

論文 / 著書情報
Article / Book Information

論題(和文)	パワー情報で重みづけた距離による単語音声認識
Title	
著者(和文)	相川 清明, 鹿野 清広, 古井 貞熙
Author	SADAOKI FURUI
出典(和文)	日本音響学会 音声研究会資料 S81-59, , , pp. 469-476
Journal/Book name	, , , pp. 469-476
発行日 / Issue date	1981, 12

パワー情報で重みづけた距離による 単語音声認識

相川 清明 鹿野 清宏 古井 貞照
(日本電信電話公社, 武蔵野電気通信研究所)

昭和56年12月22日 於 武蔵野電気通信研究所

内容梗概 スペクトル情報とパワー情報を考慮した音声フレーム間マッチング距離を用いた単語音声認識について述べる。この方法は従来のスペクトル距離と新たに導入されたパワー距離から成るフレーム間マッチング距離に基づいてDPマッチングにより単語音声と認識するものである。この方法を用いて641都市名認識実験を行った結果、標準パターンと入力音声の話者が同一の場合には、従来のスペクトル距離のみを用いた方法と同程度の認識率であったが、話者が異なる場合には、4人の話者の組合せの平均で88%の認識率を得、従来の方法より5.6%の向上を見た。

An Isolated Word Recognition Method Using Power-Weighted Spectral Matching Measure

Kiyoaki AIKAWA, Kiyohiro SHIKANO and Sadaoki FURUI
(Musashino E.C.L., N.T.T.)

December 22, 1981 at Musashino E.C.L., N.T.T.

Abstract This report describes an isolated word recognition method using inter-frame matching measure which is calculated from spectral information and power information. Each word utterance is recognized by DP-matching with word reference templates based on the inter-frame matching measure. 641 word recognition experiments show that this method provides 88% recognition accuracy in case that the speaker of input speech is different from the speaker of reference speech. This rate is 5.6% higher than that by the ordinary method using spectral distance only. When the speaker of input speech and the speaker of reference speech are identical, this method has almost the same performance as the ordinary method.

1 はじめに

本報告では、スペクトル情報とパワー情報を
用いた単語音声認識について述べる。従来、単
語音声認識にはDPマッチング法が広く用いら
れてきた。^{(1)~(6)} DPマッチング法にはさまざ
まな方法があるが、いずれの場合にも基本的
には各音声フレームのパワーを一定に正規化し
たスペクトル時系列間の時間正規化マッチング
を行なうという方法をとっている。この方法
では標準パターンと入力音声の話者が同一
の場合には同一単語間の対応する部分の
スペクトルマッチングが良いので良好な認
識能力を示す。しかし標準パターンと入力
音声の話者が異なる場合には、同一単語
間の対応する部分のスペクトルマッチング
が良好であることが保証されないため、
認識能力が劣化する。⁽⁷⁾

一方、本研究において単語音声のパワー時
系列パタンの持つ情報について検討した結
果、単語認識におけるパワー情報の有用性
が判明した。パワー情報とスペクトル情報
は一応別個の情報と考えられる。従って、
それらの併用により認識能力の向上が期
待できる。そこで、特に、多数話者音声
認識を目的として、パワーの重みをつけた
スペクトルマッチング距離を用いたDPマ
ッチング法を考案した。本報告では、この
パワー重み付け距離を用いた単語音声認
識法について述べ、多数話者による641
都市名認識実験結果について述べる。

2 パワーパタンの持つ情報

2.1 パワーパタンの有用性

従来のDPマッチングでは、音声フレーム
内のパワーを一定に正規化した入力音声と
標準パタンのフレーム間スペクトルマッ
チング距離に基づいてそれぞれのフレーム
系列の時間軸を伸縮させてマッチングを行
なっていた。この場合入力音声の話者と
標準パタンの話者が同一の場合には同一
単語のスペクトル時系列の再現性が高い
ので他の単語から良好に識別できる。し
かし、入力音声と標準パタンの話者が異
なる場合には同一単語でも同一部位の
スペクトルマッチングが良好でなく、同
一単語内のマッチング距離が小さくなら
ないため諸別しにくくなる。

表1は距離としてWLRスペクトルマッ
チング距離⁽⁹⁾を用いて求めた641都市名
の単語間マッチング距離を4人の話者の
すべての組み合わせについて求め整理し
たものである。表1では、入力音声と標
準パタンの話者が同一の場合と異なる
場合に分け、それぞれについて同一単
語のマッチング距離の平均と異なる単
語間のマッチング距離の平均を示して
いる。

	同一単語	異なる単語
同一話者	51.18	133.8
異なる話者	74.30	136.4

表1. 641都市名単語間距離の平均
(話者:4名, WLRスペクトルマッチング尺度使用)

この表からわかるように、標準パターン
と入力音声の話者が同一であっても異な
る単語間のマッチング距離はほとんど
変わらない。しかし、同一単語のマッ
チング距離では大きな差があることが
わかる。

一方、音声パワーの時系列パターンは、
母音部では高く子音部では低いという
基本的な性質を持つ。この性質は話者
が異なっても不変の性質である。図1
は4人の話者がそれぞれ2回ずつ発声
したsapporoという単語のパワーパ
タンである。この図から、パワーパ
タンは話者が変わってもあまり差
がないことがわかる。従

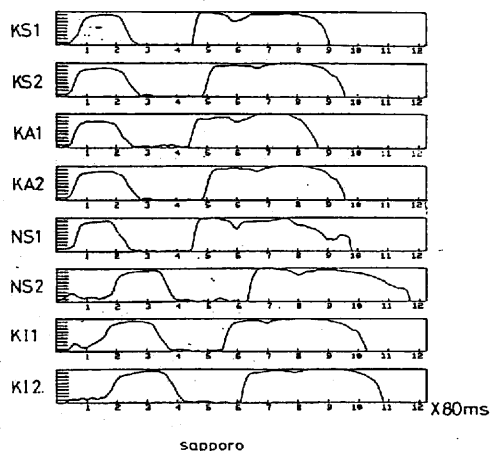


図1 4人の話者の sapporo という音声の
パワー時系列パターン。(対数パワーパ
ターンを、最大値と最小値が一定に
なるように正規化したもの)

果の方法ではパワーの情報を見捨てていたために、子音部と母音部のスペクトル距離が子音部間あるいは母音部間のスペクトル距離に比較してあまり大きくなければ、子音部と母音部がマッチングするということが起り得た。しかし、パワー情報を用いれば母音と子音をマッチングしにくくできると考えられる。

2.2 パワーパタンの正規化

パワーパタンの情報を用いるということは、標準パターンと入力音声のパワーパタンの比較を行なうという操作をしなければならぬ。このためにはパワーパタンの何らかの正規化が必要である。何故ならば、収集された音声データは、発声ごとに強度のばらつきがあり、かつ背景雑音が混入しているからである。すなわち、無声部分においては背景雑音がそのまま現われていることになる。従って、発声された音声パワーパターンが同一であっても雑音との相対強度が変わればパワーパターンは異なったものとなる。このような音声対雑音の相対強度の変化や音声パワーパターン自体のばらつきは常に起こるものであるから、これらの現象が起こっても母音部同志、子音部同志がそれぞれ良く適合するような正規化を行なう必要がある。

本研究では以上の条件をほぼ満たす正規化法として対数最大最小パワー正規化を用いた。すなわち1つの単語のパワーパターンを $p(t)$ としたとき、

$$AL = \max_t (\log(p(t))) \quad (2.1)$$

$$IL = \min_t (\log(p(t))) \quad (2.2)$$

$$P(t) = \frac{\log(p(t)) - IL}{AL - IL} \quad (2.3)$$

で与えられるものを用いた。この正規化の特徴は、

- ① 音声と背景雑音の強度比が変わっても無声部と同一の値にできる。
- ② 対数関数の性質から、パワー強度の大きい部分は圧縮率が高いので、母音部はすべてほぼ1に近い値をとる。
- ③ 対数関数の性質から、最大パワーの変動に対してパワーパタンの変形が小さい。
- ④ 欠点としては、低パワー部が拡大されるので、

低パワー部での特殊なパワーパタンの変動に対して弱くなる点がある。(ノイズのバースト等)の4つが挙げられる。③については、実際に最大パワーの変動を調べてみた結果では +100% - 50% 程度の変化をするが、対数の場合には ±10% 程度の変動にしかならないことがわっている。

2.3 パワーパタンの持つ単語個有の情報

パワーパターンが単語識別に関してどのくらいの情報を持っているかを調べた。これを調べるには、多くの単語間のパワーパタンの類似度と、同一単語のパワーパタンの類似度にどのくらいの差があるかを調べれば良い。類似度の測定には従来の時間正規化マッチング法を用いることができる。ここでは、入力音声フレーム i と標準パターンフレーム j とのマッチング距離 $D_p(i, j)$ を

$$D_p(i, j) = |P(i) - P^*(j)| \quad (2.4)$$

$*$: 標準パターンであることを示す。

$P : \rightarrow (2.3)$ 式

として傾斜制限付両端点フリーDPマッチングを行なった。⁽⁵⁾ 傾斜制限、ウィンドウ、端点領域を図2に示す。なおここで

$$IW = \frac{1}{4} \min(N, M) + 3 \quad (2.5)$$

$$IHW = \frac{1}{2} IW \quad (2.6)$$

とする。(Nは入力音声の、Mは標準パタンのフレーム数)

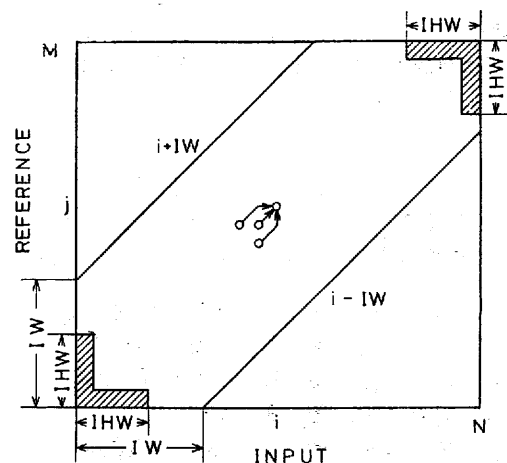


図2. 傾斜制限付、端点フリーDPマッチング
 ■ 端点領域 IW: ウィンドウ幅

これを用いて 641 都市名音声データについて、同一単語の類似度と異単語間の類似度を求めた。データ収集条件等を表 2 に示す。

語い : 641 日本国内都市名
AD 変換 : 12 bit
データウィンドウ : ハミング (幅 32 ms)
フレーム周期 : 16 ms
プリアンファシス : 1-2⁻¹
話者 : 男性 4 名 (各 2 回発声)
環境 : 計算機室 (騒音 70 dBA)

表 2 641 都市名音声データ

表 2 のデータについて計算した類似度の平均を表 3 に示す。ただし 2 つの単語の類似度を計算する場合、2 単語のフレーム数をそれぞれ、 M 、 N とすると、

$$|M - N| > \frac{3}{2} IW \quad (2.7)$$

の場合には、適合不能とみなして平均値計算から除いてある。

この表から、同一話者内では同一単語同志の類似度が、0.0374 であるのに対し、異単語間の類似度は 0.145 と大きく、また異話者間では同一単語の類似度が 0.0530 であるのに対し、異単語間では 0.147 と小さくなっていることがわかる。パワーパタンの類似度は 0 から 1 の値をとるので、パワーパタンは同一単語ではかなりの類似性があることがわかる。従って、パワー情報は単語を識別する上での有効な手がかりになると考えられる。

	SA	SS	DA	DS
同一話者内	3.74	1.33	14.53	5.25
異話者間	5.30	2.30	14.66	5.20

$\times 10^{-2}$

表 3 641 都市名のパワーパタンの類似度

SA 同一単語の類似度の平均
SS 同一単語の類似度の分布の標準偏差
DA 異単語間の類似度の平均
DS 異単語間の類似度の分布の標準偏差
同一話者内 KS, KA, NS, KI
異話者間 KS-KA, KS-NS, KS-KI,
KA-NS, KA-KI, NS-KI

2.4 パワー情報による単語の識別

次に、パワー情報の単語識別における有効性を調べるために、前記の単語間の類似度を用いて 641 単語の認識率を求めてみた。この結果を表 4 に示す。表 4 では、4 人の話者のすべての組合せについて 1 位から 10 位までの単語認識率を求め、それを同一話者同志の場合と異なる話者間の場合に分けて平均したものを示す。この表から、パワー情報のみで異話者間では 36% の単語が、また同一話者内では 67% の単語が 1 位で認識できることを示している。この認識実験では母音系列の情報さえも与えられていないのであるから、パワーパタンはかなりの単語固有の情報を持っていると言える。

RANK	1	2	3	4	5	10
同一話者内	66.9	79.0	84.4	87.3	88.9	93.1
異話者間	36.0	49.2	57.2	62.1	65.9	76.7

表 4 パワー情報のみによる 641 都市名認識実験結果 (単位 %)

3 スペクトル情報とパワー情報を用いた単語音声認識

3.1 パワー情報とスペクトル情報を用いた音声フレーム間距離

前章において正規化対数パワーパタンが同一単語と異単語ではかなりの差があることがわかった。従って、スペクトル情報を用いた従来の単語音声認識においてパワー情報と新たに付加すれば、認識能力の向上が期待できる。そこでスペクトルマッチング距離とパワー距離から成る音声フレーム間マッチング距離を手がかりにして時間正規化マッチングを行なう方法を考案した。入力音声フレーム i と標準パターンフレーム j とのスペクトルマッチング距離を $D_s(i, j)$ 、パワー距離を $D_p(i, j)$ とするとき、これらの音声フレーム間の距離 $D(i, j)$ は次式で与えられるとする。

$$D(i, j) = F(D_s(i, j), D_p(i, j)) \quad (3.1)$$

$$\frac{\partial D}{\partial D_s} \geq 0, \quad \frac{\partial D}{\partial D_p} \geq 0 \quad (3.2)$$

$$\text{ただし } D_s \geq 0, \quad D_p \geq 0 \quad (3.3)$$

とする。

ただし $F(a, b)$ は、 a と b の関数であることを示す。このような距離を用いて音声フレーム系列同志の時間正規化マッチングを行なえば、スペクトルとパワーレベルの両方が共に類似した部分がマッチングするので、母音と子音のマッチングを避けることができる。

関数 F の形としてはさまざまなものが考えられるが、最も簡単な形式として、

$$D(i, j) = D_s(i, j) + D_p(i, j) \quad (3.4)$$

$$D(i, j) = D_s(i, j) \cdot D_p(i, j) \quad (3.5)$$

の2通りが挙げられる。

$D_s(i, j)$ としては従来のケプストラム距離、WLR 距離等のスペクトルマッチング尺度を用いれば良い。

ケプストラム、あるいは WLR は、通常はパワーを正規化した計算式を用いるが理論上はパワーを含んだ計算式を導くことができる。パワーを含んだ計算式は

$$CEP = 2 \sum_{i=1}^{\infty} (C_i - C_i^*)^2 + (C_0 - C_0^*)^2 \quad (3.6)$$

$$WLR = \sum_{i=-\infty}^{\infty} (r_i - r_i^*) (C_i - C_i^*) + \frac{P}{P^*} + \frac{P^*}{P} - 2 \quad (3.7)$$

r_i i 次相関係数

C_i i 次 LPC ケプストラム係数

P 音声パワー ($C_0 = \log P$)

* 標準パターンであることを示す。

(無印は入力音声)

で表わされるが、これらの式を用いたのではパワー重み付けの因子が固定であり調整ができない。もしパワーパタンの再現性が良ければこれらの式を用いれば良いが、実際には音声強度の発声毎のばらつき、S/N 比の差異によるパワーパタンの変形、音声切出しミス、無声化等の影響があり、パワー重み付け因子の影響が強すぎるとかえってミスマッチングを起こす可能性がある。これらの影響を避け、なおかつ良好な認識率を得るためには、スペクトル因子とパワー因子の重みの調整が必要である。そこで、(3.4) 式、(3.5) 式において、 D_s はパワー成分を除いた従来のスペクトルマッチング距離を用い、パワー因子の重みを D_p として適当な関

数を与えることにより調整することにした。

パワーパターンは常にばらつきを伴っていることを考慮すると、 D_p は入力音声と標準パタンのパワー差が小さい場合にはパワー差に対する感度が鋭く、ある程度以上大きな差がある場合に感度が鋭くなるような関数が望ましい。このような関数のうち最も簡単なものは二乗曲線である。そこで D_p として入力音声フレームと標準パターンフレームのパワーの差の二乗の関数を用いることにした。

3.2 最適パワー重み付け関数の設定

スペクトル距離として WLR 尺度を用いた場合について、最適なパワー重み付け関数を 100 都市名認識実験により求めた。実験に用いた音声データは 2 章において用いた 641 都市名データのうちの最初の 100 都市名データである。話者の組合せとしては、同一話者内の例として、KS-KS を、また異話者間の例として KS-KA を用いた。フレーム間マッチング距離としては予備的な認識実験から (3.4) 式のタイプとして

$$D^2 = WLR + \alpha (P - P^*)^2 \quad (3.8)$$

と、また (3.5) 式のタイプとして

$$D^2 = WLR \cdot (1 + \alpha (P - P^*)^2) \quad (3.9)$$

を選び、それぞれ係数 α を変化させて認識率と誤差率⁽⁵⁾の変化を調べた。(誤差率：図 5 参照)

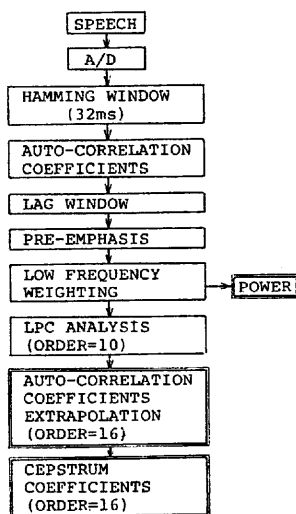


図 3 音声フレーム間マッチング距離計算用パラメータの抽出 (5x9)

ここでWLRはWLR尺度によるスペクトルマッチング距離で次式を用いている。

$$WLR = \sum_{i=1}^{16} (r_i - r_i^*) (C_i - C_i^*) \quad (3.10)$$

ただし、 C, r 等は(3.7)式と同様である。なおWLRの計算およびパワー距離計算用のパラメータの抽出方法を図3に示した。図4に基づいて(3.8)式と(3.9)式の最適係数を求めた結果、加算型では

$$D^2 = WLR + 6.0(P - P^*)^2 \quad (3.11)$$

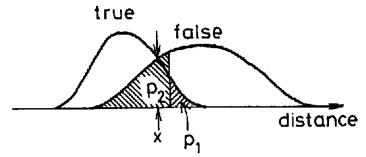
という関数が良く、乗算型では

$$D^2 = WLR \cdot (1 + 4.0(P - P^*)^2) \quad (3.12)$$

という関数が良いことがわかった。

図5 誤差率

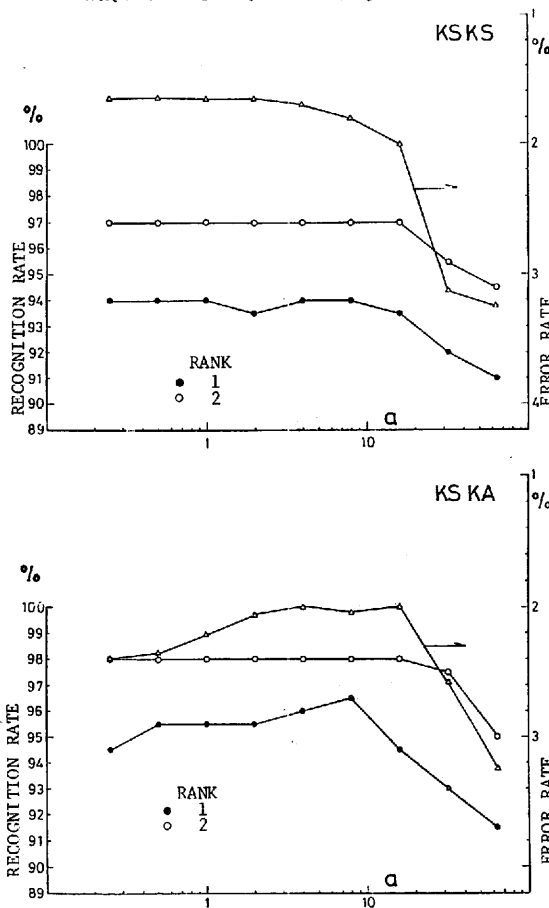
true: 同一
単語間マッ
チング距離
の分布



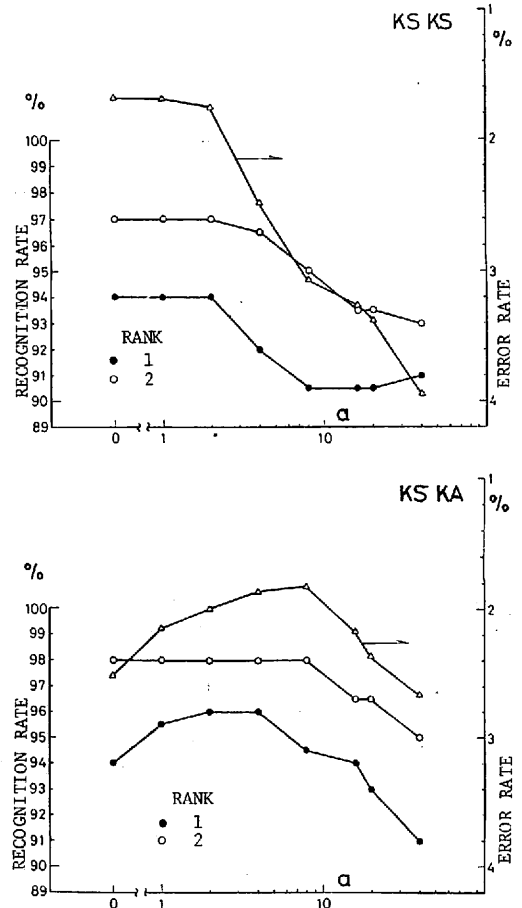
false: 異単語間のマッチング距離の分布
誤差率: $\min[(P_1 + P_2) / 2]$

4 大語い単語認識実験

スペクトル距離とパワー距離から成るフレーム間マッチング距離を用いて大語い単語音声認識実験を行ない、従来のようなスペクトル距離のみを用いた場合の結果と比較した。認識対象



(a) $D^2 = WLR + a(P - P^*)^2$



(b) $D^2 = WLR \cdot (1 + a(P - P^*)^2)$

図4 パワー重み付け量を変化させたときの10の都市名認識率と誤差率の変化。

a: パワー因子の重み係数。aが大きい方がパワーによる重みづけが強くなる。

は641都市名であり、音声データは2章において用いたものと同じである。また実験についても2章で述べたように4人の話者のすべての組み合わせについて行ない、同一話者内の場合と異話者間について整理した。この結果を表5に示す。A、B、Cはそれぞれフレーム間マッチング距離の種別を表わし、Aが従来の方法にあたる。また1位から5位までの認識率の平均を図6に示す。この結果から、

- ① (3.11)式と(3.12)式では、(3.12)式の距離を用いた方が0.2%程度1位認識率は良いが、有意な差はない。
- ② パワー情報とスペクトル情報から成る距離を用いた場合には、従来の方法に比べ、同一話者内での認識ではほぼ同程度の結果となったが、異話者間の場合では、82.4%→88.0%と大幅に認識率が改善されている。

という2点がわかる。

次にパワー重み付けにより従来の方法に比べてどのような改善が成され、どのような劣化が起こったかを調べた。まず、

- ① 従来の方法で認識結果が2位以下だった単語で、パワー重み付けにより1位に改善されたもの
- ② 従来の方法で認識結果が1位だった単語で、パワー重み付けにより2位以下に劣化したものの

を抜き出し、①については従来の方法における認識誤りを、②については本方法における認識誤りを分類した。分類項目は、

- ① モーラ数が異なる。
- ② 一部の母音のみが異なり、子音と他の母音は正しい。
- ③ 先頭子音のみが異なり他は正しい。
- ④ 一部の子音のみが異なり、母音と他の子音は正しい。
- ⑤ 一部の母音と子音が異なる。

の5種類で、①～⑤はモーラ数は等しいとする。劣化した場合にはさらにその原因が音韻の無声化によるものかどうかを調べた。この結果を表6に示す。

この表から、パワー重み付けによる認識率の改善の最も大きな理由はモーラ数の誤りの訂正であることがわかる。改善の理由は以下③④の誤りの訂正という順になっているが、いずれに

I	R	KS	KA	NS	KI	AVE.
C		95.2	97.0	97.3	97.3	96.7
B		95.2	97.5	97.0	97.5	96.8
A		95.8	97.9	97.6	97.6	97.0

(a) 同一話者内

I	KS			KA			NS	AVE.
R	KA	NS	KI	NS	KI	KI		
C	86.3	84.5	90.8	84.2	89.7	91.5	87.8	
B	85.8	83.9	90.9	85.0	90.3	91.9	88.0	
A	76.9	77.4	84.7	77.9	85.9	83.6	82.4	

(b) 異話者間

表5 641都市名認識実験結果(認識率:単位%)

I: 入力音声話者 R: 標準パターン話者

A: $D^2 = WLR$

B: $D^2 = WLR \cdot (1 + 4.0(P - P^*)^2)$

C: $D^2 = WLR + 6.0(P - P^*)^2$

してもパワー情報の利用により子音の位置が明確になり、標準パターンと入力音声の子音部分、母音部分がそれぞれ対応する部分とマッチングしやすくなったためであると考えられる。また、有聲子音と無聲子音の区別がパワー情報の利用により明確化されたことも改善の1つの理由であろう。逆に劣化の最も大きな要因もやはりモーラ数の誤りであるが、このうち約1/2が音韻の無声化によるものであることがわかる。

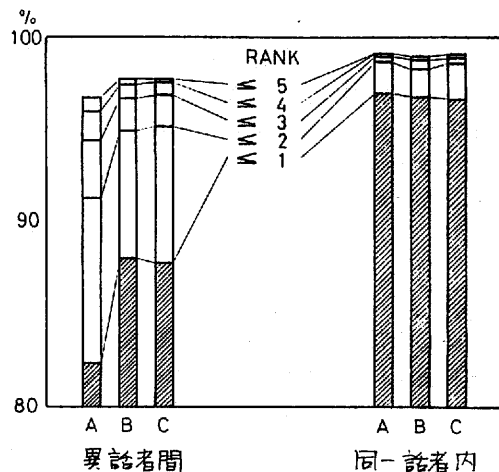


図6 641都市名認識実験平均認識率の比較

■ 1位認識率を示す。

A、B、Cについては表5参照。

5 むすび

音声フレーム間のスペクトル距離とパワー距離から成るマッチング距離を用いることにより、スペクトル距離のみを用いた場合に比べ単語音声認識において認識能力を高めることができることを示した。マッチング距離としては、スペクトル距離にパワー距離を加算するタイプと乗算するタイプについて試験を行なったが有意な差はなかった。認識実験は男性話者4人が2回ずつ発声した641都市名について行なった。パワー重み付け距離を用いた場合標準パターンと入力音声の話者が同一のときには認識率は従来のスペクトル距離のみによる方法とほぼ同程度であったが、標準パターンと入力音声の話者が異なるときには1位認識率は82.4%→88.0%と向上した。

パワー重み付け距離を用いることにより従来の方法における誤りが訂正されたものの従来の方法での誤りと分類したところ、モーラ数の誤りが最も多いことがわかった。他の誤りでは子音に関するものが多く、パワーパターンの利用により子音位置が明確化されたために以上のような誤りの訂正が行われたものと考えられる。

一方単語によっては音韻無声化のためにパワー重み付けにより新たに誤りが生ずることがあることがわかった。

従って今後は、パワーパターンの変形に対する考慮が必要である。

謝辞

日頃御指導頂く畔柳基礎部長、板倉第四研究室長に深く感謝する。また、有益な討論として頂いた四研諸氏に感謝する。

参考文献

- (1) Sakoe, Chiba, IEEE Trans. ASSP-26, 1978.
- (2) Furui, IEEE Trans. ASSP-29, 1981.
- (3) 中本, 中川, 信学研資 PRL 81-8, 1981.

- (4) Sakoe, IEEE Trans. ASSP-27, 1979.
- (5) 鹿野, 音学研資 S80-60, 1981.
- (6) 近江, 千葉, 信学全大論文集. 1260., 1973.
- (7) 鹿野, 杉山, 音学講論 3-1-19, 1981.
- (8) 杉山, 鹿野, 音学研資. S81-09, 1981.
- (9) 杉山, 鹿野, 信学論 J 64-A No.5, 1981.

		話者		モーラ	母音のみ	先頭子音のみ	子音のみ	母音+子音	計	無声化
		I	R							
同一話者内	改善	KS2	KS1			2	2		4	
		KS1	KS2	3			1		4	
		KA2	KA1	1		4	1	1	7	
		KA1	KA2	3			1		4	
		NS2	NS1						0	
		NS1	NS2			2	1		3	
		KI2	KI1	1			1		2	
		KI1	KI2	1			1		2	
		TOTAL		9		8	8	1	26	
	%		34.6		30.8	30.8	3.9	100.0		
	劣化	KS2	KS1	1		3	2	3	9	1
		KS1	KS2	4			1	2	7	3
		KA2	KA1	1		1			2	1
		KA1	KA2						0	
		NS2	NS1	1			3	1	5	
		NS1	NS2	1			2	2	5	
		KI2	KI1	1		1			2	
KI1		KI2	1		1		1	3		
TOTAL			10		6	8	9	33	5	
%		30.3		18.2	24.2	27.3	100.0	15.2		

		話者		モーラ	母音のみ	先頭子音のみ	子音のみ	母音+子音	計	無声化
		I	R							
異話者	改善	KS	KA	14	1	1	14	18	48	
		KA	KS	13		3	15	15	46	
		KS	NS	26	1	1	14	20	62	
		NS	KS	15			12	19	46	
		KS	KI	20		2	11	14	47	
		KI	KS	15		5	13	15	48	
		KA	NS	29		1	12	6	48	
	劣化	NS	KA	27		3	12	14	56	
		KA	KI	14			14	9	37	
		KI	KA	11			12	8	31	
		NS	KI	12		2	4	4	22	
		KI	NS	11			10	5	26	
		TOTAL	207		2	8	143	147	517	
		%	40.0		0.4	3.5	27.7	28.4	100.0	
同一話者	改善	KS	KA	1		1	2	3	7	
		KA	KS	5		2	2		9	2
		KS	NS	4		1	3	2	10	1
		NS	KS	4		5	3	5	17	1
		KS	KI	1		1			2	1
		KI	KS	3		3	2	3	11	1
		KA	NS	5			2	1	8	3
	劣化	NS	KA	3			2	1	6	
		KA	KI	4		5			9	4
		KI	KA	1			2	1	4	2
		NS	KI			1			1	
		KI	NS	1			3	1	5	
		TOTAL	32		19	21	17		89	16
		%	36.0		21.3	23.6	19.1		100.0	18.0

表6 パワー重みづけにより誤りが訂正された単語と新たに誤りが生じた単語の誤りの分類

I : 入力音声 R : 標準パターン