 크게 줄임으로써 전체 학습 시간을 효과적으로 단축시키고자 한다. 이 방법은 표준 에러 역전파 알고리듬뿐만 아니라 전술한 개선된 학습 알고리듬들과도 손쉽게 결합되어 운용될 수 있다는 장점을 가진다.

제안한 학습 방식은 다층 퍼셉트론의 출력층에서 계산된 평균 자승 에러와 은닉층의 각 노드에서 획득된 클래스 의존적인 적합도 (relevance)를 이용한 세 가지 선택적 주의 기준에 따라 수행된다. 적합도 (relevance) 개념은 다층 퍼셉트론의 각 은닉층 노드의 중요도를 해당 노드가 제거된 상태에서의 평균 자승 에러 함수의 변화량으로 나타낸 것으로서 주로 신경회로망의 규모 축소 (pruning) [5] 문제와 관련된 연구 분야에 사용되어 왔다. 그러나 본 논문에서는 은닉층 노드의 중요도를 주어진 입력 클래스 별로 각각 따로 나타낼 수 있는 클래스 의존적인 적합도 (relevance)를 새로이 정의하여 학습에 적용한다.

이러한 선택적 주의 방식은 초기 학습된 가중치 값들을 기반으로 하는 재학습 문제에 특히 효율적으로 적용될 수 있다. 재학습 과정에서는 초기 학습 과정에서 획득된 가중치 값으로 다층 퍼셉트론이 초기화되므로 제안된 방식은 초기 학습 데이터베이스와 그 특성이 다르게 나타나는 재학습 데이터만을 선택적으로 학습에 적용하거나 출력층의 에러를 효과적으로 감소시킬 수 있는 가중치만을 선택적으로 재학습한다. 본 논문에서는 제안된 선택적 주의 학습 방법을 음성 인식 시스템에서의 화자 적응 문제에 적용하여 그 유효성을 알아보았다. 화자 적응은 재학습 문제의 대표적인 예로서 화자 독립적인 인식시스템에 기반하여 적은 양의 특정 화자 데이터만으로 효율적인 화자 종속 인식시스템을 구축하는 학습 과정이다[6]. 실험에서는 제안한 방식을 적용하여 화자 적응을 위한 재학습 속도를 가속화하고자 한다.

II. 선택적 주의 학습

에러 역전파 알고리듬에 기반한 다층 퍼셉트론의 학습 과정은 평균 자승 에러를 계산하는 전방향 과정 (forward pass)과 그 에러를 감소시키기 위해 가중치 값들을 수정하는 후방향 과정 (backward pass)으로 크게 나눌 수 있다. 제안된 선택적 학습 방식은 주어진 세 가지 선택적 주의 기준에 따라 후방향 과정에 관련된 계산량을 감소시키는데 주안점을 두고 있다. 세 가지 선택 기준에 대한 자세한 내용은 다음과 같다.

2.1. 학습 데이터에 대한 선택적 주의 학습

에러 역전파 알고리듬은 주어진 학습 데이터에 대해 출력층에서의 에러값인 평균 자승 에러를 최소화시키는 방향으로 가중치 값을 변경하는 학습 과정을 반복적으로 수행한다. 일반적으로 학습에 사용되는 평균 자승 에러 함수는 다음과 같이 정의된다

$$E(n)=\frac{1}{2}\sum_j e_j^2(n)=\frac{1}{2}\sum_j (t_j(n)-o_j(n))^2 \qquad (1)$$

여기서 $o_j(n)$은 n번째 학습 반복 과정에서 출력층의 j번째 노드의 실제 출력값을 나타내며 t_j는 이에 대응되는 목표값을 나타낸다. 다층 퍼셉트론 내의 각 가중치, w_{ji}는 식 (1)에서 정의된 평균 자승 에러를 최소화 하기 위해 가중치 변환값, $\triangle w_{ji}(n)$을 이용하여 반복적으로 수정된다.

$$\triangle w_{ji}(n)=-\frac{\partial E(n)}{\partial w_{ji}(n)} \qquad (2)$$

식 (2)로부터, 주어진 입력에 대해 출력층에서의 에러가 0에 매우 근접한 값일 경우에는 다층 퍼셉트론의 각 가중치 값들이 거의 변경되지 않음을 알 수 있다. 따라서,

학습 데이터에 대한 선택적 주의는 이러한 입력 패턴들을 학습 과정에서 제외함으로써 전체 학습 시간을 단축시키는 방법이다.

2.2. 노드에 대한 선택적 주의 학습

노드에 대한 선택적 주의 학습은 학습 시간의 단축을 위해 다층 퍼셉트론의 전체 영역이 아닌 일부 영역만을 학습에 적용하는 방법이다. 이를 위해 다층 퍼셉트론의 각 가중치에 대한 변환값을 아래와 같이 재정의한다.

$$\triangle w_{ji}(n) = \delta_j(n)\, o_i(n) \tag{3}$$

$\delta_j(n)$은 각 노드에서의 국부 경사도 (local gradient)를 의미한다.

$$\begin{aligned}\delta_j(n) &= -\frac{\partial E(n)}{\partial e_j(n)}\frac{\partial e_j(n)}{\partial o_j(n)}\frac{\partial o_j(n)}{\partial v_j(n)} \\ &= e_j(n)\,\varphi_j'(v_j(n))\end{aligned} \tag{4}$$

여기서 $e_j(n)$과 $v_j(n)$은 출력층의 j번째 노드에서의 국부 에러값과 순 (net) 출력을 각각 나타내며 $\varphi(\cdot)$은 시그모이드 함수를 의미한다.

위의 식 (3)과 (4)로부터, 출력층 노드의 국부 에러값, $e_j(n)$이 0에 매우 가까운 값일 경우에는 그 노드에 연결된 가중치 값이 거의 조정되지 않음을 알 수 있다. 따라서 주어진 입력 패턴에 대해 다층 퍼셉트론 출력층의 각 노드 별로 국부 에러값을 계산하여 그 에러가 크게 발생하는 노드에 연결된 가중치만을 선택적으로 학습에 적용함으로써 학습에 소용되는 전체 계산량을 크게 줄일 수 있다. 이 개념은 은닉층 노드에 대해서도 쉽게 적용할 수 있다. 은닉층의 경우에는 각 노드에서의 국부 에러값을 직접 알 수는 없으나 출력층에서 발생한 에러값을 이용하여 그 값을 쉽게 추정할 수 있다.

$$\delta_j(n) = -\sum_k \delta_k(n) w_{kj}(n)\,\varphi_j'(v_j(n)) \tag{5}$$

여기서 $\delta_k(n)$은 출력층의 k번째 노드에서의 국부 경사도를 나타낸다.

전술한 학습 데이터 및 노드에 대한 선택적 주의 학습은 에러 역전파 알고리듬 상에서 쉽게 구현할 수 있음에도 불구하고 학습의 반복 횟수를 단축시키지는 않으므로 인해 그 동안 관련 분야에서 널리 적용되지는 않았다. 그러나 비교적 적은 반복 횟수만으로 학습이 완료되는 화자 적응과 같은 재학습 문제에서는 매우 효과적인 학습 기법이 될 수 있다.

2.3. 클래스 의존적인 적합도 (relevance)에 기반한 선택적 주의 학습

적합도 (Relevance) [5]는 주어진 입력에 대해 적절한 출력값을 얻기 위한 인식 과정에서 다층 퍼셉트론 내의 각 은닉층 노드에 대한 중요도를 측정하기 위해 사용된다. 예를 들어, 은닉층의 m번째 노드에 대한 적합도 (relevance), R_m은 그 노드를 인식 과정에서 제외하였을 때 출력층에서의 평균 자승 에러값의 변화량으로 정의된다.

$$\begin{aligned}R_m = &\sum_{p=1}^{P}\sum_{j=1}^{N}\left\{t_j^p - \varphi\left(\sum_{i=0}^{M} w_{ji}\rho_{mi}v_i^p\right)\right\} \\ &- \sum_{p=1}^{P}\sum_{j=1}^{N}(t_j^p - o_j^p)^2,\quad \rho_{mi} = \begin{cases} 0 & \text{if } m=i \\ 1 & otherwise \end{cases}\end{aligned} \tag{6}$$

따라서 주어진 노드의 적합도 (relevance)가 큰 경우에는 그 노드가 인식 과정에서 매우 중요한 역할을 한다는 것을 의미한다. 이러한 적합도 (relevance) 개념은 주로 신경회로망의 규모 축소 (pruning) 문제와 관련된 연구 분야에서 널리 사용되어 왔다[7,8].

본 논문에서는 전술한 적합도 (relevance) 개념을 화자 적응을 위한 선택적 주의 학습에 적용하기 위해 은닉층 노드의 적합도 (relevance)를 학습 데이터의 각 클래스별로 따로 분리하여 계산하는 방식을 제안한다. 예를 들어, ω클래스에 대한 은닉층의 m번째 노드에 대한 적합도 (relevance), $R(\omega)_m$은 다음과 같이 나타낼 수 있다

$$\begin{aligned}R(\omega)_m = &\sum_{p\in\omega}\sum_{j=1}^{N}\left\{t_j^p - \varphi\left(\sum_{i=0}^{M} w_{ji}\rho_{mi}v_i^p\right)\right\} \\ &- \sum_{p\in\omega}\sum_{j=1}^{N}(t_j^p - o_j^p)^2\end{aligned} \tag{7}$$

따라서 클래스 의존적인 적합도 (relevance)에 기반한 선택적 주의 학습은 주어진 입력 패턴에 대해 해당되는 클래스의 은닉층 각 노드별 적합도 (relevance) 값을 검사하여 그 중요도가 높은 노드에 연결된 가중치만을 선택적으로 학습에 적용하는 방법이다. 한편, 화자 적응을 위한 재학습 실험에서 각 클래스별 은닉층 노드의 적합도 (relevance)는 초기 학습 과정에서 미리 계산된다.

III. 화자 적응 실험

본 논문에서는 다층 퍼셉트론에 기반한 음성 인식시스템에서의 화자 적응 문제에 대해 제안한 선택적 주의 학습 방법을 적용하여 그 유효성을 알아보았다. 화자 적응은 화자 독립 또는 화자 종속 음성 인식시스템에서 음성의 화자 간 변질성 (inter-speaker variability) 또는 화자 내 변질성 (intra-speaker variability)으로 인한 인식 성능의 저하를 방지하기 위한 기법으로서 그 연구가 최근 들어 활발히 이루어지고 있다[6,9]. 그러나 화자 적응 기법을 실제 응용 분야에 적용하기 위해서는 인식 성능의 향상 뿐만 아니라 적응 시간을 단축시키는 문제도 중요하게 고려되어야 한다. 본 논문에서는 빠른 화자 적응을 구현하기 위해 인식시스템의 재학습 속도를 가속화시키는데 주안점을 두고 있다.

3.1. 데이터베이스 및 특징 추출

본 논문에서는 먼저 20개의 고립 단어를 인식하기 위한 기본 음성 시스템을 구축한 후에 새로운 화자의 적응 과정을 수행하면서 제안한 학습 방식의 유효성을 알아보았다. 실험에서는 같은 지역에 거주하는 10명의 화자 (남녀 각 5명)로부터 20개의 고립 단어를 10번 발음하여 획득한 총 2000개의 음성 데이터를 다층 퍼셉트론의 초기 학습을 위한 학습 데이터베이스로 사용하였다. 음성 데이터 획득은 비교적 주변 잡음이 적은 연구실에서 이루어졌으며 각 음성은 16KHz 샘플링 과정과 16bit 양자화 과정을 거쳐 입력 데이터로 사용된다. 실험에 사용된 20개의 고립 단어는 표 1에 나타낸 바와 같이 은행 업무에 관련된 숫자음과 화폐 단위음으로 구성되어 있다.

다층 퍼셉트론 기반의 음성 시스템에 대해 학습 및 인식 실험을 수행하기 위해서는 먼저 신호의 길이에 관계없이 동일한 크기의 입력 특징 벡터를 음성 데이터로부터 추출하는 것이 필요하다. 실험에서는 각 음성 데이터를 가변적인 폭을 가지는 19개의 프레임으로 나눈 후에 Hamming 창을 반 프레임씩 겹치도록 적용하여 생성된 총 37개의 프레임으로부터 각각 21차의 mel-scale filter bank 계수[10]를 추출하였다. 결과적으로 각 음성 데이터로부터 그 길이에 관계없이 차원이 777인 특징 벡터를 추출하여 다층 퍼셉트론의 입력으로 사용하였다.

표 1. 실험에 사용된 20개의 고립단어
Table 1. Isolated word set used in our experiment.

단어 구성	고립 단어
숫자음	'영', '일', '이', '삼', '사', '오', '육', '칠', '팔', '구'
화폐 단위음	'공', '십', '백', '천', '만', '억', '조', '원', '정', '금'

화자 적응 실험에서는 초기 학습 과정에 참여하지 않은 화자로부터 각 단어별 10개의 음성 데이터를 획득하여 재학습용 데이터베이스로 사용하였다. 또한 적응된 인식시스템에 대한 성능 검사는 적응 화자로부터 새로이 얻은 총 200개의 음성 데이터를 이용하여 수행된다.

3.2. 초기 인식시스템 구현

본 논문에서는 1개의 은닉층을 가지는 다층 퍼셉트론을 사용하여 초기 인식시스템을 구현하였다. 이때 입력층 노드의 수는 입력 특징 벡터의 차원과 동일한 777개이며 출력층 노드의 수는 인식하려는 단어의 수와 동일한 20개를 사용하였다. 또한 은닉층 노드의 수는 다층 퍼셉트론의 일반화 (generalization) 특성 및 학습 속도 등을 고려하여 try-and-error 방식에 의해 35개로 설정하였다. 학습 과정에서 학습률 및 관성항의 값은 각각 0.05와 0.8로 설정하였으며 화자 적응 결과의 정확한 분석 및 비교를 위해 실험 과정 동안 그 값을 변경하지 않았다. 초기 학습에서 다층 퍼셉트론 내의 각 가중치는 $[-5\times10^{-3}, 5\times10^{-3}]$ 사이의 랜덤한 값으로 초기화되며 학습은 평균 자승 에러가 0.001 될 때까지 반복적으로 수행된다. 한편 본 논문에서 수행된 모든 학습 및 인식 실험은 Pentium-III 500 PC 시스템의 리눅스 환경하에서 이루어졌다.

3.3. 화자 적응 실험

본 논문에서 제안한 세 가지 선택적 주의 기준은 학습 과정에서 각기 별도로 적용되어 사용될 수도 있으나 이를 통합하여 적용할 경우에 좀더 효과적으로 학습 시간을 단축시킬 수 있을 것으로 예상된다. 따라서 실제 화자 적응 실험에서는 아래에 기술된 세 가지 재학습 전략에 따라 학습을 수행하여 그 결과를 비교하였다.

1) 방법1: 학습 데이터에 대한 선택적 학습 적용
2) 방법2: 학습 데이터 및 노드에 대한 선택적 학습 통합 적용
3) 방법3: 학습 데이터 및 노드에 대한 선택적 학습, 그리고 클래스 의존적인 적합도 (relevance)에 기반한 선택적 주의 학습 통합 적용

재학습 실험에서 학습률과 관성항은 초기 학습과 동일한 값으로 설정되며 다층 퍼셉트론은 초기 학습 과정에서 획득된 가중치 값으로 초기화된다. 실험에서는 전술한 세 가지 방법에 기반한 적응 학습을 각각 10회 실시하여 그 결과의 평균을 비교 평가하였다. 또한 제안한 학습 방

표 2. 기본 에러 역전파 학습 알고리듬에 의한 적응 재학습 결과
Table 2. Adaptation results by standard EBP learning algorithm.

화자	인식율 (%)	반복횟수	학습시간 (sec)
A speaker	98.6	17.7	11.30
B speaker	100.0	55.4	35.32
C speaker	98.9	15.2	9.72
D speaker	98.3	5.3	3.38
E speaker	99.9	32.0	20.41

표 3. 세 가지 선택적 주의 학습 전략에 대한 적응 재학습 결과
Table 3. Adaptation results by the three selective attention methods.

화자	방법 1 (sec)	방법 2 (sec)	방법 3 (sec)
A speaker	5.03	4.05	3.69
B speaker	14.38	11.61	11.13
C speaker	3.63	3.10	2.85
D speaker	1.10	0.98	0.93
E speaker	7.83	6.97	6.43

식의 유효성을 좀더 정확하게 검증하기 위해 5명의 화자에 대해 적응 실험을 반복 수행하였다.

실험에서는 선택적 주의 학습의 성능 평가를 위한 기준으로 사용하기 위해 에러 역전파 알고리듬만을 적용한 재학습 결과를 먼저 알아보았다. 표 2에 나타난 결과에서 각 화자 별로 재학습을 통해 높은 인식 성능을 얻을 수 있지만 학습 시간이 다소 많이 소요되어 실시간 재학습이 요구되는 응용 분야에는 그 적용이 부적절함을 알 수 있다.

표 3에서는 선택적 주의 학습 방식에 의한 적응 실험 결과를 나타내었다. 표 3의 결과로부터 제안한 선택적 학습 방법이 기본 에러 역전파 학습 알고리듬에 비해 전체 학습 시간을 크게 줄일 수 있음을 확인할 수 있다. 특히 세 가지 선택적 학습 방식들을 통합하여 사용한 방법 3의 경우에 가장 효과적으로 학습 시간을 단축 시킬 수 있음을 알 수 있다.

그림 1에서는 5명의 화자에 대해 각 선택적 학습 방법의 기본 에러 역전파 학습 방식 대비 학습 시간의 감소율을 나타내었다. 표 3의 결과와 마찬가지로 방법 3을 적용한 경우에 기본 학습 방식에 비해 평균적으로 60% 이상의 학습 시간 감소 효과를 얻을 수 있음을 알 수 있다.

본 논문에서는 다층 퍼셉트론의 학습 시간을 단축시키기 위해 학습의 반복 횟수를 줄이기 보다 가중치 값의 변경과 관련된 한 반복 횟수 당 계산량을 감소시키는데 주안점을 두고 있다. 따라서 최종 실험에서는 주어진 화자

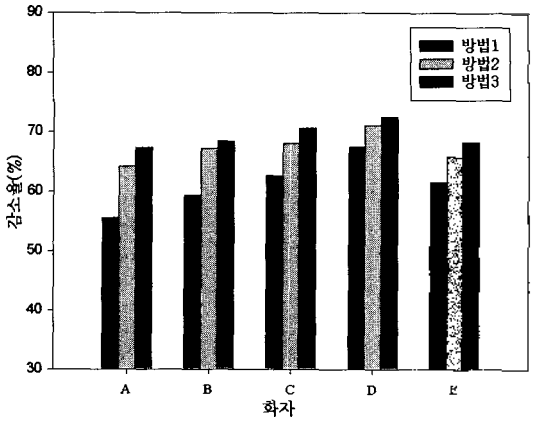

그림 1. 각 화자별 선택적 학습 방법에 따른 기본 에러 역전파 학습 방식 대비 학습 시간 감소율
Fig. 1. Reduction ratios of the three selective attention methods compared to standard EBP algorithm for five speakers.

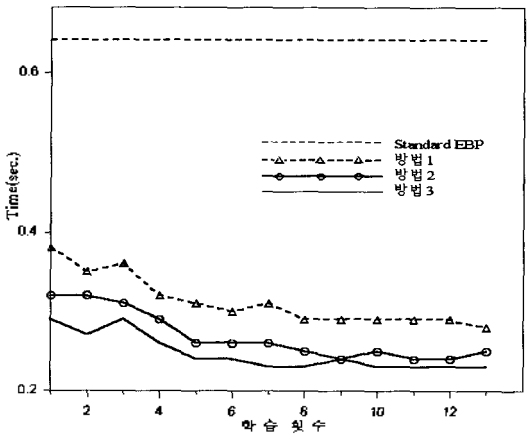

그림 2. 적응 재학습 과정에서의 한 반복 횟수 당 학습 시간 (C 화자)
Fig. 2. Learning time per iteration during adaptation process (C speaker).

에 대해 적응 재학습의 각 방법에 따른 한 반복 횟수 당 학습 시간을 계산하여 그 결과를 비교하였다. 이를 위해 먼저 각 방법 별로 동일한 학습 횟수를 보이는 결과들을 추출하였다. 실험에서는 C 화자에 대해 각 방법 별로 13회의 동일한 학습 횟수를 가지는 학습 결과들을 각각 얻을 수 있었으며 이들로부터 한 반복 횟수 당 학습시간을 계산하여 그 결과를 그림 2에 나타내었다. 기본 학습 방식은 한 학습 횟수를 수행하는데 약 0.64초가 소요되며 학습 과정 동안 그 시간이 일정하게 유지됨을 알 수 있다. 그러나 선택적 주의 학습의 경우에는 학습이 진행되는 과정에서 학습에 영향을 미치지 않는 학습 데이터 및 가중치에 대한 학습을 적절히 생략함으로써 한 학습 횟수 당 학습 시간이 점점 더 감소함을 확인할 수 있다. 특히 방법 3의 경우에는 한 학습 횟수 당 학습 시간이 가장 크

게 감소함을 알 수 있다.

IV. 결 론

본 논문에서는 다층 퍼셉트론 기반의 음성 인식시스템에서 빠른 화자 적응을 구현하기 위해 선택적 주의 학습 방식을 제안하였다. 제안된 방식은 다층 퍼셉트론의 출력층에서 계산된 평균 자승 에러와 은닉층 노드에 대한 클래스 의존적인 적합도 (relevance)를 이용하여 학습 데이터베이스 내의 특정한 학습 데이터만을 선별적으로 입력 패턴으로 사용하거나 주어진 입력 패턴에 대해 신경회로망의 특정 부분만 선택적으로 학습이 이루어지도록 함으로써 에러 역전파 알고리듬의 학습 속도를 가속화시킨다.

제안된 선택적 주의 방식은 학습 반복 횟수를 줄이고자 하는 기존의 일반적인 학습 방식과는 달리 한 반복 횟수당 학습 계산량을 크게 줄임으로써 전체 학습 시간을 효과적으로 단축시키고자 한다. 이 방법은 표준 에러 역전파 알고리듬뿐만 아니라 다른 형태의 개선된 학습 알고리듬과도 손쉽게 결합되어 운용될 수 있다는 장점을 가진다. 5명의 화자에 대한 적응 재 학습 실험 결과로부터 제안한 방법이 학습 속도를 평균 60%이상 개선시킬 수 있음을 확인하였다.

감사의 글

본 연구는 한국과학재단 목적기초연구 (1999-2-303-001-3) 지원으로 수행되었습니다.

참 고 문 헌

1. D. E. Rumelhart and J. L. McClelland, Parallel Distributed Processing, Cambridge, MA, MIT Press, 1986.
2. R. A. Jacobs, "Increased rates of convergence through learning rate adaptation," Neural Networks, Vol. 1, pp. 295-307, 1988.
3. S. E. Fahlman, "Faster-learning variations on backpropagation: An empirical study," Proc. Connectionist Models Summer School, Carnegie Mellon University, 38-51, 1988.
4. S. H. Oh, "Improving the error backpropagation algorithm with a modified error function," *IEEE Trans.* Neural Networks, Vol. 8, pp. 799-803, 1997.
5. M. C. Mozer and P. Smolensky, "Using relevance to reduce network size automatically," Connection Science, Vol. 1, No. 1, pp. 3-16, 1989.
6. C. H. Lee, C. H. Lin, and B.H. Juang, "A Study on speaker adaptation of the parameters of continuous density hidden Markov models," *IEEE Trans. Signal Processing*, Vol. 39, No. 4, pp. 806-814, 1991.
7. S. E. Erdogan, G. S. Ng, and K. H. C. Patrick, "Measurement criteria for neural network pruning," *Proc. IEEE TENCON-Digital Signal Processing Applications*, Vol. 1, pp. 83-89, 1996.
8. R. Audouin and L. Shastri, "Incremental character recognition with feature attribution," Proc. International Conference on Document Analysis and Recognition, Vol. 2, pp. 837-840, 1995.
9. 황영수, "화자 적응 방법들의 비교," 한국음향학회지, 제18권, 제1호, pp. 61-66, 1999.
10. S. B. David and P. Mermelstein, "Comparison of parametric representation for monosyllabic word recognition in continuously spoken sentences," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. 28, No. 4, pp. 357-366, 1980.

저자 약력

● 김 인 철 (In Cheol Kim)
한국음향학회지 제19권 2호 참조

● 진 성 일 (Sung Il Chien)
한국음향학회지 제19권 2호 참조