

# 상태레벨 공유를 이용한 HM-Net 적응화 시스템의 성능평가에 관한 연구

오 세 진, 김 광 동, 노 덕 규, 황 철 준\*, 김 범 국\*, 김 광 수\*\*, 성 우 창\*\*\*, 정 현 열\*\*\*  
한국전문연구원, \*대구과학대학, \*\*경운대학교, \*\*\*영남대학교  
전화 : 042-865-3280 / 핸드폰 : 011-543-0971

## A Study on Performance Evaluation of HM-Net Adaptation System Using the State Level Sharing

Se-Jin Oh, Kwang-Dong Kim, Duk-Gyoo Roh, Chul-Jun Hwang\*, Bum-Koog Kim\*,  
Kwang-Soo Kim\*\*, Woo-Chang Sung\*\*\*, Hyun-Yeol Chung\*\*\*  
KAO, \*Taegu Science College, \*\*Kyungwoon Univ, \*\*\*Yeungnam Univ.  
E-mail : sjoh@trao.re.kr

### Abstract

본 연구에서는 HM-Net(Hidden Markov Network)을 다양한 테스트에의 적용과 화자의 특성을 효과적으로 나타내기 위해 HM-Net 음성인식 시스템에 MLLR(Maximum Likelihood Linear Regression) 적응방법을 도입하였으며, HM-Net 학습 알고리즘을 개량하여 회귀클래스 생성방법을 제안한다. 제안방법은 PDT-SSS(Phonetic Decision Tree-based Successive State Splitting) 알고리즘의 문맥방향 상태분할에 의한 상태레벨 공유를 이용한 방법으로 새로운 화자로부터 문맥정보와 적응화 데이터의 발생 양에 의존하여 결정된 많은 적응 파라미터들을(평균, 분산) 자유롭게 제어할 수 있게 된다. 제안방법의 유효성을 확인하기 위해 국어공학센터(KLE) 452 음성 데이터와 항공편 예약관련 연속음성을 대상으로 인식실험을 수행한 결과, 전체적으로 음소인식의 경우 평균 34~37%, 단어인식의 경우 평균 9%, 연속음성인식의 경우 평균 78%의 인식성능 향상을 각각 보였다. 또한 적응화 데이터의 양에 따른 인식성능 비교에서, 제안방법을 적용한 인식 시스템이 적응 데이터의 양이 적은 경우에도 향상된 인식률을 보였으며, 잡음을 부가한 음성에 대한 적응화 실험에서도 향상된 인식성능을 보여 MLLR 적응방법의 특성을 만족하였다. 따라서 MLLR 적응방법을 도입한 HM-Net 음성인식 시스템에 제안한 회귀클래스 생성방법이 유효함을 확인할 수 있었다.

### I. 서론

음성인식에 있어서 불특정 화자 음향모델의 구조에는 다음과 같은 두 가지가 요구된다. 첫 번째는 학습 데이터를 효율적으로 표현하는 모델의 구조이다. 최근 음성 데이터의 확충에 따라 불특정 화자 음향모델의 작성에 대량의 학습 데이터가 사용되고 있지만 통상적으로 제한된 양의 데이터라고 할 수 있다. 따라서 이것을 가능한 효과적으로 사용하고 보다 높은 성능을 얻을 수 있는 모델의 구조를 준비할 필요가 있다. 두 번째는 소량의 데이터에도 적응화(화자, 잡음, 발생방식 등)의 수행이 쉬운 모델의 구조가 필요하다. 예를 들어, 불특정 화자 음향모델은 특정화자의 음성 데이터를 가지고 화자적응화를 수행하는 경우가 많이 있다. 고속 화자적응화를 실현하기 위해서는 소량의 음성 데이터에도 모델 전체를 목적으로 하는 화자의 특성에 적용할 수 있는 모델 구조, 즉 적응에 유리한 구조를 준비하는 것도 필요하다.

가장 이상적인 HMM을 모델의 구조 관점에서 생각해보면 이는 최소한의 자유도를 가지는 모델이라고 할 수 있는데, 이를 위해 적절한 제약조건을 주어야 하며 이것을 모델에 대해 적절한 정도로 주어지는 것이 중요하다.

모델에 제약을 주는 한가지 방법으로 모델의 구조를 공유하는 것이다. 공유는 통계적인 성질이 유사한 여러 개의 파라미터 사이에 결합(tying) 관계를 가지도록 하는 것을 말한다. 공유의 이점은 모델의 학습효율을 높일 수 있다는 것이다. 공유된 파라미터의 학습 데이터는 서로 통합되고 통계적으로 보다 신뢰성이 높은 파라미터 추정을 수행할 수 있게 된다. 따라서 공유를 통해 모델의 자유도를 줄이게 되므로 학습효율이 높아지고 강건한 인식이 가능한 모델을 작성할 수 있게 된다. 최근 모델의 공유구조는 파라미터 사이의 거리에 근거하여 수행되는 경우가 많이 있는데, 그 중에서 음소환경 클러스터링(Phoneme Environment Clustering: PEC)[1]과 일반화된 triphone 모델[2] 등의 모델레벨의 공유화, HM-Net과 같은 상태레벨의 공유화[3], tied-mixture HMM(또는 semi-continuous HMM)[4] 등의 분포레벨의 공유화가 제안되고 있다.

본 연구에서는 상태레벨의 공유방법인 HM-Net을 도입하였으며, 음향모델의 학습에는 PDT-SSS 알고리즘[7]을 사용하였다. 이 알고리즘은 SSS 알고리즘[3]의 장점과 음소결정 트리를 결합한 것으로 음소결정트리에 기반한 상태 클러스터링으로부터 상태위치에 대한 독립성을 배제하고 문맥방향과 시간방향의 상태분할을 수행하여 정밀한 문맥의존 HM-Net

음향모델의 구조를 결정할 수 있다. 즉, 미리 작성해 둔 문맥 의존 음향모델의 각 상태위치마다 음소 질의어 집합에 의해 음소결정트리를 생성하고, PDT-SSS 알고리즘에 의해 문맥 의존 음향모델의 상태열을 다시 학습하는 방법이다. 하지만, PDT-SSS 알고리즘에 의한 HM-Net은 특정 태스크에서는 우수한 성능을 보이지만 태스크가 바뀔 경우 인식을 저하를 초래한다. 이를 해결하기 위한 방법으로는 모델의 학습과정에서 PBW(Phoneme Balance Word)나 다양한 문맥정보(음소환경)를 포함한 많은 양의 음성 데이터를 이용한다면 가능하다. 하지만 음성 데이터의 수집과 모델의 학습에 많은 시간과 경비가 소요된다.

따라서 본 연구에서는 HM-Net의 단점을 보완하기 위해 적용화 방법을 HM-Net 시스템에 채용하고, HM-Net 시스템에서 상태레벨의 공유방법의 PDT-SSS 알고리즘을 MLLR[5] 적용방법에 이용하여 회귀 클래스의 생성방법을 제안하고자 한다. 특히, PDT-SSS 알고리즘의 문맥방향 상태분할한 회귀 클래스의 문맥상의 정보를 효율적으로 표현하기 위해 사용하였다. 제안방법의 유효성을 확인하기 위해 한국전자통신연구원(ETRI)과 국어공학센터(KLE)의 음성 데이터베이스를 대상으로 음소 및 단어인식 실험을 수행하고자 한다. 화자독립 기본모델은 HM-Net 모델로서 ETRI에서 제공된 대어휘 음성 데이터베이스 중 400명의 화자(남녀 각 200명)를 사용하여 문맥의존 음향모델을 작성한다. 작성한 모델을 이용하여 상태공유방법에 의해 제안한 회귀클래스 생성방법을 적용한 화자적응화는 MLLR 적용방법에 의해 KLE 남성 35명과 여성 32명의 두 번째 발성으로 적용화하고 첫 번째 발성으로 평가한다. 또한 적용 데이터의 변화에 따른 인식성능 변화를 고찰하기 위해 동일한 데이터에 대해 적용화 실험을 수행한다. 또한 연속음성에서의 적용화 실험 결과를 고찰하기 위해 본 연구실에서 채록한 항공편 예약관련 200문장 중 남성 12명이 1회 발성한 문장에 대해 각 화자별 소규모 적용화 문장을 이용하여 적용화 학습을 수행하고 학습에 참여하지 않은 문장을 평가에 사용한다.

## II. MLLR 화자 적응화

MLLR 적용방법[5]은 원래의 모델과 적용 데이터 사이에 불일치를 줄일 수 있는 변환행렬을 이용하는 것이다. 즉, Gaussian mixture HMM 시스템의 평균과 분산 파라미터들에 대한 선형 변환행렬 집합을 추정하는 것이다. 새롭게 적용된 평균을 구하는데 사용되는 변환행렬은 식 (1)과 같이 나타낼 수 있다.

$$\mu_i = A_c \zeta_i + b_c \quad (1)$$

여기서, 행렬  $A_c$ 와 바이어스 항인  $b_c$ 는 회귀 클래스  $c$ 의 MLLR 파라미터들이다. 그리고  $\zeta_i$ 는 원래 화자독립 모델의 평균 벡터를 나타낸다. MLLR 파라미터  $A_c$ 와  $b_c$ 는 클래스  $c$ 가 최대로 되는 적용 데이터의 우도에 의해 추정된다.

회귀 클래스 트리는 회귀 클래스의 계층 또는 잎(leaf)과 같은 기본 클래스의 집합으로 구성된다. 트리 노드 아래의 모든 기본 클래스들은 변환행렬을 공유하게 된다. 서로 다른 변환행렬들의 수는 사용할 수 있는 적용 데이터의 양에 따라 선택

된다. 서로 다른 변환 행렬들의 수는 이용 가능한 적용 데이터의 양에 따라서 선택되는데, 만약 매우 적은 양의 적용 데이터가 존재한다면, 트리 루트에서의 계산된 단일 변환행렬은 모든 모델에 적용된다. 좀 더 많은 적용 데이터를 사용할 수 있다면 트리의 크기는 커지게 되며 좀더 특정한 변환행렬을 계산할 수 있게 된다.

기본 클래스로부터 회귀클래스로의 분할은 최종 회귀클래스 트리가 적용 데이터의 우도를 최대로 한다던 최적이라고 간주할 수 있다. 문헌 [6]에서 제안한 반복 방법은 각 트리의 분할에 국부적인 우도를 최적화 함으로써 수학적으로 계산할 수 없는 최적 해를 추정하는 것이다 이 알고리즘은 음향학적 분할로부터 얻어진 초기 회귀클래스 트리에서 시작하게 된다.

## III. 회귀클래스 생성방법

본 연구에서는 HM-Net의 단점을 보완하고 적용학습에 널리 사용되고 있는 MLLR 적용방법을 HM-Net에 적용하기 위해 음향모델 학습에 사용된 PDT-SSS 알고리즘의 문맥방향 상태분할 과정을 개량하여 회귀클래스 트리를 생성하도록 하였다.

우선, 음소결정트리는 음소의 음향학적 변동을 확인하기 위한 것으로 출현하지 않는 문맥의 음향학적 특성을 예측하는 방법으로 트리구조 형태의 뿌리를 문맥에 독립한 단위로 2진 트리로 표현하고 뿌리에서부터 잎 방향으로 문맥 클래스의 분할을 수행한다. 이 트리는 뿌리에서 잎으로 진행함에 따라 문맥 의존성이 강한 단위를 나타내는 계층적인 구조를 가지며 일반적으로 잎의 끝 부분에 모델을 대응시킨다. 트리의 각 접점(상태)에는 경험적으로 음소의 유사성에 기반한 질문을 하게 되며, 대답으로 yes와 no에 따라 문맥 클래스를 두 개로 분할한다. 이러한 질문은 음소 환경요인과 음소집합을 대응시킨 것도 가능하다. 임의의 문맥 클래스도 트리의 뿌리에서부터 질문을 통해 반드시 잎에 대응되어야 하기 때문에 출현하지 않는 문맥에 대해서도 음향학적으로 가장 유사한 잎에 분류될 수 있다. 이를 위해 결정트리에 대한 질의어 집합을 구성하고, 분할될 접점에 대해 최적의 상태분할을 위해 어떤 질의어가 사용될 것인가를 결정해야 한다. 결정트리의 목적은 분할될 접점의 결정에 대한 불확실성을 줄이는 것이다. 아래와 같이 회귀클래스의 생성에 사용된 PDT-SSS 알고리즘도 마찬가지로 음소결정트리의 음소 질의어에 따라 트리의 루트로부터 각 접점에 대해 반복적으로 클러스터링을 수행하여 음소결정트리를 성장시키는 방법이다. 트리의 성장은 문맥독립 HMM의 상태위치에 따라 수행된다. 기본적인 개념은 모든 분할될 접점을 트리의 뿌리에 위치시키고, 상태분할은 분할 참조에 따라 최적의 질의어 집합에 의해 반복적으로 수행된 후, 각 잎에서 결정된 상태는 서로 공유하게 된다. 본 연구에서 제안한 회귀클래스 생성을 위해 이용한 PDT-SSS 알고리즘의 장점은 상태열이 음소 질의어 집합의 yes/no의 클러스터링에 따라 학습데이터에서 출현하지 않는 음소를 가지는 모든 환경에 대해서도 구성된다는 점이다. 이와 같은 방법으로 생성된 HM-Net은 문맥독립 HMM에서 문맥방향의 상태분할에 대해서만 수행된다. 따라서 이점에 착안하여 본 연구에서는 PDT-SSS 알고리즘의 문맥방향 상태분할을 이용하여 MLLR 적용화 방법의 회귀클래스 생성 방법을 다

음과 같이 제안한다.

- 1) 화자독립 학습데이터에 대해 문맥의존 음향모델 (triphone)과 상태에 대한 통계량을 학습시킨다.
- 2) 음성학적 규칙에 대한 적절한 질의어 집합을 작성하고 문맥방향의 상태분할만 있는 임의의 1 상태 HM-Net을 작성한다.
- 3) 학습데이터에 대해 식 (2)를 이용하여 최적의 분할할 상태를 선택한다.
- 4) 각 질의어에 대해 다음과 같은 문맥방향 상태분할을 수행한다.
  - a) 질의어에 대해 두 개의 단일 가우시안 분포(각 가우시안 분포는 yes/no와 관련)를 유도하고 분할될 문맥 클래스에 대해 상태분할을 수행한다.
  - b) 분할된 새로운 상태들에 각 문맥클래스와 각 가우시안을 지정한다.
- 5) 단계 1에서 구한 통계파일과 미리 정의한 상수 ( $\epsilon$ )와의 비교에서 상수보다 적을 경우 회귀클래스 생성을 종료한다.
- 6) 질의어의 문맥클래스에 근거한 문맥방향에서 최적의 회귀클래스를 선택한다. 여기서, 단계 1의 통계량과 단계 4와 5의 단계를 통해 적응화 학습에 필요한 회귀클래스를 생성한다.

단계 3에서, 분할할 상태를  $S(i)$ 로 정의한다면, 분할할 상태의 선택은 입력 음성의 프레임 단위의 평균과 분산에 따라 다음 식을 사용하여 계산한다.

$$d_i = n_i \times \sum_k \frac{\sigma_{ik}^2}{\sigma_{Tk}^2} \quad (2)$$

$$\sigma_{ik}^2 = \lambda_{1i} \sigma_{1k}^2 + \lambda_{2i} \sigma_{2k}^2 + \lambda_{1i} \lambda_{2i} (\mu_{1k} - \mu_{2k})^2$$

여기서,  $K$ 는 파라미터의 차원,  $\lambda_{1i}$ ,  $\lambda_{2i}$ 는 상태  $i$ 의 가중계수,  $\mu_{1k}$ ,  $\mu_{2k}$ 는 상태  $i$ 의  $k$ 번째 평균,  $\sigma_{1k}^2$ ,  $\sigma_{2k}^2$ 는 상태  $i$ 의  $k$ 번째 분산,  $n_i$ 는 상태  $i$ 일 때의 학습샘플의 수, 그리고  $\sigma_{Tk}^2$ 는 모든 학습샘플의  $k$ 번째 분산을 각각 나타낸다.

## IV. 인식실험 및 고찰

### 4.1 음성 데이터 및 분석조건

화자독립 문맥의존 음향모델은 ETRI에서 제공한 음성데이터 중 400명(남200:여200)이 발성한 단음절, 숫자음, 단어, 연속음성 등을 사용하였다.

기본 HM-Net 모델의 초기구조는 43개의 문맥독립 음소모델을 병렬로 연결한 126개의 상태를 가지며, 최종 학습된 모델의 상태수는 1000, 1500, 2000개이며 혼합수는 1, 2, 4, 6, 8개이다. 이 중에서 적응화 학습에는 상태수 2000개 혼합수 4개인 모델을 사용하였다.

적응화 학습과 평가에 사용된 음성데이터는 단어인식인 경우 KLE에서 작성한 PBW의 452단어 중 남자 35명과 여자 32명의 2회 발성 중에서 두 번째 발성을 적응화 학습에 사용하고 첫 번째 발성은 평가에 사용하였으며, 연속음성인식인 경우 본 연구실의 항공편 예약관련 200문장(YNU200)을 사용하였다. 적응화 학습에서 학습 데이터의 양에 따른 MLLR 적응화의

특성을 고려하여 소규모 음성 데이터를 이용하였다.

모든 음성데이터는 16kHz, 16bits로 샘플링된 후  $1-0.97z^{-1}$ 의 필터를 통과시켰으며, 입력 음성의 각 프레임에 25ms의 해밍 윈도우를 곱하여 10ms 씩 이동하면서 분석하였다. 특징 파라미터는 12차 LPC-MEL 캡스트립 계수와 정규화된 대수 에너지에 1차 및 2차의 차분성분을 포함한 총 39차를 사용하였다.

## 4.2 적응화 학습 전후 음성인식 실험

### 4.2.1 음소/단어인식 실험

제안한 PDT-SSS에 의한 회귀클래스 트리 생성방법의 유효성을 확인하기 위해 소량의 적응화 데이터에 강한 MLLR 적응방법에 적용하여 KLE 452 단어를 대상으로 음소 및 단어에 대해 잡음을 부가하기 전과 후에 대해 적응화 학습을 수행한 후 인식실험을 수행하였다.

우선, 잡음이 없는 음성에 대한 음소/단어인식 실험은 KLE 452 단어에서 남성과 여성의 두 번째 발성 중 적응화 단어수를 MLLR알고리즘 특성상 소량의 적응데이터에 강하므로 452 단어 중 임의의 10단어를 적응화 학습한 후 첫 번째 발성을 평가에 사용하였으며, 그 결과를 그림 1에 나타내었다. 또한 평가화자의 전체 452 단어를 적응화 학습한 경우의 실험결과를 그림 2에 나타내었다. 그리고 잡음음성에 대한 평가는 KLE 남성 35명의 452단어에 대해서만 SNR 20dB인 잡음을 부가한 후 수행하였다. 잡음 부가 전과 후의 비교 단어인식 실험결과를 그림 3에 나타내었다.

그림 1에서 적응화 학습에 10단어만 수행한 경우 적응화 전에 비해 인식성능이 증가하는 것을 볼 수 있었다. 또한 그림 2의 452단어를 모두 적응화 학습한 경우 음소/단어인식률이 모두 증가함을 볼 수 있으며 특히 음소인식률이 적응화 학습 전에 비하여 평균 34.37%의 인식을 향상을 보였다. 따라서 이러한 결과를 바탕으로 MLLR 적응방법의 특성인 소규모 음성 데이터에 대한 특성을 만족하는 것을 확인할 수 있었다.

그리고 그림 3에 나타난 것과 같이 20dB의 잡음을 부가한 경우의 단어인식 실험결과에서도 적응화 전의 차이는 잡음의 영향으로 낮은 인식 결과를 보였으나, 적응화 10단어를 적응화 학습을 수행한 후 인식실험을 수행한 결과 적응화 전에 비해 향상된 인식성능을 보였다.

### 4.2.2 연속음성인식 실험

본 연구에서 제안한 방법이 연속음성에서 유효한 가를 확인하기 위해 본 연구실에서 작성한 항공편 예약관련 연속음성 200문장 중 남성 12명이 1회 발성한 문장에 대해 각 화자의 10문장을 적응화 학습한 후 나머지 문장을 평가하였다. 그림 4에 나타난 것과 같이 적응화 학습전의 인식률이 비교적 낮은 데이는 모델학습에 사용된 음성은 단어이므로 연속음성의 발성 형태가 다르기 때문이다. 또한 적응화 후의 인식률도 그리 많은 향상은 보이지 않는데 이는 적응화 문장에 포함된 음소열이 적응화에 충분하지 않기 때문에 모든 음소가 적응화되지 않은 원인으로 생각된다. 하지만, 본 연구에서 제안한 방법을 적용한 경우 각 화자의 10문장을 적응화 학습한 경우 적응화 전에 비해 인식성능이 7-8%의 인식을 향상을 보였다.

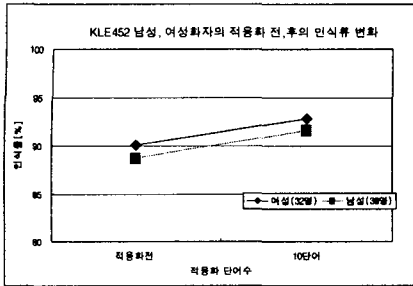


그림 1. 적응 데이터량 변화에 따른 단어인식률.

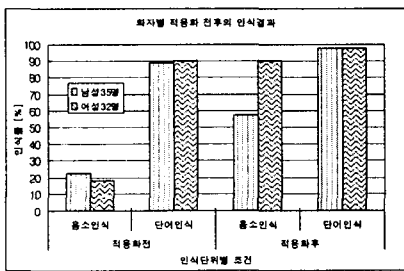


그림 2. 각 화자의 전체 452단어를 적응화 전후에 대한 인식결과.

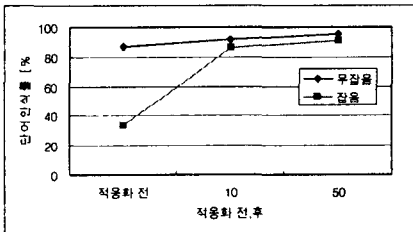


그림 3. 잡음(20dB)환경과 무잡음 환경에 대한 KLE 452 남성 35명의 적응화 전후의 단어인식률.

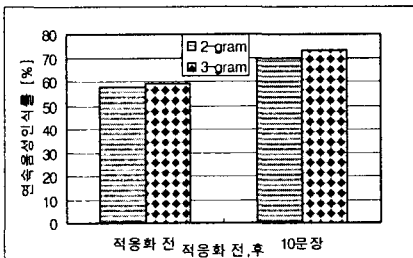


그림 4. 연속음성의 적응화 전후의 인식결과

### V. 결론

본 연구에서는 HM-Net을 다양한 태스크에의 적용과 화자의 특성을 효과적으로 나타내기 위해 HM-Net 음성인식 시스템에 MLLR 적응방법을 도입하였으며, HM-Net 학습 알고리즘을 개량하여 회귀클래스 생성방법을 제안하였다. 제안 방법은 PDT-SSS 알고리즘의 문맥방향 상태분할에 의한 상

태레벨 공유를 이용한 방법이다. 즉, 문맥방향의 각 상태에 적응화자 음성데이터에 포함된 문맥정보를 분할하여 적응화될 음소환경을 결정하는 것이다. 따라서 제안방법은 새로운 화자로부터 문맥정보와 적응화 데이터의 발생 양에 의존하여 결정된 많은 적응 파라미터들(평균, 분산) 자유롭게 제어할 수 있게 된다. 제안방법의 유효성을 확인하기 위해 국어공학센터(KLE) 452 음성 데이터와 항공편 예약관련 연속음성을 대상으로 인식실험을 수행하였다. 그 결과, 남성 35명과 여성 32명의 452단어에 대한 음소/단어인식실험에서, 음소인식의 경우 적응화 전 평균 22.8%, 18.1%를 각각 보였으나, 적응화 후 평균 57.4%, 56.0%를 각각 보여 평균 34~37% 인식률이 향상되었으며, 단어인식의 경우 적응화 전 평균 88.9%, 90.1%를 각각 보였으나, 적응화 후 평균 97.6%, 98.0% 인식률을 향상을 각각 보여 적응화 전에 비해 평균 9%의 성능향상을 보였다. 또한 적응 데이터의 양에 따른 인식성능 비교에서도 남성/여성화자 모두에서 적응 데이터의 양을 변화시킨 경우 적은 데이터에서도 향상된 인식성능을 보였다. 또한 연속음성인 경우 10문장을 적응화 학습한 경우 적응화 전에 비해 인식성능이 7-8%의 인식률 향상을 보였다. 이상의 실험결과를 바탕으로 MLLR 적응방법을 도입한 HM-Net 음성인식 시스템에 본 연구에서 제안한 회귀클래스 생성방법이 유효함을 확인할 수 있었다.

향후에는 다양한 문맥정보를 포함한 연속음성을 대상으로 연구를 수행할 예정이다.

### 참고문헌

- [1] S. Sagayama and S. Honma, "Estimation of Unknown Context Using a Phoneme Environment Clustering Algorithm," *Proc. of ICSLP90*, Vol. 1, pp. 361-364, 1990.
- [2] K.F.Lee and H.W.Hon, "Large-vocabulary speaker independent continuous speech recognition using HMM," *Proc. of ICASSP'88*, pp. 123-126, 1988.
- [3] J. Takami and S. Sagayama, "A Successive State Splitting Algorithm for Efficient Allophone Modeling," *Proc. of ICASSP'92*, 1:573-576, 1992.
- [4] S.Young and P.Woodland, "The use of state tying in continuous speech recognition," *Proc. of EUROSPEECH'93*, pp. 2203-2206, 1993.
- [5] M.J.E. Gales, "The Generation and Use of Regression Class Trees for MLLR Adaptation," *Technical Report CUED/F-INFENG/ TR263*, Cambridge University, 1996.
- [6] Reihold Haeb-Umbach, "Automatic Generation of Phonetic Regression Class Trees for MLLR Adaptation," *IEEE Trans. on SAP*, Vol. 9, No. 3, pp. 299-302, 2001.
- [7] T. Hori, M. Katoh, A. Ito, and M. Kohda, "A Study on HM-Nets using Decision Tree-based Successive State Splitting," *Proc. of ICSP'97*, Vol. 2, pp. 383-387, 1997.
- [8] 성우창, 오세진, 김광동, 정호열, 정현열, "결정트리 상태 클러스터링에 의한 MLLR 적응화의 회귀 클래스 생성에 관한 연구," 2003년도 한국음향학회 추계학술발표대회 논문집, 제22권 제1(호), pp. 121-124, 2003. 7.