

論文 / 著書情報
Article / Book Information

論題(和文)	長時間平均スペクトルの経時変化を考慮した話者認識
Title(English)	
著者(和文)	古井貞熙, 板倉 文忠, 斉藤 収三
Authors(English)	SADAOKI FURUI
出典(和文)	昭和47年度電子通信学会全国大会, , , p. 56
Citation(English)	, , , p. 56
発行日 / Pub. date	1972, 7
URL	http://search.ieice.org/
権利情報 / Copyright	本著作物の著作権は電子情報通信学会に帰属します。 Copyright (c) 1972 Institute of Electronics, Information and Communication Engineers.

56 長時間平均スペクトルの 経時変化を考慮した話者認識

古井 貞熙 板倉 文忠 斉藤 収三

(日本電信電話公社 武蔵野電気通信研究所)

I. まえがき 音声から抽出した特徴パラメータを用いて、話者の機械認識を行なう場合に重要なことは、ある程度期間においてもほとんど変動しないパラメータを選択するか、あるいは変動するパラメータに関して、その変動を適切に正規化するモデルを構成することである。我々は話者の特徴を表現するパラメータとして長時間平均スペクトルをとり上げ、その長期間にわたった変動を考慮することにより、比較的良好な認識が可能であることを確認し、種々の検討を行なってきた。⁽¹⁾⁽²⁾

II. 長時間平均スペクトルの測定 10秒程度の連続音声について 32mssecの区画毎に短時間電力スペクトルに変換し、これを全体にわたって代数平均して長時間平均スペクトルとする。測定は各話者について、短期間のデータとして数日おき、長期間のデータとして3ヶ月おきに反覆して行なう。

III. 認識方法 長時間平均スペクトル(対数)をフーリエ変換してケプストラムとし、これをN次元ベクトルC(未知パターン)及び $\bar{D}_j$ (j番目の登録話者の既知パターン)で表わす。両者の距離を次のように定める。

$$S(C, \{\bar{D}_j\}) = |M|^{-1/2} (C - \bar{D}_j) M^{-1} (C - \bar{D}_j)^T, \quad \bar{D}_j = (\sum_{i=1}^n D_{ji}) / n, \quad (n \text{ は既知パターンの数})$$

N×N行列Mとして、共分散行列の対角要素以外を0におきかえたもの(方法I)、それを話者間で平均化したもの(方法II)、話者間で平均化した共分散行列(方法III)、及び単位行列(重みなし)を用いる。又(方法IV)として Nearest Neighbor法を試みる。話者識別については9名から1名を選ぶ形をとり、話者照合については各登録話者毎に8名のImpostorを置く。

IV. 既知サンプルとして短期間のデータを用いた認識実験 既知サンプルから未知サンプルまでの時間間隔を種々に変えて認識実験を行なった。9名の話者による平均認識率を図1に示す。(a)について話者毎の結果を用いて分散分析を行なったところ、時間間隔の効果が話者の効果が0.5%有意を示した。時間間隔が短い場合には重みづけなしでもかなりよい認識が可能であるが、時間間隔が長くなるに従い、種々の重みづけを行なっても認識率が単調に低下する傾向があり、又細かく見ると話者によってパターンの安定性に差のあることを示している。

V. 既知サンプルとして長期間のデータを用いた認識実験 既知サンプルとして、時期的に未知サンプルに近い1時期(3ヶ月前)のサンプルのみを用いる場合から、溯って6時期(1年半前まで)のサンプルを用いる場合までについて認識実験を行なった。(方法II)の場合の結果を図2に示す。既知サンプルが1時期のみのときには、低い認識率しか得られないが、一方、時期的に非常に遠い時点のサンプルは、むしろ捨てた方がよいことがわかる。

VI. 考察 各登録話者の長時間平均スペクトルをある程度長期間にわたって測定して蓄積しておき、ほぼ1ヶ月程度の周期で追加測定して古いものを更新するようにすれば、長期間にわたった後のサンプルについても、かなりよい認識率を達成することが可能であると思われる。

謝辞：日原御指導、同研究所情報基礎研究部長並びに研究室の皆様に感謝する。

文献：(1) 音響学会講演論文集 2-1-2 昭和46年5月 (2) 同 3-1-18 昭和46年11月

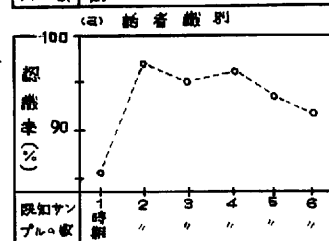
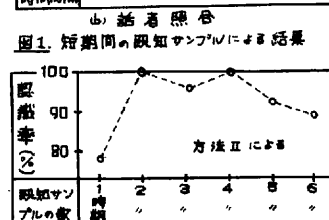
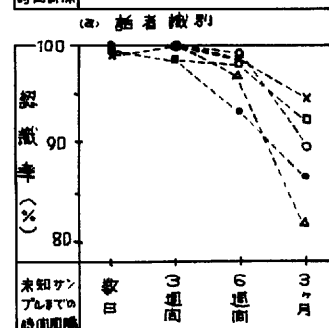
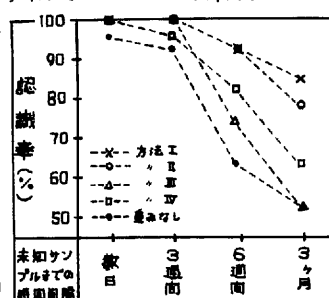


図2. 長期間の既知サンプルによる結果