

이산분포 HMM을 이용한 음성인식에서의 코드워드 Tying 알고리즘

A Codeword Tying Algorithm in Speech Recognition based on Discrete Hidden Markov Model

김도영*, 김남수*, 은종관*

(Do Yeong Kim*, Nam Soo Kim*, Chong Kwan Un*)

요 약

본 논문에서는 수형구조 분류기를 이용한 코드워드 tying 알고리즘을 제안한다. 코드워드와 상태간의 통계적 특성을 이용한 일종의 soft decision 방식이라고도 볼 수 있는 제안된 알고리즘은 빠른 트리 구성과 유일한 최적의 해를 제공하는 특징이 있다. 또한, 이산분포 hidden Markov model(HMM)을 이용한 인식시스템에 쉽게 적용이 가능하다는 장점을 가진다. 제안된 알고리즘의 성능 평가를 위한 화자독립 격리단어 인식실험에서 코드북 크기가 256과 512일 경우에 대해 각각 6%, 9%의 오차를 감소시켰으며, HMM 파라미터도 20% 정도 줄임을 확인하였다.

ABSTRACT

In this paper, we propose a new codeword tying algorithm based on a tree structured classifier. The proposed algorithm which can be viewed as a kind of soft decision using statistical properties between codewords and states has an advantage of fast construction, and guarantees a unique optimal solution. Also, it can easily be applied to any speech recognition system based on discrete hidden Markov model(HMM). Experimental results on speaker-independent isolated word recognition show error reduction of 6% for the codebook of size 256 and 9% for 512 size and also HMM parameter reduction of about 20%.

I. 서 론

Hidden Markov model(HMM)은 강력한 학습 알고리즘과 유연한 수학적 특성을 바탕으로 음성인식 분야에서 다양한 음성신호를 모델링하는데 성공적으로 이용되어 왔다. 각 모델은 음소나 단어와 같은 음성의 특징한 단위들을 나타내며, 몇 개의 상태(state)

와 상태들간의 천이로 구성된다. 상태의 출력 확률밀도 함수는 각 음성단위의 독특한 성질을 나타내며, 천이확률은 다양한 음향학적 현상들의 시간에 대한 변화와 밀접한 관련이 있다. 출력 확률밀도 함수를 이산분포로 이용하는 이산분포 HMM은 음성신호를 벡터 양자화 과정을 통해 코드워드로 바꾸어 주어 학습 및 인식하는 방식이다.

정확한 음성모델링을 위한 HMM 모델의 개선에 관한 시도는 궁극적으로 파라미터의 수를 늘리고, 시스템의 복잡도를 증가시키게 된다. 그런데, 한정된 학습데이터에 대해 과도하게 세분화된 모델은 파라미터의 통계적 신뢰도를 저하시켜, 오히려 좋지 못한

*한국과학기술원 전기 및 전자공학과 통신연구실
Communications Research Laboratory Department of
Electrical Engineering Korea Advanced Institute of
Science and Technology
접수일자: 1993년 3월 3일

결과를 초래하기도 한다. 따라서, 제한된 데이터로부터 신뢰도있는 통계치를 얻기 위한 연구에 관심이 모아지게 되는데, 이러한 문제에 대한 해결방식의 하나로 파라미터 tying 기법을 꼽을 수 있다. 파라미터 tying이란 통계적으로 또는, 인간의 지식이나 경험에 기반하여 유사한 성질을 가지는 단위들을 묶어줌으로써, 각 모델들에 대한 통계적 신뢰도를 높여주는 데이터 공유 기법이다. Lee 등은 인식단위의 선정에 이러한 기법을 적용하여 좋은 결과를 내었고[1], Huang은 출력 확률밀도 함수를 tying하는 준연속분포 HMM을 제안하였다[2].

이산분포 HMM에 이용되는 코드워드의 경우에도 tying의 개념은 적용될 수 있는데, Shirai 등은 정보이론적 방식을 이용하여 bottom-up 방식으로 코드워드를 tying하는 방식을 제안하였다[3]. 일반적으로 이산분포 HMM에서는 코드워드의 갯수가 클수록 양자화 과정에서 발생하는 오차가 작아지고 따라서 높은 성능을 낼 것으로 기대하기 쉽지만, 학습데이터에 비해 많은 코드워드는 오히려 발생빈도가 작아져서 통계적 신뢰도가 떨어지고 인식을 저하로 연결된다. 따라서, 적절한 알고리즘을 통해 코드워드를 tying시켜준다면 성능향상을 기대할 수 있다.

본 연구에서는 수형구조 분류기를 이용하여 통계적으로 비슷한 특성을 가지는 코드워드들을 tying시켜 줌으로써, robust한 시스템을 구성할 수 있는 알고리즘을 제안한다. 이 방식은 Shirai의 방식과는 달리 top-down 방식으로 진행되며 유일한 최적해를 빠른 속도로 얻을 수 있다는 장점이 있다. 또한, 제안된 알고리즘은 코드워드 변화표를 이용하여 별도의 부가적인 과정이 필요없이 기존의 시스템에 이식이 가능하다.

II. 이산분포 HMM과 수형구조 분류기

A. 이산분포 HMM을 이용한 음성인식

HMM은 상태 전이 확률과 출력 확률밀도함수로 구성되며, 다양한 인식단위의 설정이 가능하고, 음성신호의 시간적 비선형 정합에 적합하다. 또한, 강력한 학습 알고리즘을 기반으로 하므로 많은 음성인식 시스템에서 채택하고 있는 모델링 기법이다[4]. HMM은 출력 확률밀도함수를 표현하는 방법에 따라 이산분포[1], 연속분포[5], 준연속분포[2]로 구별되며, 이 논문에서는 이산분포 HMM을 이용한다.

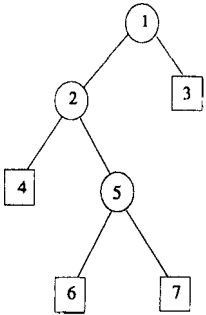
이산분포 HMM은 연속적인 음성신호의 특징벡터들을 한정된 갯수의 코드워드로 양자화하여, 학습 및 인식에 이용하는 방식으로, 연속분포에 비해 유연성이 좋고, 계산량이 작아서 실시간 처리에 유리하다. 그러나, 양자화 과정에서 발생하는 오차로 인해 인식이 저하되는 단점이 있다. 음성신호를 양자화 기법은 크게 두가지 종류로 나눌 수 있는데, 기하학적인 거리에 근거하는 방식으로 k-means 알고리즘, LBG 알고리즘[6], 수정된 k-means 알고리즘[7] 등과 음성신호의 발생 경로 자체에 의미를 부여하기 위한 supervised 벡터 양자화[8], learning vector quantization(LVQ)[9] 등이 있다.

기하학적인 거리에 의존하는 양자화 방식들은 음성 특징 벡터와 그에 대응되는 코드워드 사이의 거리를 최소화하도록 설계되며, 코드워드의 수가 늘어날수록 이러한 오차는 줄어들게 된다. 그러나, 코드워드 갯수의 증가는 파라미터의 수를 그만큼 늘리게 되어 한정된 학습 데이터를 이용한 인식기에서는 오히려 성능 저하를 초래할 수 있다. 특히, 초기값 설정에 민감하게 반응하게 되므로, 안정적인 성능을 기대하기 힘들다[10].

B. 수형구조 분류기

수형구조 분류기(tree structured classifier)는 결정트리(decision tree)라고도 하며, 다단계 혹은 연속적인 계층 결과 과정을 통해 패턴을 분류하는 방식이다. 계층적 구조를 가지는 분류기의 해석적 장점은 다양한 결정단계에서 여러 조건 및 특징을 이용하게 되므로 nearest neighbor와 같은 일단계 계층 방식에 비해 좋은 성능을 낼 수 있다는 데 있다. 수형구조 분류기는 단순한 개념과 효율적인 계산량으로 음성인식, 문자인식, 의학적 진단 등에 널리 사용된다. 본 논문에서는 이진 트리를 사용한다. 전형적인 이진 트리의 예를 그림 1에 나타내었는데, 원형의 노드는 결정 노드이고, 사각형 노드는 종말(terminal)노드로서 하나의 클래스에 대응된다. $r(i)$ 와 $l(i)$ 는 노드 i 의 좌, 우 하위노드를 나타내며, 하위 노드가 모두 0인 경우에는 종말 노드가 된다.

수형구조 분류기를 구성하기 위해서는 먼저 학습 데이터에 대해서 어떤 식의 구성을 가지며, 각 데이터가 어떤 클래스에 속하는지에 대한 정보를 알아야 한다. 수형구조 분류기의 구성과정은 노드 분할, 종말 노드 여부 판정, 클래스 지정의 3단계로 이루어지



t	l(t)	r(t)
1	2	3
2	4	5
3	0	0
4	0	0
5	6	7
6	0	0
7	0	0

그림 1. 이진트리의 예.
Fig 1. Example of binary tree.

며, Breiman 등에 의해 체계적인 이론으로 정리되었다[11]. 이 과정에서 가장 중요한 부분은 어떻게 하면 한정된 학습 데이터로부터 적정 크기의 트리를 얻느냐 하는 것으로, 초기의 연구는 주로 노드의 분할을 중단하는 기준을 찾는 데 집중되어 있었다. 그러나, Breiman 등에 의해 CART 방식으로 더 잘 알려진 보조 트리와의 이를 이용한 전지(pruning) 방식들이 소개되면서, 분할 중단 규칙은 그리 중요한 것이 아니라는 것이 알려지게 되었으며, v-fold cross validation에 의한 최적의 트리 구성방식이 이용되기 시작했다. 또한, Gelfand는 CART 방식을 보완하여 빠른 계산속도와 최적의 해를 보장하는 알고리즘을 발표하였다[12]. 본 논문에서는 Gelfand의 알고리즘을 기본으로 하여 분류기를 설계하였다. 이 방식의 기본적인 원리는 학습데이터를 2개의 부분집합으로 나누고, 각각을 트리의 성장 및 전지에 반복적으로 적용하여 수렴하게 하는 것이다. 그림 2에는 Gelfand 방식을 이용한 트리 구성의 예를 보였는데, (a)는 k-1 번째 반복에서 학습데이터 중 하나의 부분집합에 속한 데이터들을 이용하여 성장시킨 트리이고, (b)는 나머지 데이터를 부가한 다음 전지과정을 거쳐 8, 9번 노드가 제거된 상태를 나타낸다. 또한, (c)는 k번째 반복과정으로 전지된 후에 종말노드로부터 다시 성장시킨 트리이다.

III. 코드워드 Tying 알고리즘

A. 트리의 구성

수형구조 분류기의 구성을 위해 각 음성신호에 대한 분할 정보가 요구된다. 여기에서는 HMM의 학습

결과로 주어지는 파라미터를 이용하여, Viterbi 복호화과정을 통해 학습 데이터들을 분할한다. 이때 얻어지는 정보는 코드워드와 각 상태간의 동시 발생빈도로, 이 과정을 통해 주어진 코드워드 c에 대한 상태 s의 조건부 확률을 구할 수 있으며, 다음과 같이 정의된다.

$$p(s|c) = \frac{n(s, c)}{\sum_s n(s, c)}, \quad (1)$$

$$1 \leq c \leq C, 1 \leq s \leq S.$$

위 식에서 n(s, c)는 코드워드 c와 상태 s가 동시에 발생한 빈도수를 나타내며, C와 S는 각각 코드북 크기와 전체 상태의 수를 나타낸다. 다음으로는 각 노드를 나타낼 수 있는 대표 분포를 정의하는데

$$p(s|t) = \frac{\sum_{c \in t} n(c) p(s|c)}{\sum_{c \in t} n(c)} \quad (2)$$

와 같으며, 노드 t내에 존재하는 코드워드들의 통계적 특성을 함축한 형태이고, 각 코드워드의 발생빈도를 가중치로 주었다.

또한 이 대표분포로부터 노드 t의 불순도(impurity)를 다음과 같이 정의할 수 있다.

$$i(t) = \sum_s p(s|t) \log \frac{1}{p(s|t)}. \quad (3)$$

위 식은 정보량의 식과 일치하며, 직관적으로 상위 노드일수록 불순도가 높고, 하위 노드로 갈수록 낮아짐을 알 수 있다.

분할과정은 수정된 k-means 알고리즘과 유사하게 진행된다. 우선 각 분포의 유사도를 측정할 수 있는 기준이 필요하다. 분류대상이 일반적인 벡터가 아니라 확률을 나타내는 분포이므로 구별 함수를 사용하는데[13], 임의의 두 확률분포 f(x)와 g(x)에 대한 거리는 다음과 같은 구별 함수를 이용하여 나타낼 수 있다.

$$D(f(x), g(x)) = \int f(x) \log \frac{f(x)}{g(x)} dx. \quad (4)$$

구별 함수는 확률분포가 이산형태를 띄고 있을 경우에 유한한 덧셈의 형태로 나타낼 수 있으며, 하나의 분포를 기준으로 거리를 정의하는 비대칭형과, 교환법칙이 성립하는 대칭형으로 나누어 정의할 수 있다.

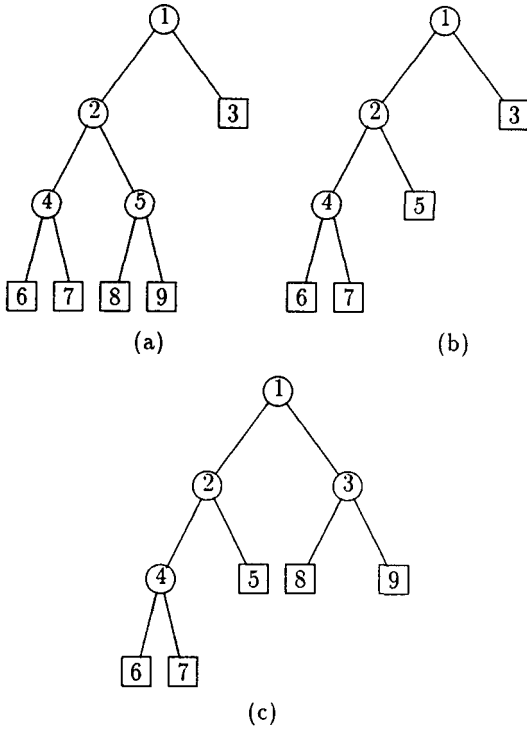


그림 2. 반복적인 성장과 전지.
Fig 2. Iterative Growing and Pruning.
(a) T_{k-1} , (b) T_{k-1}^* , (c) T_k

$$D_{대칭}(c_i, c_j) = \sum_p p(s|c_i) \log \frac{p(s|c_i)}{p(s|c_j)}, \quad (5)$$

$$D_{대칭}(c_i, c_j) = \sum_p [p(s|c_i) - p(s|c_j)] \log \frac{p(s|c_i)}{p(s|c_j)}. \quad (6)$$

분할은 노드 t 내에 존재하는 모든 코드워드들을 2개의 클래스(하위 노드)로 분류한 것으로, 각 코드워드들 중 서로간의 거리가 가장 큰 두 개를 찾아내어 초기 대표분포로 한 후 수정된 k-means 알고리즘과 같이 반복적으로 분류하는 방식을 사용한다.

분할은 몇가지 조건의 점검을 통해 중단 여부가 결정되는데, 첫째, 노드에 속하는 코드워드의 수가 1개 뿐일 경우, 둘째, 노드의 분할로 생긴 하위 노드들의 불순도가 현재 노드의 불순도와 같을 경우에는 분할이 중단된다. 두번째 조건은 아래식과 같이 나타낼 수 있다.

$$\Delta i(t) = i(t) -$$

$$\left[\frac{\sum_{c \in r(t)} n(c)}{\sum_{c \in t} n(c)} i(r(t)) + \frac{\sum_{c \in l(t)} n(c)}{\sum_{c \in t} n(c)} i(l(t)) \right]. \quad (7)$$

위 식에서 $\Delta i(t)$ 가 0일 경우 분할을 멈추며, 이때 $r(t)$ 와 $l(t)$ 는 각각 노드 t 의 좌, 우 하위 노드를 나타낸다. 마지막으로 노드의 불순도가 일정치 이하일 경우에 분할을 중단할 수도 있으나, 여기에서는 적용하지 않았다.

트리의 성장은 이와 같이 분할, 분할 중단, 그리고 종말 노드에 대한 클래스 지정으로 이루어지며, 여기에서 클래스 지정은 단순히 새로운 코드워드 번호를 설정해 주는 것이다. 전체 코드워드가 모여 있는 최상위 노드로부터 시작하여 반복적인 분할을 거쳐, 분할 중단 규칙에 해당되는 노드들이 차례로 종말 노드가 되면 성장은 완료된다.

B. 코드워드 tying 알고리즘

코드워드 tying을 위한 수형구조 분류기의 구성은 다음과 같은 단계를 거쳐 수행된다.

1. 학습데이터를 둘로 나누고 각각에 대해 $n_i(c, s)$ 와 $p_i(s|c)$ 를 구한다. ($i=1, 2$)
2. $k=1$.
3. $p_k(s|c)$ 를 이용하여 트리를 구성한다.
4. k 가 홀수이면 $k=2$, k 가 짝수이면 $k=1$ 로 해준다.
5. 구성된 트리에 $p_k(s|c)$ 를 대입하고, 전지 한다.
6. 이전 반복시기와 종말 노드의 갯수가 다르다면 단계 3으로 가고 같으면 중단한다.

빠른 트리 구성이 가능한 이유는 전지과정을 거친 이후의 성장과정이 최상위 노드로부터 새로 시작되는 것이 아니라, 전지과정 결과로 생긴 새로운 종말 노드로부터 시작하면 되기 때문이며, CART 방식에 비해 $\frac{1}{5}$ 정도의 시간이면 구성이 가능하다[12]. 또한, CART 방식이 보조 트리를 이용한 불연속적인 최적화 방식임에 비해 연속적인 성장, 전지과정이 반복되므로 최적의 해를 얻을 수 있다.

전지 과정은 다음과 같은데, \bar{t} 는 종말 노드의 집합을 나타내며, 최하위 노드 t_L 에서 부터 시작하여 상위 노드 t_1 쪽으로 진행된다.

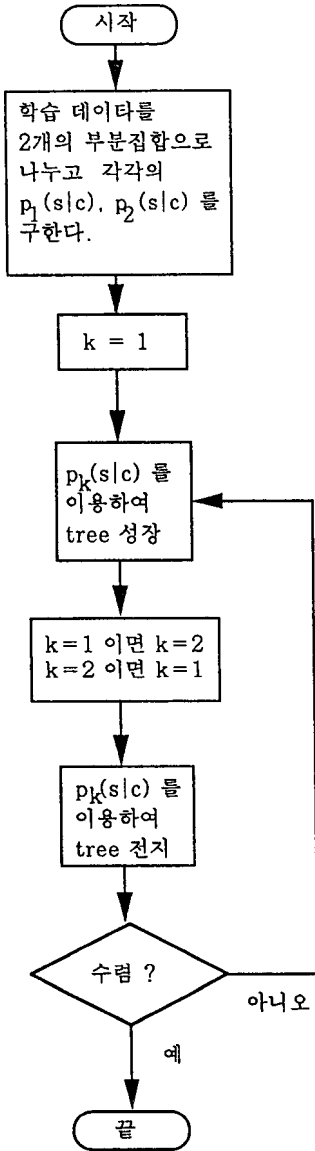


그림 3. 코드워드 tying을 위한 수형구조 분류기의 구성.
Fig 3. Construction of tree-structured classifier for code-word tying.

전지과정 : $T = T_0$

For $t = t_L, \dots, t_1$

{If $t \in \hat{T}$ then

{ $S(t) = i(t)$

If $t \in T - \hat{T}$ then

$$S(t) = \frac{1}{\sum_{c \in I(t)} n(c)} [\sum_{c \in I(t)} n(c) i(r(t)) + \sum_{c \in I(t)} n(c) i(l(t))]$$

If $i(t) \leq S(t)$ then
 $\{T = T - (T_{lu} \cup T_{ru}),$
 $l(t) = 0, r(t) = 0\}$

$$T_0 = T$$

구해진 트리는 각 종말 노드에 비슷한 통계적 성질을 가지는 코드워드들이 존재하게 되며, 이들을 이용하여 HMM 파라미터들을 조정해 준다. 즉, 코드워드 c_i, c_j, c_k 가 t 라는 종말 노드로 분류되었고, 이를 새로운 코드워드 \tilde{c} 라고 하면, tying 과정은 단순히 확률분포를 더해주는 것만으로 가능하다.

$$p(\tilde{c} | s) = p(c_i | s) + p(c_j | s) + p(c_k | s) \text{ for } \forall s \quad (8)$$

이상의 과정은 다른 각도에서보면, 벡터양자화로 인한 오차를 줄이기 위한 soft decision의 일환이라 할 수도 있다[2][14]. 즉,

$$p(x | s) = \sum_i w_i p(c_i | s) \quad (9)$$

과 같이 나타낼 수 있으며, 특징벡터 x 가 어떤 코드워드 양자화되었을때, 그 코드워드와 tying된 코드워드에 대해서는 가중치 w_i 를 1로 해주고, 나머지에 대해서는 0으로 하는 것과 같다. 기존의 soft decision 방식들과 다른 점이 있다면, 기하학적인 거리에 의존해서 가중치가 결정되는 것이 아니라 각 코드워드의 통계적 성질에 근거한다는 점을 들 수 있다.

IV. 인식 실험 결과

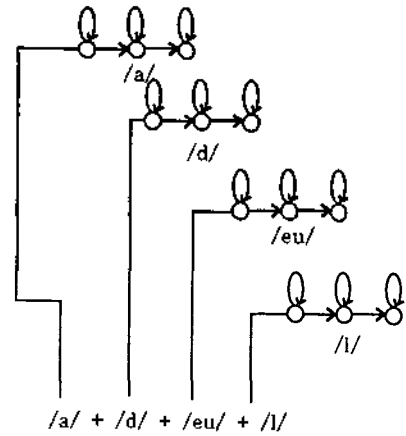
표 1에는 인식 실험에 사용된 75개의 격리단어 목록을 나타내었다. 학습에는 서로 다른 5명의 남성 화자가 1회씩 발음한 375개의 단어가 이용되었으며, 인식 실험에는 학습에 참가하지 않는 2명의 화자가 1-2회씩 발음한 216개의 단어를 사용하였다. 75개의 단어는 32개의 유사음소단위로 나타내어지며, 각 음소가 균형을 맞추어 발생하도록 한 것이다. 특징벡터로는 선형예측 분석에 기반한 12차 켈스트럼 계수를 이용하였으며, 벡터 양자화 방식은 수정된 k-means 알고리즘을 사용하였다.

HMM은 음소단위로 3개씩의 상태를 가지도록 하였고, 각 단위는 천이, 정상, 천이의 3가지 상태로 구

성된다. Left-to-right 모델을 기본으로 하였으며, 단어모델은 그림 4와 같이 음소모델의 연결로 구성된다. 초기 분할 정보는 hand-labeling에 의해 주어지는 것이 아니라, 각 단어를 균일하게 분할하는 방식을 사용하였으며, 학습 결과를 이용하여 반복적으로 재분할하게 된다.

표 1. 격리단어인식에 사용된 단어 목록.
Table 1. Word list for isolated word recognition.

1	아들	26	감기	51	납기
2	애기	27	감자	52	날뛰다
3	밥	28	간판	53	남산
4	바퀴	29	간식	54	늑대
5	뽕	30	글	55	농비
6	비행	31	꿀	56	옷
7	보리	32	고삐	57	옷밥
8	창	33	곡식	58	왼쪽
9	달	34	구리	59	풀
10	다리	35	구웠다	60	사람
11	딸	36	괜찮다	61	샘
12	들깨	37	굴	62	쌀
13	등발	38	하나	63	투구
14	된장	39	홀러	64	왔다
15	들다리	40	획기적	65	완수
16	동백	41	자리	66	웬일
17	동이	42	찾새	67	원고
18	동쪽	43	갓송이	68	약속
19	동태	44	찌개	69	양
20	의사	45	줄기	70	예
21	가보	46	칼	71	역사
22	값이	47	마음	72	연도
23	가구	48	맹새	73	욕
24	가족	49	목	74	용산
25	갈치	50	나	75	육성



(b) 단어 단위의 HMM(아들)
(c) Word level HMM

그림 4. HMM의 구조.
Fig 4. The Structure of HMM.

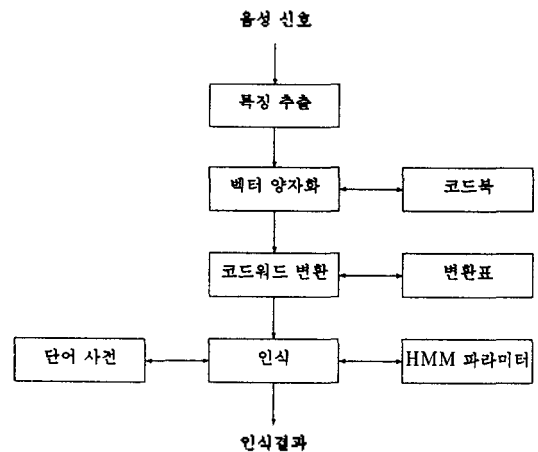
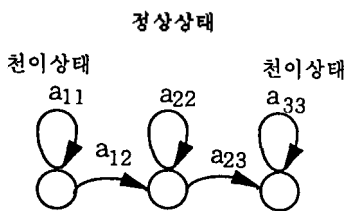


그림 5. 코드워드 tying을 이용한 인식.
Fig 5. Recognition using Codeword Tying.



(a) 음소 단위의 HMM
(a) Phoneme level HMM

인식과정은 그림 5과 같이 진행된다. 먼저 음성신호를 특징추출하고 양자화하고 코드워드로 바꾼 다음, 수형구조 분류기를 통해 얻어진 코드워드 변환표에 의해 변환시켜 준다. 이후의 과정은 기존의 시스템에서와 동일하며, HMM 파라미터는 tying에 의해 구해진 것을 사용한다.

표 2에는 코드북 크기가 256일 경우에 대한 실험 결과를 나타내었다. 분포간의 거리를 재는 척도로 대

표 2. 코드북 크기가 256일 경우의 인식 결과. (a) 기본 시스템, (b) 대칭형 구별함수를 사용하였을 경우, (c) 비대칭형 구별함수를 사용하였을 경우.

Table 2. Recognition Results of 256 size Codebook. (a) Baseline System, (b) Using Symmetric Discriminant Function, (c) Using Asymmetric Discriminant Function.

	(a)	(b)	(c)
코드북 크기	256	238	222
인식률(%)	76.9	78.2	78.2

표 3. 코드북 크기가 512일 경우의 인식 결과. (a) 기본 시스템, (b) 대칭형 구별함수를 사용하였을 경우, (c) 비대칭형 구별함수를 사용하였을 경우.

Table 3. Recognition Results of 512 size Codebook. (a) Baseline System, (b) Using Symmetric Discriminant Function, (c) Using Asymmetric Discriminant Function.

	(a)	(b)	(c)
코드북 크기	512	440	429
인식률(%)	75.0	77.3	76.9

칭형, 비대칭형 구별 함수를 이용하여 각각에 대한 결과를 보였는데, 대칭형 구별 함수의 경우는 6%의 오차 감소를 나타내었고, 비대칭형의 경우 역시 6%의 오차를 줄임을 알 수 있다. 또한, 각각의 경우에 대해 18개, 34개 씩의 코드워드들이 tying됨으로써 인식에 필요한 HMM 파라미터의 수도 그 만큼 줄어들게 된다.

코드북의 크기가 512일 경우에 인식실험 결과를 표 3에 나타내었는데, 여기서 주목해야 할 사항은 코드북 크기가 커졌음에도 불구하고 오히려 인식률이 저하되었다는 점이다. 즉, 실험에 사용된 데이터베이스에서는 512의 코드북 크기가 오히려 과다한 파라미터를 요구하여 적당하지 않다는 것을 알 수 있다. 그러나, 이 경우에도 제안된 알고리즘을 적용하면 대칭형, 비대칭형 함수에 대해 각각 9%, 8%씩의 오인식률이 감소된다.

V. 결 론

제안된 알고리즘은 수형구조 분류기를 이용하여 코드워드들을 분류하고, 동일한 종말 노드로 분류된 코드워드들을 tying시켜 주어 적정규모의 파라미터를 가지게 한다. 화자독립 격리단어 인식실험에 적용

해 본 결과 6%-9%의 오인식율이 개선되었으며, 빠른 트리 구성과 주어진 조건에서 최적의 해를 찾는 장점을 보였다. 또한, 이산분포 HMM을 이용하는 음성인식 시스템에 쉽게 적용이 가능하다.

제안된 알고리즘은 또한 벡터 양자화 과정에서 발생하는 오차를 줄이기 위한 soft decision 방식의 일환으로도 볼 수 있는데, 이때 기하학적인 거리에 의존하는 기존의 방식들과 구별되는 점은 코드워드와 상태간의 통계적 성질에 의해 가중치가 결정된다는 것이다.

참 고 문 헌

1. K. F. Lee, *Automatic Speech Recognition*, Kluwer Academic, 1989.
2. X. D. Huang, Y. Ariki, and M. A. Jack, *Hidden Markov Models for Speech Recognition*, Edinburgh University Press, 1990.
3. K. Shirai, N. Aoki, and N. Hosaka, "Multi-Level Clustering of Acoustic Features for Phoneme Recognition Based on Mutual Information," *Proc. of ICASSP*, pp. 604-607, 1989.
4. L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proc. of IEEE*, vol. 77, pp. 257-286, Feb. 1989.
5. C. H. Lee, E. Giachin, L. R. Rabiner, R. Pieraccini, and A. E. Rosenberg, "Improved Acoustic Modeling for Large Vocabulary Continuous Speech Recognition," *Computer Speech and Language*, vol. 6, pp. 103-127, Apr. 1992.
6. J. Makhoul, S. Roucos, and H. Gish, "Vector Quantization in Speech Coding," *Proc. of IEEE*, vol. 73, pp. 1551-1588, Nov. 1985.
7. J. G. Wilpon and L. R. Rabiner, "A Modified K-means Clustering Algorithms for Use in Isolated Word Recognition," *IEEE Trans. on ASSP*, vol. 33, pp. 587-594, June 1985.
8. G. Yu, W. Russel, R. Schwartz, and J. Makhoul, "Discriminant Analysis and Supervised Vector Quantization for Continuous Speech Recognition," *Proc. of ICASSP*, pp. 685-688, 1990.
9. T. Kohonen, *Self-Organization and Associative Memory (2nd ed.)*, Springer-Verlag, 1989.
10. 김도영, Hidden Markov Model을 이용한 음성인식 시스템의 Codebook 최적화에 관한 연구, 석사학위 논문,

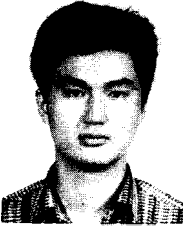
한국과학기술원, 1993.

11. L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Wadsworth International Group, 1984.
12. S. B. Gelfand, C. S. Ravishankar, and E. J. Delp, "An Iterative Growing and Pruning Algorithm for Classification Tree Design," *IEEE Trans. on PAMI*,

vol. 13, pp. 163-174, Feb. 1991.

13. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, 1973.
14. H. P. Tseng, M. Sabin, and E. Lee, "Fuzzy Vector Quantization Applied to Hidden Markov Modeling," *Proc. ICASSP*, pp. 641-644, 1987.

▲김도영



1968년 8월 7일생
 1991년 2월 : 연세대학교 전자공학과 졸업(공학사)
 1993년 2월 : 한국과학기술원 전기 및 전자공학과 졸업(공학석사)
 1993년 3월~현재 : 한국과학기술원 전기 및 전자공학과 박사과정

▲김남수

1965년 10월 18일생
 1988년 2월 : 서울대학교 전자공학과 졸업(공학사)
 1990년 2월 : 한국과학기술원 전기 및 전자공학과 졸업(공학석사)
 1990년 3월~현재 : 한국과학기술원 전기 및 전자공학과 박사과정

▲은종관 : 10권 3호 참조