

분절 특징 HMM의 매개 변수 수의 감소에 관한 연구

Reduction of Number of Free Parameters in Segmental-feature HMM

윤 영 선*, 오 영 환*
(Young-Sun Yun*, Yung-Hwan Oh*)

*한국과학기술원 전산학과

(접수일자: 2000년 8월 29일; 채택일자: 2000년 10월 16일)

음성 인식에 많이 사용되는 HMM (hidden Markov model)을 개선하기 위하여 분절 특징을 사용한 분절 특징 HMM은 성능이 우수하다고 발표되었다. 그러나, 분절 길이가 증가하고 회귀 차수가 높아질수록 분절 특징 HMM을 표현하는 매개 변수의 수도 같이 증가된다. 따라서, 본 연구에서는 상태에서 관측 가능한 분절의 분산을 분절 내의 모든 프레임에 대하여 공통적으로 표현하는 고정 분산 방법을 통하여 성능의 저하 없이 매개 변수의 수를 줄이도록 시도하였다. 실험 결과, 두 혼합 밀도인 경우 고정 분산을 이용한 분절 특징 HMM의 성능과 시변 분산을 이용한 성능의 차이가 거의 없어, 제안된 방법의 유효성을 입증하였다.

핵심용어: 분절 특징 HMM, 분절 특징, 파라미터 수, 음성 인식

투고분야: 음성처리 분야 (2.5)

It was published that the performance of segmental feature HMM (hidden Markov model), which uses segmental features in place of frame features, is better than that of widely used conventional HMM. However, the number of free parameters is increased by longer segment length and higher regression order in SFHMM (segmental-feature HMM). Thus, in this paper, we propose a modified fixed variance estimation, which commonly uses single variance in a segment, to reduce the number of parameters. Experimental results show that the difference of the performances between time-varying variance and fixed variance estimation is small if the appropriate number of mixtures is chosen and also indicate that the proposed approach is useful.

Key words: Segmental feature HMM, Segmental feature, Number of parameter, Speech recognition

Subject classification: Speech signal processing (2.5)

I. 서 론

HMM(hidden Markov model)은 음성 인식을 비롯한 많은 응용 분야에서 널리 사용되며 그 성능을 인정받고 있다[1]. 그러나, HMM을 이용한 음성 인식 시스템이 좋은 성능을 보이고는 있지만, HMM에 사용된 일부 가정은 실제 문제에 잘 맞지 않다고 지적되고 있다. 따라서, 1990년대에서부터는 HMM의 방식이나 HMM을 개선한 방법에 대해 활발히 연구되고 있다. 그렇지만, 여러 실제 응용 분야에서는 아직도 HMM이 많이 사용되고 있기 때문에 HMM의 약점을 보완하기 위한 연구나 HMM과 같은 통계적 알고리즘과 수학적 모델을 정의할 수 있는 모델링 방법에 대하여 많은 연구를 하고 있다.

HMM의 기본이 되는 가정은 일반적으로 두 가지로 분류된다. 하나는 현재 상태는 바로 이전의 상태에 대해

서만 조건 확률을 갖는다는 1차 마코프(1st Markov) 가정이다. 이 가정으로 인하여 상태에서의 출현 확률은 상태에서의 지속 시간에 따라 지수적(exponentially)으로 감소하게 되므로, 음성 신호의 시간적 구조를 제대로 표현하지 못하게 된다. 다른 또 하나의 가정은 관측 독립(observation independence) 가정으로, 입역의 상태를 독립된 관측 벡터를 균등한 분포로 생성하는 안정된 근원(stationary source)으로 간주한다. 이러한 안정 상태에 대한 가정은 그 상태가 공명음(sonorant)이나 마찰음(frictive)과 같이 짧은 음을 표현하는 경우에는 어느 정도 합리적이거나 파열음(plosive)과 같이 긴 음을 표현하는 경우에는 불충분하다고 알려져 있다[2]. 위의 약점을 보완하기 위하여 HMM의 상태에서의 지속 시간을 다항식의 회귀 함수(polynomial regression function)를 이용하여 모델링하거나[2,3], 프레임 특징(frame feature) 대신 분절 특징(segmental feature)을 이용하여 독립 관측의 가정을 완화시키고자 하는 연구가 발표되었다. 일반적으로 음성 인식에 많이 사용되는 동적 특징(dynamic feature)도 프레임 특징에 의한 한계를 극복

하기 위하여 여러 프레임에서 회귀 함수를 통해 구해진 미분 값을 이용하고 있다[4]. 분절 특징을 이용한 대표적인 연구로는 Gish가 발표한 다항식에 의한 분절 모델링(parametric trajectory model)[5], Russell과 Gales가 독립적으로 연구한 분절 HMM(segmental HMM's)[6,7], Ostendorf에 의한 분절 모델(segmental model)[8] 등이 있으며, 1996년에 Ostendorf 등에 의하여 포괄적인 정리가 이루어졌다[9]. 그러나, 기존의 방법은 알고리즘이 복잡하여 변수의 추정이나 평가단계에서 계산 시간이 많이 걸리거나, 상태에서의 관측 확률 분포에 대한 특정 가정을 기반으로 하고 있다. 또한, 음소 단위의 전체 음향학적 벡터를 정규화하는 과정을 필요로 하고 있어 연속 음성 인식에 부적합한 면이 있다. 이러한 문제점들을 해결하기 위하여 인식 모델은 HMM에 바탕을 두고, 분절 특징을 입력으로 이용하는 분절 특징 HMM(segmental-feature HMM; SFHMM)이 제안되었다[10,11]. 그러나, 이 방법 역시 분절 길이(segment length)와 회귀 차수(regression order)가 증가할수록 많은 변수와 계산 시간을 필요로 한다. 본 논문에서는 기존의 방법을 개선하여 성능의 저하 없이 변수의 수를 줄이는 방법을 제안한다.

II. 분절 모델링

음성 신호의 연속적인 음향 특징 벡터들간의 관계는 특징 공간에서 궤적의 형태로 근사될 수 있다는 기본적인 생각에서 출발한 분절 모델링은 구형 방법에 따라 모수적(parametric) 또는 비모수적(non-parametric) 방식으로 분류된다. 분절 특징 HMM에서는 모수적 방법이 여러 음성 단위에서 궤적의 평활화 효과를 가져오기 때문에 잡음이나 환경 변화, 화자 변화에 강인한 성질을 보이고 있어, 분절 모델링에 모수적 방식을 이용하고 있다[10,11].

시간 t 에서 지속 구간 N 을 갖는 음성 분절 C_t 는 모수적 방식에 의하여 다음과 같이 표현될 수 있다.

$$C_t = ZB_t + E \tag{1}$$

여기에서 각각의 프레임은 D 차원의 특징 벡터로 구성되어 있으며, Z 는 사용된 윈도우의 형태를 결정하는 $N \times R$ 크기의 디자인 행렬이다. 또한, B_t 는 궤적 계수를 나타내는 $R \times D$ 행렬을 나타내며, E 는 궤적 추정에서의 잔차 오차(residual error)를 표현한다. R 은 궤적의 복잡도를 나타내는 회귀 차수(regression order)를 의미한다 [5,10,11].

Yun 등이 제안한 분절 특징 HMM은 Gish 등이 제안한 모수적 궤적 모델(parametric trajectory model)의 약점으로 파악되었던 경계 문제(boundary problem)를 해결하기 위하여 식 (2)와 같은 디자인 행렬을 이용하여 현재 시간에서의 관측 벡터가 분절의 중앙에 오도록 조정하였다.

이 식에서 볼 수 있듯이 제안된 방법은 현재 관측 벡터가 분절의 중앙에 위치하고, 이전 시간의 관측은 음수의 항에 의하여 지배를 받으며, 이후 시간의 관측은 양수

의 항에 의하여 지배를 받는다. 이 디자인 행렬에 의하여 표현되는 분절은 시간의 흐름에 따라 한 프레임 또는 여러 프레임(분절 길이보다 작아야 함) 단위의 이동에 의하여 분석이 되며, 이전 분절의 뒤 부분과 현재 분절의 앞 부분이 중복되어 모델링된다. 만약 음향학적 모델(acoustic model)이 문맥 독립 모델(context-independent model)로 표현이 되더라도 인접한 분절끼리는 중복이 허용되므로, 음소 단위의 모델인 경우 경계 부분에서 인접한 음소의 전이(transition) 부분을 표현할 수 있다.

$$Z = \begin{bmatrix} 1 & \left(-\frac{M}{2M}\right) & \left(-\frac{M}{2M}\right)^2 & \dots & \left(-\frac{M}{2M}\right)^{R-1} \\ 1 & \vdots & \vdots & & \vdots \\ 1 & \left(-\frac{m}{2M}\right) & \left(-\frac{m}{2M}\right)^2 & \dots & \left(-\frac{m}{2M}\right)^{R-1} \\ 1 & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & 0 & 0 \\ 1 & \vdots & \vdots & & \vdots \\ 1 & \left(\frac{m}{2M}\right) & \left(\frac{m}{2M}\right)^2 & \dots & \left(\frac{m}{2M}\right)^{R-1} \\ 1 & \vdots & \vdots & & \vdots \\ 1 & \left(\frac{M}{2M}\right) & \left(\frac{M}{2M}\right)^2 & \dots & \left(\frac{M}{2M}\right)^{R-1} \end{bmatrix} \tag{2}$$

특징 추출 단계에서 입력되는 음성 신호는 분절 단위로 묶여지고, 각 분절은 식 (2)에서와 같이 형성된 디자인 행렬에 의해 궤적으로 표현된다. 음성 신호의 분절 C_t 가 입력되면, 그 분절은 궤적 ZB_t 로 추정되고, 그 차이에 의하여 적합도(goodness-of-fit) χ^2 는 표현된다. 따라서, 음성 입력이 들어오면 프레임 별로 기본 특징을 추출하고, 다시 분절로 묶은 후에 분절 특징을 표현하는 궤적 ZB_t 과 적합도 χ^2 로 표현이 된다. 이들 두 특징이 분절 특징이 되며 다음 절에서 설명되는 분절 특징 HMM의 입력이 된다[10,11].

III. 분절 특징 HMM (SFHMM)

전 절에서 설명된 분절 특징을 음향학적 모델에 적용하여 인식 과정에 사용하기 위해서는 HMM 알고리즘과 구조를 개선하여야 한다. 분절 특징 HMM에서는 기존의 Gales와 Russell 등이 제안한 분절 HMM을 분절 특징에 맞게 수정하고 그에 맞는 수학적 모델링과 추정 알고리즘을 제안하였다. 본 절에서는 Yun 등이 제안한 분절 특징 HMM을 간략히 소개한다.

3.1. 분절 무도

기존의 분절 HMM에서는 분절을 표현하는 확률 분포를 제한하여 외적 분절 변이(extra-segmental variation)를 고정 분산이나 선형 시스템으로 나타내고, 내적 분절 변이(intra-segmental variation)는 가우스 분포(Gaussian distribution)로 표현하였다. 그러나, SFHMM에서는 외적 분절의 분포를 분절을 표현하는 궤적에 의한 확률 분포로 나타내기 때문에 평균 궤적과 그 분산으로 표현하였다. 또한, 내적

분절은 추정된 궤적에 의하여 외적 분절의 변이를 나타내는 과정에서 외적 분절 변이에 대한 가중치로서 작용하는 궤적의 추정 오차로서 표현되었다. 이는 입력 분절에 대해 먼저 궤적을 추정하고 추정된 궤적의 정확도에 따라 분절 분포의 추정에 반영한다는 것을 의미한다. 따라서, 추정 오차가 작은 궤적, 즉, 분절을 잘 표현하는 궤적은 외적 분절 분포의 계산에 많이 반영되며 추정 오차가 큰 궤적은 적게 반영된다.

분절 특징 HMM에서 최종적으로 외적 분절 변이와 내적 분절 변이의 곱에 의하여 표현되는 시간 t 에 상태 j 에서의 분절 관측 확률은 다음과 같이 계산된다.

$$b_j(C_t) = P(C_t | s_j, \lambda) = P(ZB_t | s_j, \lambda) P(C_t | ZB_t, s_j, \lambda)$$

여기에서 λ 는 모델을 나타내며, $P(ZB_t | s_j, \lambda)$ 는 외적 분절 변이를, $P(C_t | ZB_t, s_j, \lambda)$ 는 내적 분절 변이를 나타내는 확률로서 다음과 같이 표현된다.

$$\begin{aligned} P(ZB_t | s_j, \lambda) &= P(ZB_t | ZB_t, \Sigma_j) \\ &= \prod_{r=1}^M \frac{1}{(2\pi)^{D/2} |\Sigma_{r-1}|^{1/2}} \cdot \\ &\quad \exp\left\{-\frac{1}{2} \{z_r(B_t - B_j)\} \Sigma_{r-1}^{-1} \{z_r(B_t - B_j)\}\right\}. \end{aligned} \quad (3)$$

$$\begin{aligned} P(C_t | ZB_t, s_j, \lambda) &= P(C_t | ZB_t) = \exp\left\{-\frac{1}{2} x_t^2\right\} \\ &= \exp\left\{-\frac{1}{2N} \sum_{r=1}^M (c_r - z_r B_j)(c_r - z_r B_j)\right\}. \end{aligned} \quad (4)$$

여기에서 B_j 는 상태 s_j 에서의 평균 궤적 계수 행렬을 나타내며, Σ_j 는 궤적의 분산을 나타내는 분산 행렬의 열을 나타낸다. 분산 행렬의 열 Σ_j 는 분절의 각 프레임에 대한 분산 $\{\Sigma_{-M,j}, \dots, \Sigma_{0,j}, \dots, \Sigma_{M,j}\}$ 로 표현되기도 한다.

3.2. 변수 추정

SFHMM의 구성에 필요한 변수는 각 상태에서의 평균 궤적과 그 궤적에 대한 분산으로 나눌 수 있다. 물론, 한 상태를 다중 혼합 밀도(multiple mixture)로 표현하는 경우에는 각 혼합 밀도에 대한 가중치도 포함되나 일반 HMM의 가중치를 추정하는 방법과 같기 때문에 평균 궤적과 분산에 대해서만 고려하기로 한다.

임의의 상태에 대한 평균 궤적은 모든 분절에 대한 상태 j 에서의 혼합 밀도 k 에 머무를 기대치와 그 기대치와 궤적의 곱으로 얻을 수 있는 기대 평균치로 표현된다.

$$\bar{Z} B_{jk} = \frac{\sum_{i=1}^T \xi(i, k) ZB_i}{\sum_{i=1}^T \xi(i, k)} \quad (5)$$

SFHMM의 각 상태에 대한 평균 궤적이 추정되면, 이

궤적으로부터 분산을 구할 수 있다. 기존의 연구[11]에서는 분산의 추정 방법을 시변 분산(time-varying variance)과 특징 또는 궤적에 대한 고정 분산(fixed variance of features or trajectories)으로 나누고 각각에 대한 분산 추정 방법을 소개하였다. 그러나, 고정 분산을 이용한 경우, 시변 분산에 비하여 낮은 성능을 보여주고 있다. 따라서, 본 절에서는 시변 분산의 추정 방법을 소개하고 다음 절에서는 개선된 고정 분산 방법을 이용하여 성능을 향상시키는 방법에 대하여 제안하고자 한다.

시변 분산은 한 상태에서 관측 가능한 분절에 포함된 각 프레임별로 별도의 분산을 구하는 방법이기 때문에 각 상태의 모델은 1개의 평균 궤적과 N 개의 분산으로 표현된다. 다음 식은 각 프레임에 해당되는 분산의 추정 방법을 나타내고 있다.

$$\bar{\Sigma}_{n,ik} = \frac{\sum_{j=1}^T \xi(j, k) \{z_n(B_t - \bar{B}_{jk})\} \{z_n(B_t - \bar{B}_{jk})\}}{\sum_{j=1}^T \xi(j, k)} \quad (6)$$

여기에서 $z_n B_t$ 와 $z_n \bar{B}_{jk}$ 는 각 시간 t 에서 주어진 분절의 궤적과 상태 j 의 k 번째 혼합 밀도의 평균 궤적으로부터 복원된 프레임 특징을 나타낸다.

IV. 개선된 고정 분산

일반적으로 분절은 인접하는 프레임의 집합으로 구성된다. 음향학적 모델이 프레임 특징을 이용하지 않고 분절 특징을 이용한다면, 이론상으로 분절 안에 포함된 프레임의 수만큼 매개 변수의 양이 증가하게 된다. 비록 모수적 방법을 이용한다고 하더라도 매개 변수의 증가를 막을 수는 없다. 표 1에서는 상태를 표현하기 위하여 필요한 매개 변수를 일반 HMM과 SFHMM에서 비교하였다.

표 1. 각 상태를 표현하는데 필요한 매개 변수의 수 비교 (V 는 분산의 차수: 대각 분산 또는 완전 분산에 따라 $1 \times D$ 또는 $D \times D$)

Table 1. Comparisons of the number of free parameters in a state (V = covariance dimension: $1 \times D$ or $D \times D$ by diagonal or full covariance).

모델	평균 벡터 (궤적)	분산
일반 HMM	$1 \times D$	V
SFHMM	$R \times D$	$N \times V$

표에서 보는 바와 같이 SFHMM을 이용하는 경우, 상태를 표현하기 위해서는 평균 궤적은 일반 HMM에 비하여 R 배의 매개 변수를, 분산은 N 배의 매개 변수를 필요로 한다. 본 연구에서는, SFHMM의 매개 변수 수를 줄이기 위하여 분산의 표현에 변수를 줄이고자 시도하였다.

시변 분산은 각 상태에서 관측 가능한 분절의 프레임 별로 분산을 구하는 방법이다[10,11]. 기존의 고정 분산 [11]은 특징 공간과 제적 공간으로 구별하여 분절의 모든 프레임에 대해 공통적으로 분산을 적용하는 방법이다. 이 경우, 추정된 제적과 평균 제적의 차이를 행렬 연산에 의하여 정의하였으나, 본 연구에서는 고정 분산을 시변 분산의 각 프레임에 대한 평균 분산으로 정의하여 표현하였다.

$$\bar{\Sigma}_{n,jk} = \frac{1}{N} \sum_{n=1}^N \bar{\Sigma}_{n,jk} \quad (7)$$

여기에서 $\bar{\Sigma}_{n,jk}$ 는 상태 j 의 k 번째 혼합 밀도에서 관측 가능한 분절들에 대한 시변 분산의 프레임별 분산을 나타내며, 분절 길이는 $N=2M+1$ 로서 표현된다. 다시 식 (7)은 다음과 같이 확장될 수 있다.

$$\bar{\Sigma}_{n,jk} = \frac{\sum_{i=1}^T \xi_i(j,k) \{ (ZB_i - Z\bar{B}_{jk})'(ZB_i - Z\bar{B}_{jk}) \}}{N \sum_{i=1}^T \xi_i(j,k)} \quad (8)$$

이 식은 기존의 연구에서 제안된 특징에 대한 고정 분산을 분절 길이 N 으로 나눈 식과 같다. 이와 같이 분산의 표현을 수정하면 일반 HMM에 비해 평균 제적의 표현에만 더 많은 매개 변수를 필요로 한다.

V. 실험 및 결과

제안된 방법의 성능을 비교하기 위하여 TIMIT자료에 기반한 모음 분류 실험을 하였다. 모음 분류 실험은 13개의 단모음(*iy, ih, ey, eh, ae, aa, ah, ao, ow, uw, uh, ux, er*)과 3개의 복모음(*ay, oy, aw*)을 이용하였다. 학습과 실험에 필요한 자료는 TIMIT에서 문장을 음소별로 분할한 음소 표기 열에서 추출하여, 총 41,429개의 음소가 학습에 사용되었으며 15,119개의 음소가 실험에 사용되었다. 각 단계에서 사용된 음성 신호는 20msec의 분석구간과 10msec의 구간이동 조건에서 12차의 MFCC와 로그 에너지를 구한 후, 다시 각 특징에 대한 1차 미분 계수를

표 2. 단일 혼합 밀도에서의 SFHMM의 분산 추정에 따른 성능 비교

Table 2. Comparison of performance of SFHMMs with single mixture by different variance estimation.

모델	시변 분산	고정 분산
일반 HMM	55.23	-
N=3,R=2	56.86	56.12
N=3,R=3	57.21	56.12
N=5,R=2	58.77	56.99
N=5,R=3	58.79	56.89
N=5,R=4	58.76	56.70
N=5,R=5	58.74	56.68

포함한 총 26차의 특징으로 표현된다. 모음 분류의 성능에 대한 기준이 되도록 일반 HMM에 대한 성능도 포함시켰으며 이 경우에는 26차의 특징을, SFHMM인 경우 26차의 특징에 기반한 분절 특징을 이용하였다.

각 실험에서는 분절 길이와 회귀 차수를 변경하며 실험하였으며, 단일 혼합 밀도(single mixture)와 두 혼합 밀도(two mixtures)로 분리하여 실험하였다. 표 2에서는 단일 혼합 밀도에서의 모음 분류율을 보여주고 있으며, 표 3에서는 두 혼합 밀도에서의 모음 분류율을 보여주고 있다.

표 3. 두 혼합 밀도에서의 SFHMM의 분산 추정에 따른 성능 비교

Table 3. Comparison of performance of SFHMMs with two mixtures by different variance estimation.

모델	시변 분산	고정 분산
일반 HMM	58.44	-
N=3,R=2	59.56	59.90
N=3,R=3	59.64	59.87
N=5,R=2	60.43	59.90
N=5,R=3	60.84	60.14
N=5,R=4	60.88	61.10
N=5,R=5	60.86	60.94

실험 결과, 단일 혼합 밀도를 이용한 경우나 두 혼합 밀도를 이용한 경우, 일반 HMM에 비하여 좋은 성능을 보여주고 있다. 단일 혼합 밀도의 경우 시변 분산을 이용한 방법이 고정 분산을 이용한 방법보다 더 좋은 성능을 보이고 있으나, 혼합 밀도의 수를 증가시킨 경우 성능의 차이는 미미하였다. 이것은 고정 분산을 이용하여 상태를 모델링하더라도 혼합 밀도의 수가 충분하다면 시변 분산을 이용하는 경우와 성능이 비슷해진다는 것을 알 수 있다. 일반적으로 분절 모델링에서 시변 분산을 이용하는 경우의 성능이 우수하다고 알려져 있으나, 본 연구와 같이 작은 단위의 음성 분절에 대해서는 고정 분산을 이용한 방법도 혼합 밀도의 조정에 의하여 시변 분산을 이용한 성능과 비슷해짐을 알 수 있다.

VI. 결론

본 논문에서는 기존의 제안된 SFHMM이 성능은 우수하나, 일반 HMM에 비해서 매개 변수의 양이 급격히 증가하기 때문에 성능의 저하 없이 매개 변수의 수를 줄이는 방법을 제안하였다. 제안된 방법은 상태에서 관측 가능한 분절을 표현할 때, 분절을 구성하는 각각의 프레임에 대하여 분산을 계산하지 않고, 분절 내에서는 공통된 분산을 사용하게 된다. 고정 분산인 경우 분절의 분산을 표현하는데 1개의 분산 행렬만을 필요로 하기 때문에, 분절의 길이에 따라 선형적으로 분산 행렬이 증가하게 되는 시변 분산에 비하여 매개 변수의 양을 줄일 수 있다. 실험 결과, 제안된 방법이 단일 혼합 밀도인 경우에는 시변 분산을 이용한 경우보다 성능이 떨어지나, 두 혼합 밀도인 경우

에는 성능이 비슷하게 되었다. 이 결과로부터 길이가 작은 분절에 대해서는 혼합 밀도의 수가 적당하다면 고정 분산을 이용하더라도 성능의 저하 없이 매개 변수의 수를 줄일 수 있음을 알게되었다.

참 고 문 헌

1. X.D. Huang, Y. Ariki, M.A. Jack, "Hidden Markov Models For Speech Recognition," Edinburgh University Press, 1990.
2. L. Deng, M. Aksmanovic, D. Sun, J. Wu, "Speech recognition using hidden Markov models with polynomial regression functions as non-stationary states," IEEE Trans. on Speech and Audio Proc. 2 (4), pp. 507-520, 1994.
3. L. Deng, "A generalized hidden Markov model with state-conditioned trend functions of time for the speech signal," Signal Processing, 27, pp.65-78, 1992.
4. S. Furui, "Speaker-Independent Isolated Word Recognition Using Dynamic Features of Speech Spectrum," IEEE Trans. on Acoustics, Speech and Signal Processing, 34 (1), pp. 52-59, 1986.
5. H. Gish, K. Ng, "A segmental speech model with application to word spotting," In International Conference on Acoustics, Speech and Signal Processing 1993, 2, pp. 447-450, Minneapolis, Minnesota, 1993.
6. M. Russell, "A segmental HMM for speech pattern modeling," In International Conference on Acoustics, Speech and Signal Processing 1993, 2, pp. 499-502, Minneapolis, Minnesota, 1993.
7. M.J.F. Gales, S.J. Young, "The Theory of Segmental Hidden Markov Models," CUED/F-INFENG/TR 133, Cambridge University Engineering Department, Trumpington Street, Cambridge CB2 1PZ, England, 1993.
8. M. Ostendorf, S. Roukos, "A stochastic segment model for phoneme-based continuous speech recognition," IEEE Trans. on Acoustics, Speech, and Signal Processing, 37 (2), pp. 1857-1869, 1989.
9. M. Ostendorf, V. Digalakis, O.A. Kimball, "From HMM's to Segmental Models: A Unified View of Stochastic Modeling for Speech Recognition," IEEE Trans. on Speech and Audio Processing, 4 (5), pp. 360-378, 1996.
10. Y.-S. Yun, Y.-H. Oh, "A Segmental-Feature HMM For Speech Pattern Modeling," IEEE Signal Processing Letters, 7 (6), pp. 135-137, June 2000.
11. 윤영선, 오영환, "모수적 궤적 기반의 분절 HMM을 이용한 연속 음성 인식," 한국음향학회지 19 (3), pp. 35-44, 2000.

▲ 윤 영 선(Young-Sun Yun)



1990년 2월: 한국과학기술원 전산학과 졸업(공학사)

1992년 2월: 한국과학기술원 전산학과 졸업(공학석사)

1992년 3월~1995년 7월:(주)헨다 소프트웨어 기술연구소
주임연구원

1995년 9월~현재: 한국과학기술원 전산학과 박사과정

※ 주관심분야: 음성 인식

▲ 오 영 환(Yung-Hwan Oh)

한국음향학회지 제17권 6호 참조