

TR-I-0315

ベクトル場モデル適応方式を用いた
雑音環境下音声認識の検討

Speech Recognition Based on Vector Field Model Adaptation
In a Noisy Environment

土井啓輔 大倉計美
Keisuke DOI Kazumi OHKURA

1993.3.3

概要

ベクトル場モデル適応方式を用いた雑音環境下における音素認識実験を行った。本実験では、従来方法のコードブック間の対応付けヒストグラムに基づく話者適応方式との比較を行い、ベクトル場モデル適応方式が従来方法よりも優れた雑音抑圧能力をもつことを確認した。

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

© (株)ATR 自動翻訳電話研究所 1993
© 1993 by ATR Interpreting Telephony Research Laboratories

目次

1	はじめに	2
2	話者適応法	2
2.1	ヒストグラムに基づく話者適応方式	2
2.2	ベクトル場モデル適応方式	3
2.3	ファジィベクトル量子化	4
3	実験	5
3.1	実験条件	5
3.2	雑音抑圧性能の評価	6
3.3	雑音環境下における話者適応性能評価	7
4	まとめ	8
5	付録	10

1 はじめに

ATR 自動翻訳電話研究所では、コードブック間の対応付けヒストグラムに基づく話者適応方式 [1](以下、CBH と略す) がデモ装置に組み込まれ稼働している。また、本話者適応方式が雑音への適応能力をもつことも示されている [2]。ヒストグラムに基づく話者適応方式では 25 単語という比較的少量の学習単語で適応を行なえるが、実際の使用状況を考慮すると、更に学習単語数を削減する必要がある。この学習単語数の問題を解決するため、少数学習データを用いた話者適応方式としてベクトル場モデル適応方式 (以下、VFM と略す) が提案され、その有効性が示された [3]。本報告では、雑音環境下における VFM の性能を CBH との比較により評価し、その有効性を示す。

2 話者適応法

本章では、初めに VFM と CBH のアルゴリズムについて説明し、次に学習データの増加を避けるための方法として用いたファジィベクトル量子化の説明をする。

2.1 ヒストグラムに基づく話者適応方式

本方法は、ベクトル量子化が特徴空間の離散表現になっていることを利用し、有限個の離散点の話者間の関係を見いだすことにより、話者適応を行うものである。

話者間の対応関係は、未知話者のコードブックを用いてベクトル量子化された未知話者の学習音声と、標準話者のコードブックを用いてベクトル量子化された標準話者の学習音声間で、非線形マッチング (DTW) を行うことにより求められ、対応付けヒストグラムにより表現される。以下に、本方法のアルゴリズムを示す。

(話者適応アルゴリズム)

[学習]

- ステップ 1: 未知話者 (A) の学習単語より未知話者のベクトル量子化コードブック $\{v_{iA}\}$ を作成する。
- ステップ 2: 未知話者の学習音声を $\{v_{iA}\}$ を用いてベクトル量子化する。
- ステップ 3: 標準話者 (B) のコードブック $\{v_{jB}\}$ でベクトル量子化した標準話者の学習単語と未知話者の学習単語間で DTW を行い、最適パスを求める。
- ステップ 4: DTW の最適パスに従い、 $\{v_{iA}\}$ と $\{v_{jB}\}$ 間のベクトルの対応回数を求め、対応付けヒストグラム (h_{ij}) を求める。
- ステップ 5: 対応付けヒストグラムの値を重みとし、未知話者の空間を標準話者の空間に写像するための変換コードブック $\{v_{iA-B}\}$ を求める。
- ステップ 6: $\{v_{iA}\}$ を $\{v_{iA-B}\}$ に入れ換える。

ステップ7: DTW 時の距離が収束していないならば、3へ戻る。

学習時には、2.3 節に示すファジィ量子化を用いた。

(ヒストグラムに基づく出力確率の変換)

認識時には学習時に求めたヒストグラム h_{ijn} を両者のコードベクトルにおける対応付けの確からしさと見なし、次式に示すように標準話者の HMM のコードベクトルの出力確率 b_{kjn} と各特徴パラメータ毎に求めたヒストグラム h_{ijn} の積をとることにより、標準話者のコードベクトルの出力確率 $\omega_t(k)$ を変換することにより認識を行う。

$$\omega_t(k) = \prod_{n=1}^P \left(\sum_{j=1}^{M_n} \left(\sum_{i=1}^{C_n} u_{in} h_{ijn} \right) b_{kjn} \right)$$

ただし、 n は特徴パラメータ番号 (CEP, Δ CEP, Δ POW)、 t は時刻、 k は HMM のステート番号、 i は入力コードベクトルの番号を表す。

2.2 ベクトル場モデル適応方式

本方法を学習時と認識時に分けて説明する。

[学習]

1. 入力話者の学習用音声と標準話者の音声を DTW を用いて対応付け、標準話者のコードベクトルに対応付けられた入力話者のベクトルの平均を求める。次に、対応付けが求められたコードベクトル間の差分を用いて対応付けの行なわれなかったベクトルの差分ベクトルの推定を行なう。

(差分ベクトルの推定)

- ステップ1: 標準話者コードブック C^R を変換コードブック C^T の初期値とする。
- ステップ2: 未知話者の入力ベクトル列と、標準話者のコード列とを変換コードブックを用いて DTW により時間対応付けを行う。
- ステップ3: m 番目のベクトル C_m^T に対応付けられた入力ベクトルの集合を M とする。 M に属する入力ベクトル x の平均値と C_m^R の差分ベクトル V_m を求める。

$$V_m = \frac{1}{N_m} \sum_{x \in M} x - C_m^R$$

ただし、 N_m は M に属するベクトルの個数である。

- ステップ4: $N_n = 0$ であるベクトル C_n^R について、 $N_k > 0$ であるコードベクトル C_k^R のファジィ級関数 $\mu_{n,k}$ を求める。 C_n^R の差分ベクトル V_n を、 C_k^R の差分ベクトル V_k と $\mu_{n,k}$ を用いて、以下の式により計算する。

$$V_n = \sum_k \mu_{n,k} V_k$$

ステップ5: 変換コードブック中のすべてのベクトル C_n^T を差分ベクトル V_n を用いて更新する。

$$C_n^T = C_n^R + V_n$$

ステップ6: DTW 時の距離が収束していないならば、2へ戻る。

2. 以上の方法で求めた差分ベクトルに対して、ファジィ級関数に基づくスムージングを行なうことにより変換コードブックを求める。

(スムージングアルゴリズム)

ステップ1: ベクトル C_n^R に対する、 $N_k > 0$ であるすべてのベクトル C_k^R のファジィ級関数 $\mu_{n,k:k \neq n}$ を求める。

ステップ2: コードベクトル C_n^R の差分ベクトルを次式によって計算する。

$$V_n = \frac{\sum_{k:k \neq n} \mu_{n,k} N_k^\alpha V_n + \beta N_n^\alpha V_n}{\sum_{k:k \neq n} \mu_{n,k} N_k^\alpha + \beta N_n^\alpha} \quad (1)$$

ここでは、 N_k を対応付けの信頼度と考え、差分ベクトルへの重みとしている。 α は N_k の寄与度を $\mu_{n,k}$ と同じ程度にするための定数であり、 β はあらかじめ定められる V_n の信頼度である。

ステップ3: スムージングされた差分ベクトル V_n を用いてすべての変換コードブックを更新する。

本実験では、fuzziness=2.0、 $\alpha = 0.0$ 、 $\beta = 1.0$ とした。

[認識]

変換コードブックを用いて未知話者の音声を量子化し、標準話者のHMMにより認識を行う。

2.3 ファジィベクトル量子化

本実験で用いているファジィベクトル量子化について説明する。

ファジィベクトル量子化は、入力ベクトルを既存のコードベクトルに対する帰属度を用いて表現する方法である。帰属度を表す級関数 (u_i) は、次式で表される。

$$u_i = \frac{1}{\sum_{j=1}^c (d_i/d_j)^{1/(m-1)}}$$

m はファジネスとよばれる曖昧度を表す値である。 c は入力ベクトルを幾つのコードベクトルで表すかを示す k -近傍数である。 d_i は入力ベクトル (x) と未知話者のコードベクトル (v_{iA}) の距離を表す。話者変換後のコードベクトルは次式に従い、ファジィ級関数とファジネスおよびコードベクトルより求められる。

$$x \Rightarrow x' = \frac{\sum_{i=1}^c (u_i)^m v_{iA}}{\sum_{i=1}^c (u_i)^m} \Rightarrow b' = \frac{\sum_{i=1}^c (u_i)^m v_{iA-B}}{\sum_{i=1}^c (u_i)^m}$$

ただし、 x は入力ベクトル、 x' は入力ベクトル x をファジィ VQ することにより求めたベクトル、 b' は入力ベクトル x に対する話者変換後のベクトル、 v_{iA} は未知話者のコードベクトル、 v_{iA-B} は変換コードベクトルを表している。

本実験では、 $m = 1.5$ 、 $c = 6$ とした。

3 実験

音素認識実験 (/b,d,g,m,n,N/) により、VFM の雑音抑圧性能及び、雑音環境下における話者適応性能を評価する。3.1 節では実験条件を述べる。3.2 節では、雑音重畳音声空間 (MAU) を同一話者 (MAU) の雑音の無い空間へマッピングすることにより、VFM の雑音抑圧性能を評価する。3.3 節では、雑音重畳音声空間 (MAU) を異話者 (MHT) の雑音の無い空間へマッピングすることにより、雑音環境下における話者適応性能を評価する。

3.1 実験条件

分析条件を表 1 に、本報告で行なう実験における共通の条件を表 2 にそれぞれ示す。

表 1: 分析条件

標本化周波数	12 kHz
フレーム長	256 点 (21.3ms)
フレームシフト幅	36 点 (3ms)
高域強調	$1 - 0.97z^{-1}$
LPC 分析	14 次
CEP, Δ CEP 計算打ち切り次数	16 次

表 2: 実験条件

HMM	4 状態 3 ループ / b,d,g,m,n,N / の 6 音素 / b,d,g/ は語頭と語中の 2 種類
HMM 学習資料	重要語 5240 単語の偶数番目の単語中の音素 各 HMM 当たり最大 200 サンプル
認識資料	重要語 5240 単語の奇数番目の単語中の音素 各音素当たり最大 200 サンプル
特徴量 (コードブックサイズ)	cep(256)、 Δ cep(256)、 Δ 対数 power(64)
標準話者用コードブック作成資料	音素バランス 216 単語の先頭から 100 単語
雑音重畳資料	有色雑音を音声波形に重畳

3.2 雑音抑圧性能の評価

本実験では、雑音が重畳した MAU の音声空間を、雑音の無い MAU の音声空間にマッピングすることにより、VFM の雑音抑圧能力を評価する。また、CBH 及びマッピング等の雑音処理を行わない場合の認識実験を比較として行う。

HMM は、MAU の 1 回目発声の音素バランス 216 単語の先頭から 100 単語を用いて作成されたコードブックと 5240 単語中の偶数番目単語を用いて学習した。認識資料は、MAU の 5240 単語中の奇数番目単語に有色雑音を計算機上で重畳したものをを用いた。

VFM の雑音適応学習資料には、MAU の 2 回目発声の音素バランス 216 単語の先頭から 5、10、25 単語に、SNR が ∞ 、30、20(dB) の雑音を重畳したものをを用いた。比較として行う CBH の雑音適応学習資料には、MAU の 2 回目発声の音素バランス 216 単語の先頭の 25 単語を用いた。また、雑音処理を行わない場合 (Speaker Dependent: 以下 S.D) の認識結果との比較により VFM 及び CBH の雑音抑圧性能を検討した。S.D とは、MAU の 1 回目発声の音素バランス 216 単語の先頭から 100 単語を用いて作成されたコードブックを用いて、雑音の重畳した入力音声を量子化することにより雑音重畳音声を認識する方法である。

実験結果を図 1 に示す。図中の認識率は、HMM の学習にファジィベクトル量子化 (Fuzzy VQ) を用いた場合と用いない場合 (Hard VQ) 及び、入力音声を Fuzzy VQ した場合としない場合の 4 通りの組合せのうち、認識率の最も高かったものを示してある。各組合せによる認識率は、付録として表 3、表 5 に示した。

図 1 より、高雑音下 (SNR=20dB) では VFM 及び CBH は S.D よりも高い認識率を示しており、両マッピング手法は高雑音下において有効な手段であることが分かる。また、VFM は CBH よりも高い雑音抑圧能力をもつことが分かる。雑音の無い環境 (SNR= ∞ dB) では、S.D との比較において、VFM は CBH よりも認識率の低下が少ないことが分かる。これは、本実験が同一話者間におけるマッピング (少数単語のなす狭い空間から、同一話者の広い空間を推定する) であるため、VFM の差分ベクトル推定が高精度に行なわれたためであると考えられる。CBH は、学習単語 (25 単語) から得られる狭い空間 (コードブック) を用いて、未知入

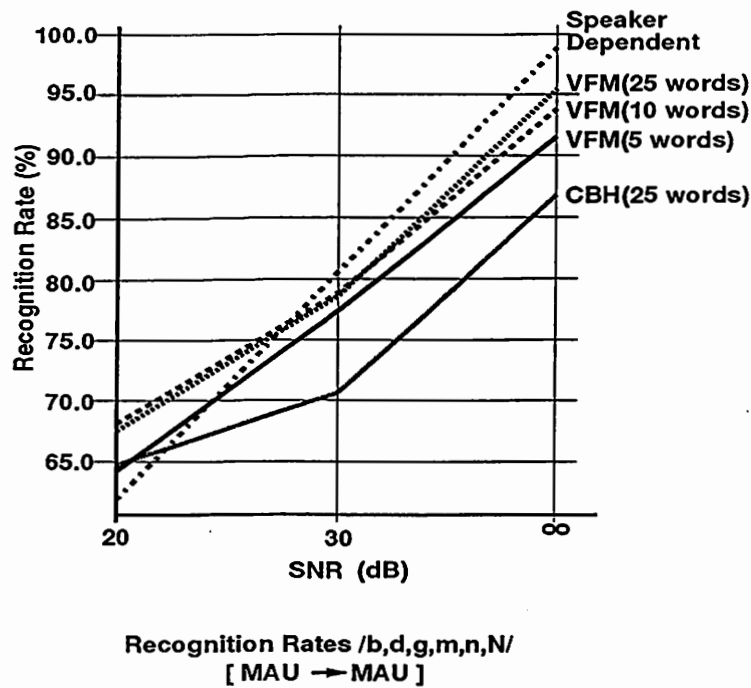


図 1: 雑音抑圧性能評価実験結果

力を量子化するため、未知入力 of 量子化時の歪みが大きくなることと、狭い空間から標準話者の広い空間へのマッピングを行なうことによる曖昧性のため、認識率が低下するものと考えられる。

学習単語数の増加により、SNR=20dB における VFM の認識率の変化を検討するため、学習単語を 30 単語及び 50 単語に増加した場合の認識実験を行なったが、認識率の向上は認められなかった。(30 単語:68.19%、50 単語:67.43%)

また、SNR=∞ dB の環境で 1 回目発声を学習単語に用いた場合の VFM の認識結果を表 7 に示しておく。本実験では、学習時において同一の発声資料が用いられるため、DTW の対応づけが高精度に行なわれることによる認識性能の向上が考えられるが、認識性能に変化はなかった。

3.3 雑音環境下における話者適応性能評価

本実験では、雑音環境下における話者適応性能の評価を行なう。

HMM は、MHT の音素バランス 216 単語の先頭から 100 単語を用いて作成されたコードブックと 5240 単語中の偶数番目単語を用いて学習した。認識資料は、MAU の 5240 単語中の奇数番目単語に有色雑音を計算機上で重畳したものをを用いた。

VFM の雑音適応学習資料には、MAU の 1 回目発声の音素バランス 216 単語の先頭から 5、10、25 単語に、SNR が ∞、30、20(dB) の雑音を重畳したものをを用いた。比較として行う CBH の雑音適応学習資料には、MAU の 1 回目発声の音素バランス 216 単語の先頭の 25 単語を用いた。また、話者適応を行わない場合 (Without Adaptation: 以下 W.A) の認識結果

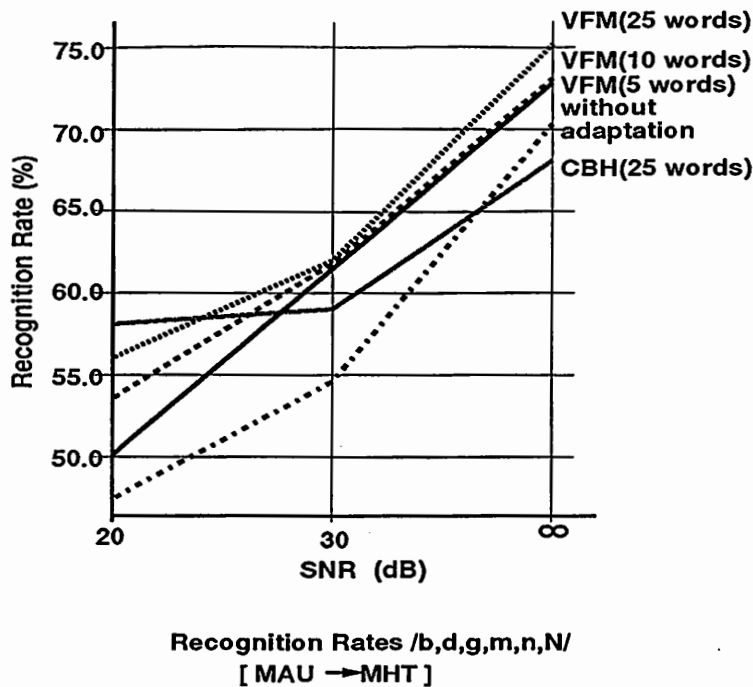


図 2: 雑音環境下における話者適応性能評価結果

との比較により、VFM 及び CBH の雑音環境下における話者適応性能を検討した。W.A とは、MHT の音素バランス 216 単語の先頭から 100 単語を用いて作成されたコードブックを用いて、雑音の重畳した MAU の入力音声量子化し、MHT の HMM で MAU の雑音重畳音声を認識する方法である。

図 2 に実験結果を示す。図中の認識率は、HMM の学習にファジィベクトル量子化 (Fuzzy VQ) を用いた場合と用いない場合 (Hard VQ) 及び、入力音声を Fuzzy VQ した場合としない場合の 4 通りの組合せのうち、認識率の最も高かったものを示してある。各組合せによる認識率は、付録として表 4、表 6 に示した。

図 2 より、雑音環境下において両適応方式とも話者適応を行わない場合よりも高い認識率を示しており、雑音環境下においても話者適応方式は機能していることが分かる。また、VFM は HMAP よりも高い話者適応能力をもつことが分かる。SNR=∞ において HMAP の認識率は話者適応を行わない場合の認識率よりも低くなっている。これは標準話者と未知話者が共に男性であり、話者間の音声空間が近く、話者適応を行わない場合よりも少量学習単語を用いて作成した話者空間を表すには不十分なコードブックを用いた適応誤差の方が大きくなったためであると考えられる。SNR=20dB では、VFM は HMAP よりも低い認識率を示している。これは、移動ベクトルに対する連続性の拘束条件の強さが SNR により異なるためであり、拘束条件を制御するためのファジネスを今後検討する必要がある。

4 まとめ

音素認識実験 (/b,d,g,m,n,N/) により、VFM の雑音抑圧性能及び、雑音環境下における話者適応性能を評価した。本実験結果より、以下のことが明らかになった。

1. 雑音抑圧性能比較実験結果

- (a) 雑音の無い環境 (SNR= ∞ dB) では、VFM は CBH よりも認識率の低下が少ない
- (b) 高雑音下 (SNR=20dB) では VFM 及び CBH は雑音抑圧能力を示す。
- (c) VFM は CBH よりも高い雑音抑圧能力をもつことが分かる。

2. 雑音環境下における話者適応性能評価結果

- (a) 雑音環境下においても、VFM を用いた話者適応は可能
- (b) 約 30dB 以上の環境において VFM は CBH よりも高い話者適応能力をもつ

今回の実験では、VFM に関するパラメータ (式 (1) 中の α, β 及び fuzziness 等) の雑音環境下における検討を行なわなかったが、今後検討していく必要がある。

謝辞

実習の機会を与えて頂いた ATR 自動翻訳電話研究所榎松明社長をはじめ嵯峨山茂樹室長ならびに御討論いただいた音声情報処理研究室の皆様へ感謝致します。

参考文献

- [1] 中村, 花沢, 鹿野: "ベクトル量子化話者適応アルゴリズムの HMM 音韻認識による評価," 信学技報, SP88-106 (1988).
- [2] 大倉, 服部, 杉山, 鹿野: "コードブックマッピングを用いた雑音環境下での音声認識," 信学技報, SP90-12, pp.25-32 (Jun. 1990).
- [3] 服部, 嵯峨山: "少量学習データを用いたコードブックマッピングによる話者適応化," 音講論集, 1-5-23 (Mar. 1991).

5 付録

[1] 雑音抑圧性能の評価

表 3: /b,d,g,m,n,N/ の音素認識結果
 hard:HardVQ
 fzy:FuzzyVQ
 fuzziness1.5 近傍数 6

VFM (5 単語)

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	91.60	99.41	99.92
	fzy	hard	91.52	98.73	99.58
	hard	fzy	91.18	99.32	99.92
	hard	hard	87.11	97.46	99.24
30	fzy	fzy	72.35	86.85	92.88
	fzy	hard	77.35	90.42	95.67
	hard	fzy	76.08	91.69	96.61
	hard	hard	73.96	90.50	95.67
20	fzy	fzy	64.21	83.55	91.94
	fzy	hard	63.44	84.22	93.21
	hard	fzy	64.04	85.07	94.74
	hard	hard	60.81	80.07	89.65

VFM (10 単語)

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	92.79	99.41	99.83
	fzy	hard	93.81	99.41	100.00
	hard	fzy	92.88	99.41	100.00
	hard	hard	90.92	98.81	99.92
30	fzy	fzy	73.79	86.43	91.09
	fzy	hard	78.88	91.43	95.93
	hard	fzy	75.06	88.72	95.42
	hard	hard	74.89	91.18	96.86
20	fzy	fzy	65.65	82.87	90.16
	fzy	hard	68.11	83.63	92.37
	hard	fzy	64.29	84.65	93.47
	hard	hard	62.26	81.17	90.67

VFM (25 単語)

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	93.47	99.32	100.00
	fzy	hard	95.42	99.83	100.00
	hard	fzy	94.06	99.41	100.00
	hard	hard	92.54	99.32	100.00
30	fzy	fzy	74.22	86.01	91.01
	fzy	hard	78.54	91.52	96.18
	hard	fzy	76.17	90.42	95.50
	hard	hard	78.63	92.45	97.29
20	fzy	fzy	67.51	83.88	91.86
	fzy	hard	66.58	84.05	94.40
	hard	fzy	67.18	84.73	93.89
	hard	hard	63.36	83.21	92.28

CBH (25 単語)

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	86.77	98.22	99.83
	fzy	hard	81.17	95.67	98.73
	hard	fzy	86.34	97.54	99.49
	hard	hard	70.82	89.31	97.29
30	fzy	fzy	70.65	84.99	93.04
	fzy	hard	65.48	86.34	94.32
	hard	fzy	68.45	88.04	95.67
	hard	hard	61.58	82.27	91.35
20	fzy	fzy	64.72	82.53	93.13
	fzy	hard	57.76	82.10	94.06
	hard	fzy	64.29	86.09	94.83
	hard	hard	56.57	77.27	89.14

[2] 雑音環境下における話者適応性能評価

表 4: /b,d,g,m,n,N/ の音素認識結果
 hard:HardVQ
 fzy:FuzzyVQ
 fuzziness1.5 近傍数 6

VFM (5 単語)

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	72.86	90.75	97.54
	fzy	hard	64.63	86.34	95.17
	hard	fzy	66.92	87.62	96.52
	hard	hard	62.68	82.87	92.88
30	fzy	fzy	61.49	78.71	88.38
	fzy	hard	57.34	79.13	91.01
	hard	fzy	57.76	78.80	91.52
	hard	hard	54.71	76.00	88.63
20	fzy	fzy	47.84	68.62	84.56
	fzy	hard	50.13	70.74	85.41
	hard	fzy	46.56	69.89	84.05
	hard	hard	44.19	62.51	79.13

VFM (10 単語)

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	73.11	91.18	97.63
	fzy	hard	68.96	89.40	96.61
	hard	fzy	68.96	90.16	97.12
	hard	hard	65.73	84.90	94.49
30	fzy	fzy	61.83	79.56	89.14
	fzy	hard	58.35	79.22	90.33
	hard	fzy	57.42	79.73	90.84
	hard	hard	57.08	78.46	89.91
20	fzy	fzy	53.60	74.13	86.60
	fzy	hard	52.42	73.28	86.85
	hard	fzy	51.31	75.66	89.14
	hard	hard	46.40	67.94	83.04

VFM (25 単語)

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	75.15	92.45	97.12
	fzy	hard	68.11	89.06	95.93
	hard	fzy	72.77	92.28	97.96
	hard	hard	68.53	87.11	94.49
30	fzy	fzy	62.00	81.00	90.75
	fzy	hard	61.24	81.68	91.94
	hard	fzy	60.05	82.10	94.74
	hard	hard	59.29	80.75	90.92
20	fzy	fzy	56.06	74.81	86.68
	fzy	hard	54.79	75.66	88.55
	hard	fzy	55.39	78.54	90.33
	hard	hard	50.81	72.09	84.90

CBH (25 単語)

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	68.11	87.96	96.35
	fzy	hard	62.43	86.17	95.00
	hard	fzy	61.75	83.72	93.38
	hard	hard	59.29	81.42	91.01
30	fzy	fzy	59.03	81.93	93.30
	fzy	hard	57.34	79.13	91.18
	hard	fzy	56.57	80.24	93.13
	hard	hard	50.64	73.28	86.85
20	fzy	fzy	58.10	80.41	91.86
	fzy	hard	52.08	75.74	87.70
	hard	fzy	55.73	79.30	92.45
	hard	hard	43.34	64.89	79.73

[3] Speaker Dependent

表 5: /b,d,g,m,n,N/ の音素認識結果
 hard:HardVQ
 fzy:FuzzyVQ
 fuzziness1.5 近傍数 6

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	94.57	99.66	99.92
	fzy	hard	96.27	99.83	99.92
	hard	fzy	95.76	99.92	100.00
	hard	hard	98.81	100.00	100.00
30	fzy	fzy	76.34	89.74	94.32
	fzy	hard	78.88	90.59	95.84
	hard	fzy	77.10	91.52	97.54
	hard	hard	80.58	93.04	98.30
20	fzy	fzy	61.92	83.46	92.11
	fzy	hard	60.98	79.81	91.43
	hard	fzy	56.74	79.39	93.04
	hard	hard	58.95	78.37	89.48

[4] Without Adaptation

表 6: /b,d,g,m,n,N/ の音素認識結果
 hard:HardVQ
 fzy:FuzzyVQ
 fuzziness1.5 近傍数 6

SNR (dB)	VQ 方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
∞	fzy	fzy	70.31	91.09	97.12
	fzy	hard	63.61	85.84	93.47
	hard	fzy	69.64	89.06	97.20
	hard	hard	61.75	82.77	91.60
30	fzy	fzy	54.62	76.17	86.17
	fzy	hard	51.15	72.69	86.09
	hard	fzy	50.64	74.22	86.34
	hard	hard	47.24	69.80	84.22
20	fzy	fzy	44.27	63.87	77.86
	fzy	hard	47.50	70.74	83.97
	hard	fzy	39.19	59.12	78.12
	hard	hard	39.53	60.81	78.12

[5] 1回目発声のバランス単語 (MAU) を用いた適応

表 7: /b,d,g,m,n,N/ の音素認識結果
 hard:HardVQ
 fzy:FuzzyVQ
 fuzziness1.5 近傍数 6

単語数	VQ方法		Recognition Rate (%)		
	学習	認識	1st	2nd	3rd
5	fzy	fzy	92.88	99.32	99.92
	fzy	hard	94.74	99.49	99.83
	hard	fzy	93.38	99.32	99.92
	hard	hard	92.11	98.64	99.83
10	fzy	fzy	94.06	99.41	99.92
	fzy	hard	96.10	99.66	99.92
	hard	fzy	94.74	99.58	100.00
	hard	hard	95.00	99.49	99.92
25	fzy	fzy	93.98	99.41	99.92
	fzy	hard	96.10	99.49	99.92
	hard	fzy	95.42	99.66	100.00
	hard	hard	94.74	99.75	100.00