

TR-I-0108

高精度HMMを用いた英単語認識  
English Word Recognition By HMM  
Using Multiple Codebooks

丸山 活輝 花沢 利行 川端 豪 鹿野 清宏

Katsuteru MARUYAMA, Toshiyuki HANAZAWA,  
Takeshi KAWABATA and Kiyohiro SHIKANO

1989 . 9

概要

自動翻訳電話の要素技術である英語音声認識を実現するためにHidden Markov Modelを用いて英単語認識実験を行ってきたが、学習時に音韻モデルを結合、分解しながらラベルなしで学習を行う方法により約80%の認識率を得ている。さらに認識率の向上をはかるために複数の特徴パラメータを用いるセパレートベクトル量子化とHMM出力確率の平滑化手法を導入して実験を行った。その結果、セパレートベクトル量子化を用いた場合は認識率が85.5%まで改善され、有効性が確認できた。一方、HMM出力確率の平滑化手法はあまり改善がみられず、学習サンプル数に応じた使い分けが必要であることがわかった。

ATR Interpreting Telephony Research Laboratories  
ATR 自動翻訳電話研究所

## 目次

1. はじめに	1
2. HMMを用いた英単語認識	1
2.1 HMM音韻連結学習による英単語認識	1
2.2 音声データ	3
3. セパレートベクトル量子化を用いたHMM	4
3.1 複数のコードブックを用いたHMMの学習	4
3.2 コードブックの種類	5
4. HMMの出力確率の平滑化手法	6
4.1 近傍のn個のベクトルの出力確率の荷重和により平滑化する方法	6
4.2 近傍のn個のベクトルで距離による重みをつけて平滑化する方法	6
4.3 ファジィベクトル量子化(ファジィVQ)による方法	7
5. 認識実験	8
5.1 セパレートベクトル量子化手法を用いた英単語認識実験	8
5.2 出力確率の平滑化手法を用いた英単語認識実験	13
6. まとめ	20
参考文献	20

## 1.はじめに

自動翻訳電話の要素技術である英語音声認識を行うために、日本語の音声認識において既に実績のあるHidden Markov Model<sup>[1]</sup>を用いて英単語の認識実験を行ってきたが、音韻モデルを基本単位としてそれらを連結した単語モデルを用いて学習する方法により約80%の単語認識率を得ており、学習用データの発声内容がわかっているラベリングなしでもある程度の認識率を得られることが確認できた<sup>[2]</sup>。また、花沢らによって種々の高精度HMMの検討が行われてきた。本報告では、モデルを高精度化し認識率を向上させることを目的として、日本語の音声認識において有効性が確認された、複数の音響的特徴パラメータを用いるセパレートベクトル量子化手法<sup>[3]</sup>、及び少ない学習データに対して疑似的にデータ数を増やす効果のあるHMMの出力確率平滑化手法<sup>[4]</sup>を導入して認識実験を行ったので報告する。

## 2.HMMを用いた英単語認識

### 2.1 HMM音韻連結学習による英単語認識

英語の音韻のラベル付けは日本語に比べてはるかに困難であるため、ラベリングされていない音声データを用いて実験を行っている。従って、単語内の時間情報がなく、音韻を正確に切り出すことはできないため、音韻単位での学習はできない。そこで学習用データの発音記号に応じた音韻モデルを連結して単語モデルを構成して学習を行った後、分解して音韻モデルを作る方法を導入する。

学習の対象となる音韻モデルは、無音モデル1個を含む57個である。まず、単語の発音記号を記載した単語-音韻変換テーブルを用いて、学習用単語の音韻列に従い単語モデルをつくる。この際、音韻モデルを連結したモデルの両端に無音モデルをつけ音声始末端も自動的に決定できるようにする。学習用単語について、それぞれ単語モデルを作り、forward-backwardアルゴリズム<sup>[5]</sup>によりHMMパラメータの学習を行う。単語モデルで学習した後、もと通り音韻モデルに分解する(図1)。この操作を学習用単語1487語について10回繰り返す。

認識実験は、学習された音韻モデルを連結して辞書単語(542語)の単語モデルを作り、評価用単語に対する生起確率を求め、辞書単語中で確率の最も大きい単語をその評価用単語の認識結果とする。(図2)

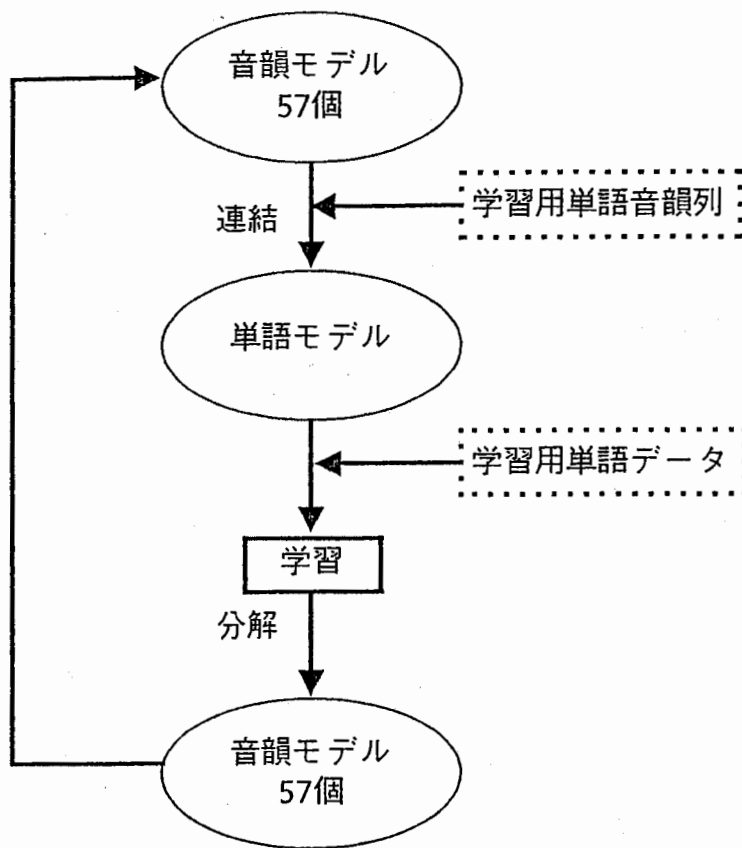


図1 HMM音韻連結学習

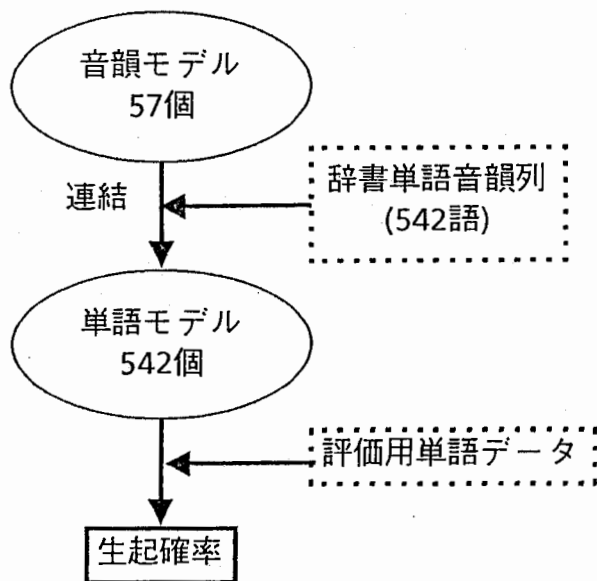


図2 音韻モデルの連結による単語認識

## 2.2 音声データ

特別の発声訓練を受けていない米国人男性1人により発声された英単語2834語(542種類)を使用する。発声内容は、国際会議の問合せや、ホテルの予約等の電話英会話文377文を単語に区切って発声したものである<sup>[6]</sup>。これらのデータにはラベリングは施されていない。音声データは、377文を10文毎に区切り、交互に学習用データと評価用データとし、3.2において説明するコードブックを用いてコード番号列に変換して実験用データとした。その結果、学習用データは1487単語、評価用データは1347単語となった。

### 3.セパレートベクトル量子化を用いたHMM

セパレートベクトル量子化は、スペクトル、パワー等の特性の異なる幾つかの音声の特徴パラメータに対し複数個のコードブックを作成し、それらのコードラベルの組合せで入力ベクトルを表現する方法で、Kai-Fu Lee<sup>[7]</sup>らの提案によるものであり認識能力の向上が報告されている。また、ATRでも日本語の音韻認識において有効性が確認されている<sup>[3]</sup>。

#### 3.1. 複数のコードブックを用いたHMMの学習

音声データは、セパレートベクトル量子化によって各フレーム毎に複数個のコードラベルで表現されるので、HMMのパラメータは遷移確率 $a_{ij}$ と、各コードブック $c$ ごとのコードラベル $v_k$ の出力確率 $b^{(c)}_{ijv_k}$ となる。時刻 $t$ での出力確率 $B_{ijt}$ は各コードラベルの出力確率 $b^{(c)}_{ijv_k}$  ( $v_k$ :時刻 $t$ での出力コードラベル)の積として計算される。すなわちコードブック数を $N$ 個とすると、

$$B_{ijt} = \prod_{c=1}^N b^{(c)}_{ijv_k}$$

である。従って、前向き確率、後向き確率はそれぞれ

$$\alpha(j, t) = \sum_i \alpha(i, t-1) a_{ij} B_{ijt}$$

$$\beta(i, t) = \sum_j \beta(j, t+1) a_{ij} B_{ijt+1}$$

となる。遷移確率 $a_{ij}$ 、出力確率 $b^{(c)}_{ijv_k}$ は以下のように推定される。

$$\text{遷移確率 : } a_{ij} = \frac{\sum_t \alpha(i, t-1) a_{ij} B_{ijt} \beta(j, t)}{\sum_j \sum_t \alpha(i, t-1) a_{ij} B_{ijt} \beta(j, t)}$$

$$\text{出力確率 : } b^{(c)}_{ijv_k} = \frac{\sum_{t: v_k} \alpha(i, t-1) a_{ij} B_{ijt} \beta(j, t)}{\sum_t \alpha(i, t-1) a_{ij} B_{ijt} \beta(j, t)}$$

### 3.2 コードブックの種類

本報告では、単一のコードブックを用いた場合と、複数のコードブックを用いた場合について実験を行った。コードブックの作成には、学習用単語中の300単語を用いて次に述べるコードブックを作成した。

#### (1) 単一のコードブック

##### ① パワー重み付きスペクトル(PWLR)

PWLR距離尺度による256個のベクトルからなるコードブック。

#### (2) 複数のコードブック

##### ① スペクトル(WLR)

WLR距離尺度による256個のベクトルからなるコードブック。

##### ② スペクトルの時間変化(DCEP)

古井<sup>[8]</sup>が不特定話者単語認識を対象としてスペクトル等と組み合わせることによって良好な結果を得ている特徴パラメータであり、本報告でもほぼ同様のパラメータを用いる。すなわちLPCケプストラム係数の1次から16次までの時間方向の線形回帰係数である。時間変化の計算範囲は古井の検討を参考にして48msecとした。コードブックはユークリッド距離による256個のベクトルからなるものとした。

##### ③ パワー(POW)

PWLRのパワー項の距離尺度による64個のスカラーからなるコードブック。

## 4.HMMの出力確率の平滑化手法

HMMは音韻のゆらぎを統計的に表現するのでその学習には多量のデータが必要であり、学習データが不十分な場合はHMMのパラメータ推定の精度が低下し認識率に大きく影響する。ここで使用している学習データにおいてもデータ数が不十分な音韻がいくつかあり、そのような音韻を含む単語の認識が困難であることがわかっている<sup>[2]</sup>。そこで花沢<sup>[4]</sup>らによって有効性が確認されているHMMの出力確率の平滑化手法を導入を試みた。平滑化手法は花沢らの検討と同様に次の3種類の方法を採用した。

### 4.1 近傍のn個のベクトルの出力確率の荷重和により平滑化する方法<sup>[9]</sup>

#### (平滑化手法①)

コード番号kのベクトルを $v_k$ 、 $v_k$ の近傍のn個のベクトルを $v_{k1}, v_{k2}, \dots, v_{kn}$ とし平滑化前の $v_k$ の出力確率を $bv_k$ 、平滑化後 $b'v_k$ とすると、

$$b'v_k = (1-w_1)bv_k + \sum_{i=1}^n [(w_1/n)bv_{ki}]$$

として平滑化する。 $w_1$ は荷重比を決める定数であり、 $0 < w_1 < 1$ の値をとる。

### 4.2 近傍のn個のベクトルで距離による重みをつけて平滑化する方法

#### (平滑化手法②)

ベクトル $v_k$ と近傍ベクトル $v_{ki}(i=1,2,\dots,n)$ との距離を $d_{ki}$ とするとき

$$b'v_k = \frac{\sum_{i=0}^n [10^{-(w_2*d_{ki})}bv_{ki}]}{\sum_{i=0}^n 10^{-(w_2*d_{ki})}}$$

$$bv_{k0} = bv_k$$

として平滑化する。 $w_2$ は重みを決める定数であり、 $0 < w_2$ である。

平滑化手法①、②はともに通常のforward-backwardアルゴリズムで学習を終了したのちに平滑化を行っている。



### 4.3 ファジィベクトル量子化(ファジィVQ)による方法<sup>[10]</sup>

通常のVQでは1フレーム分の音声パラメータをコードブック中の1つのベクトルで表すが、ファジィVQでは複数のベクトルの一次結合で表現する。すなわち入力音声のベクトルを $\mathbf{X}$ 、コードブックのベクトルを $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M$ 、 $\mathbf{X}$ と $\mathbf{v}_k$ との距離を $d_k$ とするとき

$$\mathbf{X} \xrightarrow{\text{ファジィVQ}} \sum_{i=1}^M m_i \mathbf{v}_i$$
$$m_i = (1/d_i)^{(1/F-1)} / [\sum_{k=1}^M (1/d_k)^{(1/F-1)}], \quad F > 1$$

とする。 $F$ はファジネスと呼ばれる定数であり $F$ が1に近い程、 $\mathbf{X}$ と最小の距離をもつベクトル $\mathbf{v}_k$ に対する $m_k$ が大きくなる。線形和を取るベクトル数はコードブックのサイズ $M$ までとれるが、近傍のベクトル $n(<M)$ 個までで打ち切ってもよい。このように1フレーム分の音声パラメータが複数のベクトルとその係数 $m_k$ で表現されるので疑似的にデータ数を増加させる効果があり学習データが少ない場合の認識率の向上が報告されている<sup>[10]</sup>。

## 5. 認識実験

### 5.1 セパレートベクトル量子化手法を用いた英単語認識実験

単一のコードブック(PWLR)、及び複数のコードブック(WLR, DCEP, POW)を用いて認識実験を行った結果を表1に示す。複数のコードブックを用いた場合は単一コードブックの場合に比べて4.5%高い認識率を得ており、複数のコードブックを用いることの有効性が確認できた。

表2、3に評価単語の生起確率が辞書単語の中で何番目であるかをまとめた累積正答率を示す。複数のコードブックを用いた場合、1位正答率では、単一コードブックの場合に比べて高い認識率を得ているが、5位までではほぼ同じ正答率となっている。すなわち、複数のコードブックを用いることは、認識順位の低い致命的な誤りに対しては効果はないが、1~3位くらいの類似した単語の識別には有効であることがわかった。

表4には、音韻数別の誤認識率を示した。音韻数の多い単語は誤認識率が低くなっている。音韻数が多いほど単語の冗長度が増し、他の単語と識別しやすくなると思われる。複数のコードブックを用いた場合は、単一コードブックを用いた場合に比べてどの音韻数においても誤認識率が減少しているが、音韻数が多いほど誤認識率が低いという傾向は変わらない。

3個以上誤認識された単語について、評価単語、学習単語それぞれの誤認識の割合、その単語に対して認識した単語について表5にまとめた。単一のコードブックを用いた場合に致命的に認識できていなかった単語は、複数のコードブックを用いてもほとんど改善がみられないが、very, no, pay, this, and, so, be, youについてはかなり誤認識が減っており、複数のコードブックを用いることがHMMの学習において有効であることが確認できた。

表1 単語認識率

	評価用単語 1347語		学習用単語 1487語	
	正解 単語数	認識率 [%]	正解 単語数	認識率 [%]
単一のコードブック PWLR	1091	81.0	1371	92.2
複数のコードブック WLR & DCEP & POW	1151	85.5	1417	95.3

表2 累積正答率(単一のコードブック)

順位	単語数	累積単語数	累積正答率[%]
1	1091	1091	81.0
2	103	1194	88.6
3	38	1232	91.5
4	38	1270	94.3
5	21	1291	95.8
6	6	1297	96.3
7	2	1299	96.4
8	1	1300	96.5
9	5	1305	96.9
10	5	1310	97.3

表3 累積正答率(複数のコードブック)

順位	単語数	累積単語数	累積正答率[%]
1	1151	1151	85.5
2	77	1228	91.2
3	26	1254	93.1
4	22	1276	94.7
5	13	1289	95.7
6	6	1295	96.1
7	7	1298	96.4
8	1	1299	96.4
9	1	1300	96.5
10	1	1301	96.6

表4 音韻数別誤認識率

音韻数	評価 単語数	単一の コードブック PWLR		複数の コードブック WLR & DCEP & POW	
		誤認識 単語数	誤認識率	誤認識 単語数	誤認識率
1	159	32	20.1	25	15.7
2	422	90	21.3	72	17.0
3	339	64	18.9	50	14.7
4	144	36	25.0	25	17.4
5	93	18	19.4	11	11.8
6	65	11	16.9	9	13.8
7	32	3	9.4	0	0.0
8	26	1	3.8	3	11.5
9	38	1	2.6	1	2.6
10	27	0	0.0	0	0.0
11	2	0	0.0	0	0.0

表5 誤認識の多い単語

単語	単一コードブック PWLR			複数のコードブック WLR & DCEP & POW		
	誤認識数 全数 評価 単語	誤認識数 全数 学習 単語	認識した単語 ( )内は度数 (評価単語)	誤認識数 全数 評価 単語	誤認識数 全数 評価 単語	認識した単語 ( )内は度数 (評価単語)
bank	5/5	0/0	thank(3),think(2)	3/5	0/0	thank(3)
I'm	3/3	7/7	from(2),time(1)	3/3	4/7	time(2),fine(1)
tour	6/6	0/1	for(2),record(2), cover(1),before(1)	6/6	0/1	record(3),cover(2), four(1)
very	7/7	7/7	vary(6),interpreter(1)	2/7	1/7	three(1),vary(1)
way	3/3	1/2	away(2),while(1)	2/3	0/2	away(1),waiting(1)
on	7/8	0/7	foreign(2),join(2),one(2) from(1)	6/8	0/7	one(3),from(2), foreign(1)
any	6/7	2/2	ATR(4),Japanese(1), sending(1)	6/7	2/2	Japanese(3),ATR(1), fifteen(1),taking(1)
do	15/18	6/7	to(12),June(3)	18/18	7/7	to(15),June(3)
no	5/6	1/2	double(3),myself(1), number(1)	2/6	0/2	myself(1),Tomoko(1)
attend	4/5	0/3	intend(2),content(1), understand(1)	2/5	2/3	intend(1),content(1)
pay	5/8	0/6	take(4),paid(1)	2/8	1/6	paying(1),take(1)
this	3/5	0/2	Ms.(1),office(1),tours(1)	0/5	0/2	
the*	6/11	5/11	me(3),fee(2),be(1)	4/11	5/11	be(2),fee(1),Miki(1)
if	4/9	3/5	it(1),is(1),live(1),six(1)	4/9	0/5	live(1),it(1),itself(1), since(1)
and	9/23	4/32	end(3),attend(2),send(2), chance(1),sent(1)	3/23	0/32	end(1),planned(1), send(1)
so	3/8	0/3	social(2),Tomoko(1)	0/8	0/3	
will	7/19	9/25	well(7)	9/19	10/25	well(9)
I	15/47	14/37	five(4),I'd(4),like(3), I've(2),apply(1),I'm(1)	12/47	7/37	I'd(4),fine(3),I've(2), type(2),like(1)
have	4/13	0/12	however(3),happy(1)	2/13	12/12	haven't(2)
be	4/17	2/25	before(1),fee(1), need(1),speak(1)	1/17	1/25	me(1)
you	8/65	3/50	Yukiko(3),future(2), Yamamura(2),used(1), venue(1)	2/65	0/50	future(1),June(1)

(\* [ði:]のthe)

## 5.2 出力確率の平滑化手法を用いた英単語認識実験

前述した3種類の平滑化手法を用いて認識実験を行った。但し、平滑化手法①、②は単一コードブックの場合のみ実験を行った。また、ファジィVQによる方法は、学習時、認識時、学習時と認識時の両方に用いる3つの場合について実験を行った。

認識実験では、以下の値をそれぞれの平滑化手法に用いた。

- ・平滑化手法①  $w_1=0.1$
- ・平滑化手法②  $w_2=5$
- ・ファジィVQ  $F=1.5$

そして、平滑化に用いたベクトル数 $n$ は各手法とも5個とした。

単一コードブック、複数コードブックそれぞれの場合での単語認識率(1位正答率)、5、10位までの累積正答率を表6、7に示す。単一コードブックの場合では、平滑化手法①、②とファジィVQを認識時に行った場合の単語認識率は平滑化を行わない場合とほぼ同じであるが、ファジィVQを学習時、学習・認識時の両方に行った場合の単語認識率は約4%低下している。また、複数コードブック場合は、ファジィVQを行った場合いずれも単語認識率が低下しており、特にファジィVQを学習・認識時の両方に行った場合は9.3%も低下している。従って、この実験では単語認識率において平滑化手法①、②は、改善効果がほとんどなく、ファジィVQは認識率を低下させている。

そこで、どの単語が認識率を低下させているのか分析してみた。誤認識の数が3個以上の単語について各種平滑化手法別にまとめたものを表8、9に示す。表8は単一コードブックの場合で、平滑化なし、平滑化手法①、平滑化手法②、ファジィVQ(学習時)、ファジィVQ(認識時)、ファジィVQ(学習・認識時両方)の6通り、表9は複数コードブックの場合で、平滑化なし、ファジィVQ(学習時)、ファジィVQ(認識時)、ファジィVQ(学習・認識時両方)の4通りで比較した。単語名横のカッコ内の数字は評価用単語中に含まれるその単語の数である。

表8より、平滑化手法①、②は、どちらも傾向はほぼ同じで、a, and, do, have, I, in, thank, willなどの誤認識が増えている。逆に、he, can, youなどは誤認識が減っている。

ファジィVQを行った場合は、平滑化手法①、②で誤認識が増加した単語に加え、I'll, it, need, no, that, the2, this, very, whatなどの誤認識が増えている。

る。特に、学習時と学習・認識時両方にファジィVQを行った場合が多く増加している傾向がある。また、誤認識が減った単語としては、august, can, if, it's, much, no, on, pay, transferなどがある。

平滑化を行って誤認識が減少、及び増加したおもな単語の音韻列(発音記号)を表10、11に示す。また、学習データに含まれる各音韻の数を表12に示す。平滑化を行って誤認識が減少した単語にはts, tʃ, ɔ:などの音韻が含まれており、これらの音韻は学習データに少ししか現れていないことから、少ない学習データに対しては平滑化が有効であることがわかる。また、平滑化を行って誤認識が増加した単語に含まれる音韻はどれも学習データに多く含まれており、学習データ数が十分な場合には平滑化の効果よりもデータをぼかしてしまう逆効果の方が大きいため誤認識すると考えられる。従って、音韻の学習データ中に含まれる数が少ない場合は平滑化を行い、多い場合は行わないように平滑化手法を用いる方法が良いと思われる。



表6 平滑化手法を用いた単語認識率  
(単一コードブック)

	1位	5位	10位
平滑化なし	81.0	95.8	97.3
平滑化手法①	81.5	97.1	98.4
平滑化手法②	80.5	96.7	98.4
ファジィVQ[学習時]	77.1	96.1	98.8
ファジィVQ[認識時]	81.1	96.8	98.5
ファジィVQ[両方]	77.2	96.9	99.0

表7 平滑化手法を用いた単語認識率  
(複数コードブック)

	1位	5位	10位
平滑化なし	85.5	95.7	96.6
ファジィVQ[学習時]	80.6	95.3	99.0
ファジィVQ[認識時]	82.1	95.8	97.3
ファジィVQ[両方]	76.2	93.2	96.8

表8 各種平滑化手法による誤認識単語  
(単一コードブック)

平滑化	なし	①	②	学習	認識	両方
エラー単語数	256	249	263	309	255	307
a (22)	1	5	6	13	5	13
and (23)	9	11	12	13	9	8
any (7)	6	6	7	3	7	7
are (9)	2	2	2	2	3	3
at (6)	1	2	2	3	2	3
attend (5)	4	3	3	4	3	5
bank (5)	5	5	5	5	5	5
be (17)	4	2	1	8	1	6
can (16)	2	0	0	3	0	0
do (18)	15	18	18	18	18	18
give (5)	1	1	1	1	3	2
goodbye (8)	2	3	3	2	3	4
have (13)	4	8	9	6	6	4
hundred (4)	2	2	2	3	3	3
I (47)	15	18	20	29	16	25
I'll (4)	1	2	2	3	4	3
I'm (3)	3	2	3	3	2	3
if (9)	4	4	4	4	1	2
in (17)	2	4	5	8	5	8
it (21)	2	1	3	2	3	7
no (6)	5	5	4	3	5	3
on (8)	7	7	8	8	8	5
pay (8)	5	5	4	5	2	3
right (3)	1	2	2	2	3	3
see (12)	1	1	2	3	1	1
so (8)	3	2	2	2	2	2
thank (19)	2	4	4	5	4	4
the2 (11)	6	6	5	5	5	5
this (5)	3	3	2	4	2	4
three (9)	1	1	0	4	1	2
to (47)	1	2	2	2	3	3
tour (6)	6	5	4	5	3	5
very (7)	7	7	7	6	7	7
way (3)	3	3	2	3	3	3
what (6)	2	2	2	6	2	6
will (22)	9	11	14	22	15	22
you (65)	8	6	6	14	5	8

表9 各種平滑化手法による誤認識単語  
(複数コードブック)

平滑化	なし	学習	認識	両方
エラー単語数	196	262	241	321
a (22)	3	21	10	21
and (23)	3	3	2	3
any (7)	6	7	7	7
are (9)	3	5	3	4
at (6)	0	2	0	4
attend (5)	2	3	3	3
august (5)	3	0	2	0
bank (5)	3	3	4	5
be (17)	1	2	3	2
dash (11)	0	0	0	4
do (18)	18	18	18	18
hello (5)	0	2	0	3
I (47)	12	26	24	33
I'll (4)	1	4	4	4
I'm (3)	3	3	3	3
if (9)	4	6	5	5
in (17)	0	6	0	8
it (21)	2	6	1	14
it's (5)	5	2	2	0
may (5)	1	2	5	1
much (11)	7	4	5	4
name (6)	0	1	1	3
need (4)	0	3	3	3
no (6)	2	5	3	6
on (8)	6	4	3	2
or (3)	2	1	3	3
pay (8)	2	1	3	3
that (13)	0	2	1	7
the2 (11)	4	8	8	10
this (5)	0	4	1	2
tour (6)	6	6	6	6
transfer (5)	5	2	4	3
very (7)	2	7	6	7
way (3)	2	3	3	3
what (6)	1	2	3	5
will (22)	12	15	17	17
you (65)	2	1	3	5

表10 誤認識が減った単語の発音記号

単語	発音記号
it's	i ts
much	m ʌ tʃ
on	ɔ: n

表11 誤認識が増えた単語の発音記号

単語	発音記号
a	ə
i	ai
in	i n
it	i t
the2	ð i:
will	w i l

表12 学習データに含まれる各音韻の数

音韻 音韻数	音韻 音韻数	音韻 音韻数	音韻 音韻数
i: 155	i 408	e 195	æ 151
a: 1	a 93	ɔ: 32	u: 72
u 33	ʌ 78	ə: 27	ə 357
ə 62	ɛ 3	o 26	ei 104
ai 129	ɔi 5	au 31	ou 79
ju: 58	ju 3	iə 5	ɛə 16
aə 20	ɔə 39	uə 1	juə 10
aia 1	auə 0	p 99	b 92
t 345	d 176	k 221	g 35
f 126	v 95	θ 56	ð 128
s 268	z 127	ʃ 48	ʒ 0
h 74	ts 13	dz 22	tr 9
dr 17	tʃ 27	dʒ 30	m 147
n 411	ŋ 67	l 196	r 159
j 36	w 103	無音 2974	

## 6. むすび

複数の特徴パラメータを用いるセパレートベクトル量子化により単語認識率が81.0%から85.5%に向上し、セパレートベクトル量子化の有効性が確認できた。また、HMM出力確率の平滑化手法を用いた実験では、部分的には効果があったが、平滑化の逆効果の方が大きく、認識率の改善はみられなかった。HMM出力確率の平滑化は学習サンプル数に応じた使い分けが必要であると考えられ今後検討したい。

## 参考文献

- [1] 花沢、川端、鹿野：「Hidden Markov Modelを用いた日本語有声破裂音の識別」, 信学技報 SP87-98(1987.12)
- [2] 丸山、花沢、川端、鹿野：「Hidden Markov Modelを用いた英単語認識」、ATR Technical Report TR-I-0047(1988.10)
- [3] 花沢、川端、鹿野：「HMM音韻認識におけるセパレートベクトル量子化の検討」, 音講論集2-P-8(1988.10)
- [4] 花沢、川端、鹿野：「HMM音韻認識におけるモデル学習の諸検討」, 信学技報 SP88-22(1988.6)
- [5] Levinson,S.E, Rabiner,L.R. and Sondhi,M.M. : "An Introduction to the Application of the Theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition", The Bell System Technical Journal, vol.62, No.4 (April,1983)
- [6] 丸山、鹿野：「英語音声データベースの作成」、ATR Technical Report TR-I-0048(1988.10)
- [7] Lee,K.F., Hon,H.W. : "Large Vocabulary Speaker-Independent Continuous Speech Recognition Using HMM", ICASSP88(Aprl,1988)
- [8] Furui,S. : "SpeakerIndependent Isolated Word Recognition Using Dynamic Feature of Speech Spectrum", IEEE Trans.ASSP-34,1(1986)
- [9] Sugawara,K., Nishimura,M., Toshioka,K., Okouchi,M., and Kaneko,T. : "Isolated Word Recognition Using Hidden Markov Models", ICASSP85 (March 1985)
- [10] Tseng, Ho-ping, Sabin, Michael,J., Edward,A. : "Fuzzy Vector Quantization applied to Hidden Markov Modeling", ICASSP87(April 1987)