

論文 / 著書情報
Article / Book Information

論題(和文)	話者認識におけるスペクトルパラメータの時期的変動の効果
Title(English)	
著者(和文)	古井 貞熙, 板倉 文忠
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1973年講演論文集, , , pp. 159-160
Citation(English)	, , , pp. 159-160
発行日 / Pub. date	1973, 5

2-1-18 話者認識におけるスペクトルパラメータの時期的変動の効果*

古井貞照 坂倉文忠 (日本電信電話公社 武蔵野電気通信研究所)

I. まえがき 音声の基本周波数と部分自己相関係数 (PARCOR) の統計的パラメータを、複数の単語について組合わせることにより、よい認識率で話者の認識が行なえることをすでに報告した。⁽¹⁾⁽²⁾⁽³⁾ 本報告では、この方式におけるパラメータの時期的変動の効果について述べる。

II. 認識方法 種々の単語音声を、基本周波数と1~12次のPARCORを組合わせた13次元ベクトルの時系列で表現する。有声音区画における平均値 μ と、そのまわりでの標準偏差 σ 、パラメータ間相関行列 Λ を求める。 μ と Λ の成分から m 成分を選択し、これを μ と組合わせて13+m次元の特徴ベクトル Z とする。時期的変動を考慮した重みつき距離 D によって、話者識別(9名)および話者照合(詐称者26名)を行なう。

$$D(Z, \theta_k) = (Z - \theta_k) V^{-1} (Z - \theta_k)^T \quad \left(\begin{array}{l} V: \text{時期的変動に関する共分散行列} \\ \theta_k: \text{登録話者}k\text{の標準パターン} \end{array} \right)$$

III. 認識結果 既知サンプルが10日間程度の間に求めたものである場合、未知サンプルまでの時間間隔が6週間程度までであれば、100%の認識が可能であるが、3か月程度たつと誤りが急激に大きくなる。⁽²⁾ しかし既知サンプルが9か月程度にわたったものであれば、最終の既知サンプルから未知サンプルまでの時間間隔が3か月程度になっても、4単語を組合わせれば98~99%の認識率を得ることができる。⁽²⁾

IV. 長期間わたったのちの未知サンプルの認識 既知サンプルとして、3か月おきに4時期にわたって求めたものを用い、最終の既知サンプルから3か月後~12か月後のサンプルの認識を試みた。図1は、 $m=20$ とし、4単語を組合わせたときの認識結果である。このように、長期間わたった後のサンプルに關しても、時期がたつに従って、認識率のやや低下する傾向が見られる。

V. 既知サンプルと未知サンプルの時間的順序関係の効果 図2に示すように、既知サンプルと未知サンプルの時間的順序関係を5通りに変えて、認識実験を行なった。 $m=0$ すなわち μ 成分のみを用い、4単語を組合わせたときの認識結果を図3に示す。識別と照合のどちらの場合も、現実的なタイプAの場合よりも、中程のタイプの方が良い認識率を示し、全体的に山形の結果を示している。この結果は、用いている特徴が、3か月以上の長期間にわたっても、ゆっくり変化する性質を有することを示している。

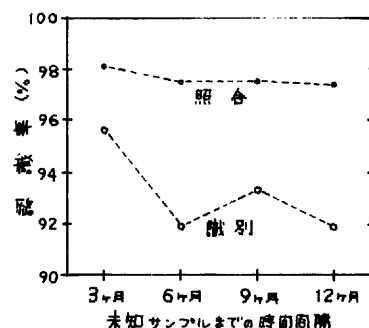


図1. 長期間わたったのちのサンプルの認識結果

VI. パターンの長期間にわたる変化 各話者の各単語ごとに、長期間(約3年間)

* Effect of Long-term Variations of Spectral Parameters on Talker Recognition. By Sadaoki Furui and Fumitada Itakura (Musashino E.C.L., N.T.T.)

と短期間(約4か月間)のそれぞれの場合について、 μ の成分に因する主成分分析を行ない、第1、第2主成分からなる二次元平面上に、各サンプルを射影した。分析は、パターンの変動に関する共分散行列を、相関行列に変換してから行なった。結果の一例を図4に示す。短期間のサンプルの時間的関係は下の通りである。

図4からわかる $a \rightarrow b \rightarrow c \rightarrow d \rightarrow e \rightarrow f \rightarrow g$
 ように、ほぼ10
 3月 2月 2月 2月 3月 3月
 日 日 日 日 週 週
 日間のb, c, d, e

の点は、他の点と比べてかなり接近しており、パターンが、短い期間内では安定しているが、長い期間たつと大きく変わってくる性質を示している。長期間の場合の各サンプル点は、3か月の間隔を置いて、かなりランダムに変動している。

次に、種々の時間間隔を置いたときの μ の値の変化量を、ユークリッド距離(の二乗)を用いて求め、

図5に示した。4単語と9名の話者について平均した結果を示した。この図を見ると、時間間隔が9か月程度になる

までは、時間間隔が長くなるに従って単調に変化量が大きくなる傾向があり、それ以上になると、時間間隔に無関係にほぼ一定になっている。 μ を構成する13成分のそれぞれについて同様の検討を行なったところ、どの成分も類似の性質を有することがわかった。

VII. むすび 特徴パラメータの時間的変動特性と、話者認識におよぼす効果が明らかになった。

ここで用いている特徴は、3か月以上の長い期間でも相関を有する不規則変動をなしているということができた。

[謝辞] 日御指導頂く 野田基経研究部長、青藤才4研究室長ならびに研究室の諸氏に感謝する。

[文献] 山古井・坂倉：音響学会講演集 2-2-20 (昭47-10) (2) 同：音声研究会資料 (昭47-11) (3) 同：信学会大 1258 (昭48-03)

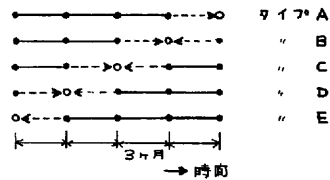


図2. 時間的順序関係

{ ○ 未知サンプル
● 既知 "

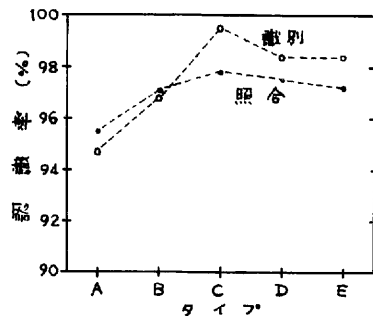


図3. 時間的順序関係と認識率の関係

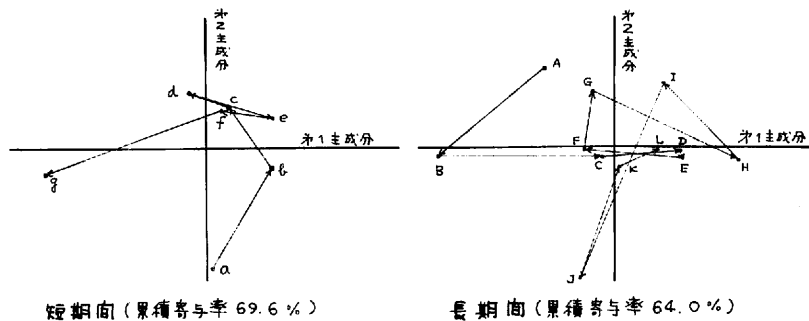


図4. パターン(μ)の長期間にわたる変化

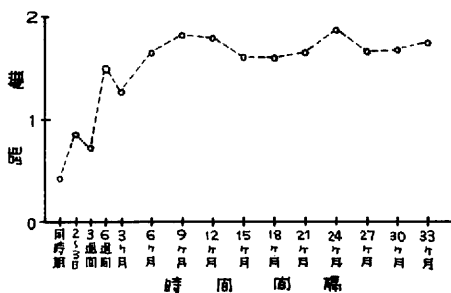


図5. μ の変化量