

고립단어 인식에서 유사단어 정보를 이용한 단어의 검증

백창흠*, 이기정**, 홍재근*

* 경북대학교 전자전기공학부, ** 포항1대학 전자계산기과
wb@palgong.kyungpook.ac.kr

Speech Verification using Similar Word Information in Isolated Word Recognition

Chang-Heum Baek*, Ki-Jeoung Lee**, and Jae-Keun Hong*

* Dept. of Electronics, Kyungpook National University

** Dept. of Electronic Computer Science, Pohang College
wb@palgong.kyungpook.ac.kr

Abstract

Hidden Markov Model (HMM) is the most widely used method in speech recognition. In general, HMM parameters are trained to have maximum likelihood (ML) for training data. This method doesn't take account of discrimination to other words. To complement this problem, this paper proposes a word verification method by re-recognition of the recognized word and its similar word using the discriminative function between two words. The similar word is selected by calculating the probability of other words to each HMM. The recognizer having discrimination to each word is realized using the weighting to each state and the weighting is calculated by genetic algorithm.

해 본 논문에서는 유사단어에 대한 정보와 두 단어 사이에 변별력을 가지는 함수를 사용하여, 인식된 단어와 유사단어만을 대상으로 재인식하는 과정을 통해 단어를 검증하는 방법을 제안하였다.

유사단어는 각 단어의 HMM에 다른 단어의 훈련음성으로 확률값을 계산하여 가장 유사한 단어를 구했으며, 단어간에 변별력을 가지는 인식기는 각 상태에 하중값을 가지는 인식기를 사용하여 구현하였다. 단어간에 변별력을 가지는 하중값은 유전자 알고리즘 (genetic algorithm)을 사용하여 구하였다. 유전자 알고리즘은 매개변수 최적화 문제에 대하여 자연의 진화원리를 모방한 알고리즘으로 최근에 복잡한 탐색공간에서 최적의 해를 찾는 문제에 효과적으로 적용되고 있다.

I. 서론

현재 HMM[1]은 음성인식에서 가장 널리 쓰여지는 방법이다. 대부분의 경우 HMM의 파라미터는 훈련데이터에 대해 maximum likelihood (ML)를 가지도록 훈련된다. 이러한 방법은 다른 단어들에 대한 변별력을 고려하지 않는 단점이 있다. 이러한 단점을 보완하기 위

II. HMM을 이용한 단어의 인식

N 개의 상태를 가지는 단순 좌우 HMM의 구조를 그림 1에 나타내었다. HMM은 음성의 특징이 하나의 상태 i 에서 다른 상태 j 로 천이될 확률(상태전이확률, a_{ij})과 상태 j 에서 특징벡터 o_j 가 나타날 확률(관측심부확률, $b_j(o_j)$)에 의해 음성을 표현한다. ML방법을 사용하는 경우 Baum-Welch 재추정식을 사용하여 HMM에 대한 훈련데이터의 확률값이 최대가 되도록 모델(λ)의 파라미터를 추정한다.

각 단어 ν 에 대해서 HMM(λ^ν)의 파라미터를 추정 한 후에 인식과정에서 음성 O 가 입력되면

$$\nu^* = \arg \max_{\nu} [P(O | \lambda^\nu)] \quad (1)$$

에 의해 최대의 확률값을 가지는 단어 ν^* 로 인식하게 된다. 식 (1)의 확률값은 Viterbi 알고리즘을 사용하여 계산한다.

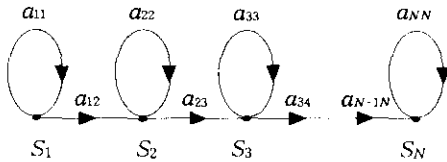


그림 1. 단순 좌우 HMM의 구조

III. 유사단어 정보를 이용한 단어의 검증

1. 유사단어 정보를 이용한 검증기

ML방법을 사용하여 구한 각 단어의 HMM에 다른 단어의 훈련데이터에 대한 확률값이 가장 높은 단어가 그 모델의 유사 단어이다.

$$\nu' = \arg \max_{\nu, \nu \neq a} [P(O' | \lambda^\nu)] \quad (2)$$

식 (2)에서 단어 a 의 유사단어는 ν' 이다. ML방법을 사용하여 훈련하는 경우에 각 단어 사이의 변별력을 고려하지 않기 때문에, 단어 ν' 가 단어 a 로 인식되는 형태로 인식과정에서 많은 오인식이 나타난다. 본 논문에서는 기존의 인식기에서 단어 a 를 인식하였을 때 단어 a 와 단어 ν' 사이의 변별력을 가지는 인식함수로 재인식하여 단어를 검증하였다. (그림 2)

유사단어를 찾는 문제에서 a 모델의 유사단어가 b 단어 로 결정되더라도 b 모델의 유사단어가 a 단어가 되는 것은 아니다. 본 실험에서 약 67%정도의 단어가 서로 유사단어의 특징을 가지고 있었다.

두 단어간에 변별력을 가지는 인식기에 사용되는 함수는, 각 상태에서 프레임의 평균 확률값에 하중값을 주어

$$f(O, \lambda) = \sum_{i=1}^T w_{q_i} [\log a_{q_i-1, q_i} + \log b_{q_i}(O_i)] / n_{q_i} \quad (3)$$

로 구현하였다. 여기서 w_{q_i} 는 상태 q_i 의 하중값이며, n_{q_i} 는 상태 q_i 에 속하는 프레임의 개수이다. w_{q_i} 는 유전자 알고리즘을 사용하여 두 단어간에 변별력을 가지도록 하였다. 본 실험에서는 식 (2)에서의 확률값과 식 (3)의 값을 더한

$$R(O, \lambda) = \alpha \cdot P(O | \lambda) + \beta \cdot f(O, \lambda) \quad (4)$$

을 사용하여 재인식에 사용하였다. 여기서 α 와 β 는 각 함수의 하중값이다.

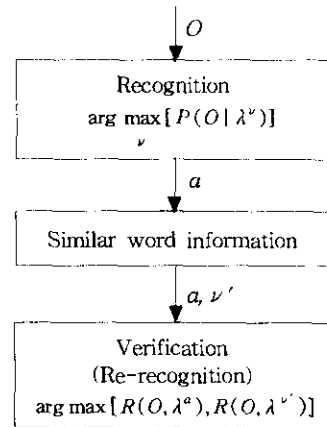


그림 2. 유사단어 정보를 이용한 검증기

2. 유전자 알고리즘을 이용하여 단어간 변별력을 가지는 인식기 구현

유전자 알고리즘[2]은 매개변수 최적화문제에 대하여 자연의 진화원리를 모방한 알고리즘으로, 염색체(chromosome) 형태로 표현된 개체군(population) 중에서 환경에 대한 적합도(fitness)가 높은 개체가 높은 확률로 살아남아 재생(reproduction)하게 되며, 두 염색체간에 정보를 교환하기 위한 교배(crossover) 연산과 전역적인 탐색이 가능하도록 하는 돌연변이(mutation) 연산 후에 다음 세대 개체군을 형성하게 되고, 이러한 과정을 반복하면서 최적의 개체를 구하게 된다. 유전자 알고리즘의 구조를 그림 3에 나타내었다.

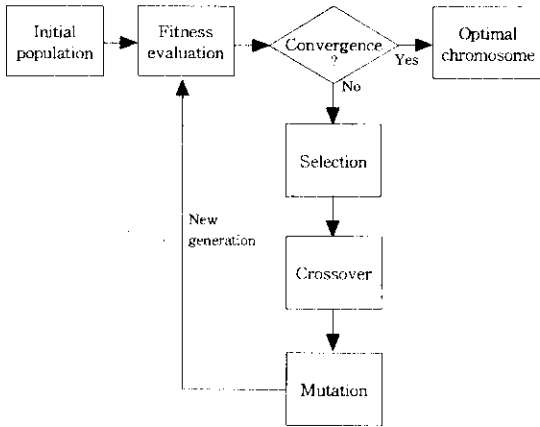


그림 3. 유전자 알고리즘의 구조

1) 적합도

N개의 상태를 가지는 HMM에서 각 상태의 하중값을

$$W = (w_1, w_2, w_3, \dots, w_N) \tag{5}$$

의 부동소수점수 (floating point number)의 염색체 형태로 표현하였다. 모델 a의 유사단어가 b일 때 두 모델 사이에 변별력을 향상시키기 위한 하중값을 두 종류의 적합도함수를 사용하여 구하였다. 첫 번째 방법은

$$F_1(W_{ab}) = f(O_a, \lambda_a) - f(O_b, \lambda_a) \tag{6}$$

$$F_1(W_{ba}) = f(O_b, \lambda_b) - f(O_a, \lambda_b)$$

를 적합도함수로 사용하여 다른 단어에 변별력을 가지는 하중값을 각각 구하였다. 여기서 W_{ab} 는 단어 b에 대해 변별력을 가지는 모델 a의 하중값이고, W_{ba} 는 단어 a에 대해 변별력을 가지는 모델 b의 하중값이다. 두 번째 방법은

$$F_2(W_{ab}) = \{f(O_a, \lambda_a) - f(O_b, \lambda_a)\} + \{f(O_b, \lambda_b) - f(O_a, \lambda_b)\} \tag{7}$$

를 적합도함수로 사용하였다. 여기서 W_{ab} 는 모델 a, b의 하중값이다. 이 방법은 두 모델사이에서 변별력을 가지도록 하는 하나의 하중값을 구한다.

2) 유전자 알고리즘의 연산

유전자 알고리즘으로 파라미터를 구할 때 반복의 초기

에는 각 개체의 적합도의 분산이 크고 시간이 증가하면서 적합도 간의 분산은 작은 값을 가진다. 본 논문에서는 개체를 선택하는 강도를 일정하게 해주는 sigma scaling 방법[2]을 이용하여 선택연산을 하였다. 교배확률 (crossover probability)에 의해 임의로 선택된 두 염색체 사이의 교배는 산술교배를 사용하였고, 돌연변이 확률 (mutation probability)에 의해 선택된 염색체는 파라미터중에 하나의 값에 불균등돌연변이 (non-uniform mutaton) 연산을 수행하였다[3]. 개체군의 크기는 50으로 하였고, 300번 반복하여 0.1에서 1.9의 범위에서 상태의 하중값을 구하였다.

IV. 실험 및 결과

1. HMM

HMM은 관측심볼확률을 가우스 혼합밀도함수로 표현하는 CDHMM을 사용하였고, 6개 7개 8개의 상태를 가지는 단순최우모델에 대해 각각 실험하였으며, 가지수는 3개로 하였다.

2. 실험 데이터

훈련 및 인식실험에 사용된 데이터는 실험실 환경에서 15개의 2음절의 한국 도시명을 10명의 화자(남녀 각각 5명)가 10회씩 발음한 단어이다. 샘플링 주파수는 12kHz이고, 12bit로 양자화하였다. 300샘플 길이의 Hamming 윈도우를 사용하였고, 100샘플간격으로 이동하며 12차의 LPC-cepstrum과 delta-cepstrum을 구하여 24차의 특징벡터를 사용하였다. 인식실험은 10명의 화자를 5개의 그룹(각각 남녀 1명)으로 나누어 round-robin 방식으로 하였다.

3. 실험 결과

상태수에 따른 그룹전체의 오인식률과 유사단어에 의한 오인식률을 표 1에 나타내었다. 전체 오인식 단어의 50% 이상이 유사단어에 의해 발생됨을 알 수 있다.

표 1. 전체 오인식률과 유사단어에 의한 오인식률

Num. of state	6	7	8
Err. rate	7.5%	8.0%	10.4%
Err. rate by similar word	4.3%	4.0%	6.8%

기존의 방법에 의해 인식된 단어를 본 논문에서 제안한 방법에 의해 유사단어와 비교하여 재인식하는 경우의 인식률을 표 2에 상태수 별로 나타내었다. 식 (6)에 의한 적합도로 하중값을 구하는 방법 1과 식 (7)에 의한 적합도를 사용하여 하중값을 구하는 방법 2에서 오인식률이 약 25% 감소하였다. 유사단어에 의한 오인식이 전체 오인식의 50%인 경우(N=7)에 유사단어에 의한 오인식이 반정도 감소했음을 의미한다.

그룹 5에 대해서는 인식률이 오히려 감소하는 결과를 보이기도 한다. 단어를 검증하는 과정에서 인식기가 기존의 인식기의 HMM을 그대로 사용하여 기존의 인식기에 의존적인 성질을 가지기 때문이다.

표 2. 각 그룹에 대한 인식률

① N = 6

	ML	Verification	
		Method 1	Method 2
Group 1	91.3%	94.7%	96.0%
Group 2	90.7%	93.7%	92.7%
Group 3	97.7%	98.0%	98.0%
Group 4	98.7%	99.0%	98.3%
Group 5	84.0%	83.7%	82.7%
Average	92.5%	93.8%	93.5%

② N = 7

	ML	Verification	
		Method 1	Method 2
Group 1	91.3%	97.3%	98.0%
Group 2	93.0%	94.0%	95.0%
Group 3	93.7%	96.3%	96.0%
Group 4	98.0%	97.7%	97.7%
Group 5	84.0%	84.0%	84.0%
Average	92.0%	93.9%	94.1%

③ N = 8

	ML	Verification	
		Method 1	Method 2
Group 1	85.3%	91.7%	90.7%
Group 2	88.3%	93.3%	93.7%
Group 3	93.3%	96.0%	96.3%
Group 4	97.7%	98.7%	99.3%
Group 5	83.3%	81.7%	82.0%
Average	89.6%	92.3%	92.4%

V. 결론

본 논문에서는 ML 방법을 사용하여 훈련된 HMM에 다른 단어의 훈련데이터로 유사단어를 찾고, 유사단어와의 변별력을 가지는 상태당 하중값을 사용한 단어 검증기를 구현하였다. 제안한 방법은 기존 HMM의 파라미터를 그대로 사용하므로 검증기를 구현할 때 계산량이 적고, 상태당 하중값과 유사단어 정보만 저장하면 되므로 구현이 간단하다. 유사단어에 의한 오인식은 전체 오인식의 50% 이상이며, 본 논문에서 제안한 방법으로 유사단어에 의한 오인식을 반정도 감소시켰다. 모델간의 변별력을 더 많이 가지는 합수를 구현한다면, 오인식률을 더욱 감소시킬 수 있을 것으로 생각된다.

참고 문헌

- [1] L. R. Rabiner, "A tutorial on hidden Markov model and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257-286, Feb. 1989.
- [2] M. Mitchell, *An Introduction to Genetic Algorithms*, The MIT Press, 1996.
- [3] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Program*, Springer-Verlag, 1995.
- [4] D. E. Goldberg, *Genetic Algorithms in search, Optimization and Machine Learning*, Addison-Wesley, 1989.
- [5] 장수현, 윤병주, "유전자알고리즘에서의 실수처리 방법 비교," *정보처리논문지*, 제5권, 제2호, pp. 361-371, 1998년 2월.
- [6] B. H. Juang, W. Chou, and C. H. Lee, "Minimum classification error rate methods for speech recognition," *IEEE Trans. on Speech and Audio Processing*, vol. 5, no. 3, pp. 257-265, May 1997.
- [7] L. R. Rabiner and B. H. Juang "An introduction to hidden Markov models," *IEEE ASSP Mag.*, pp. 4-16, Jan. 1986.