

論文 / 著書情報
Article / Book Information

論題(和文)	長時間平均スペクトルによる話者照合
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)	日本音響学会1971年講演論文集, , , pp. 129-130
Citation(English)	, , , pp. 129-130
発行日 / Pub. date	1971, 5

2-1-2 長時間平均スペクトルによる話者照合*

古井貞熙 板倉文忠 斉藤収三 (日本電信電話公社 電気通信研究所)

1. はじめに 本報告は、音声の個人的特徴を表わすパラメータとして、今まであまり注目されることのなかった長時間平均スペクトルをとり上げ、男性9名について、長期間にわたってそのパターンの変化を調べるとともに、実際にそれを用いた話者の照合を試みたものである。

2. 長時間平均スペクトルの測定 下図のような方法で、長時間平均スペクトルの測定を行なった。

音声サンプル：単音から30秒程度の連続音声まで種々

話者：男性9名

測定回数：昭和45年5,6,7,12月, 昭和46年2月 (計5回)

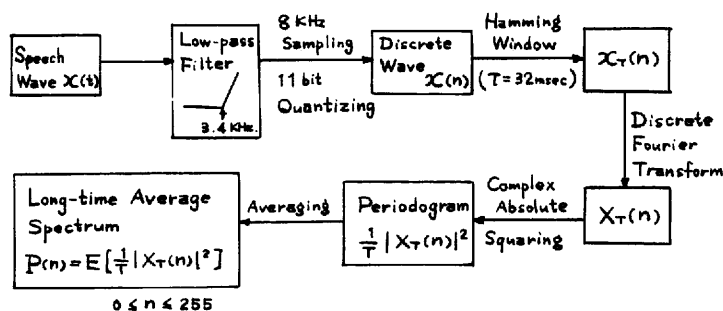


Fig. 1. Block diagram of measuring system

その結果、話者によって有意な差のあるパターンが得られた。Fig. 2 に 10~12秒程度の連続音声の 5, 6, 7, 12月に測定した長時間平均スペクトルの平均パターンの例を示す。

3. 話者照合 長時間平均スペクトルのケプストラムを N 次元の特徴ベクトルと考え、その N 次元超空間での重みづき距離による照合を行なう。この方法の特徴はケプストラム領域での距離は、容易に対数スペクトル領域での距離に対応づけられる上、ケプストラムに変換することによって、スペクトルパターンの大雑把な構造と微細構造を分離して取扱えるという点にある。

話者の照合は登録話者の既知パターンと未知パターンとの重みづき距離に閾値を設けて、話者の異同を判定する方法を用いる。未知パターンの j 次のケプストラムを $C(j)$ 、 r 番目の登録話者の i 番目の既知パターンの j 次のケプストラムを D_{rij} 、 r 番目の登録話者の j 次のケプストラムに対する重みを w_{rjj} とおくと、重みづき距離として次のようなものを用いる。

* Speaker Verification by the Long-time Average Spectrum.

By S. Furui, F. Itakura and S. Saito (E.C.L., N.T.T.)

$$S(C, \{D_{ri}\}) = \sum_{j=1}^N w_{rj}^2 (C(j) - \bar{D}_r(j))^2$$

$$\left(\begin{array}{l} \bar{D}_r(j) = \frac{1}{m} \sum_{i=1}^m D_{ri}(j) \\ m: r \text{ 番目の登録話者の既知} \\ \text{パターンの数} \\ w_{rj} = \left(\prod_{i=1}^N \sigma_{ri} \right)^{-\frac{1}{N}} \frac{1}{\sigma_{rj}} \\ \sigma_{rj}: r \text{ 番目の登録話者の } j \text{ 次の} \\ \text{ケプストラムの標準偏差} \end{array} \right)$$

4. 結果 男性9名の各話者について
次のような形で話者照合を行なった場合
の認識率は 9名で平均 97.2%であった。

登録話者の既知パターン $\{D_{ri}\}$:
5, 6, 7, 12月の測定データ
未知パターン $\{C\}$:
2月に発声した登録話者を含む
27名のデータ
登録話者のケプストラムの標準偏差 $\{\sigma_{ri}\}$:
5, 6, 7, 12月のデータから推定

Fig. 3は 登録話者Sの場合の、話者の
異同判定を行なう距離の閾値と、2種の
Error Rate との関係を示す。即ち、
Aは登録話者(データの数は30)を棄却
する誤り、Bは27名のうちの登録話者以
外の26名を 登録話者と間違える割合
を表わす。

5. むすび 今までに発表された種々
の話者照合の方法に比べて 簡単な方法
ながら よい認識を行なうことができた。
今後 最適な識別関数の決定などについ
て 更に実験と検討を進めて行きたい。

日頃御指導頂く当研究所聴覚基礎研究
部長 ならびに研究員の諸氏に感謝の意
を表する。

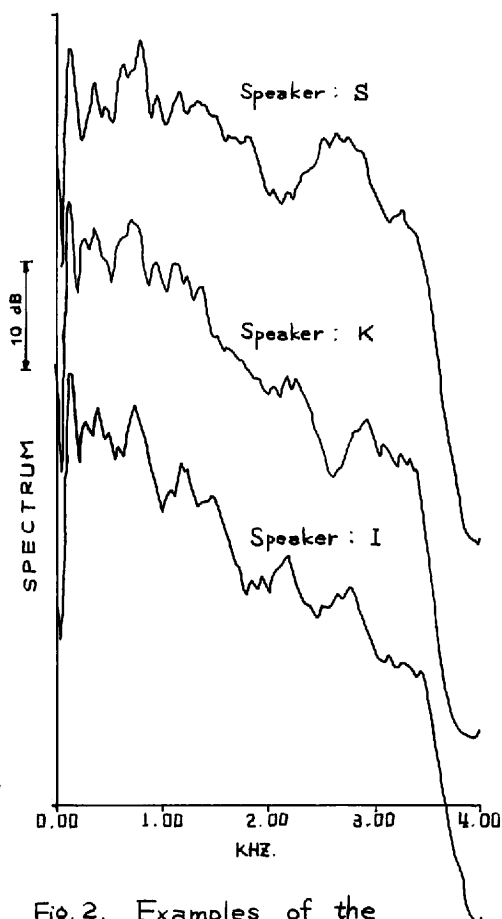


Fig. 2. Examples of the average pattern of the spectra

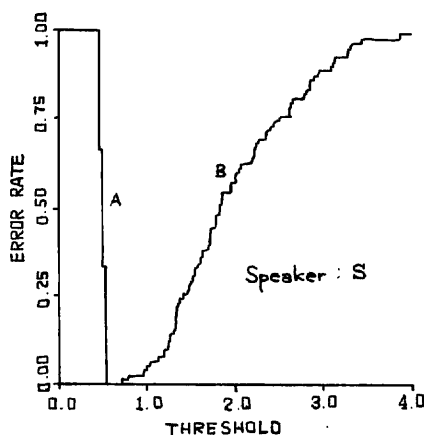


Fig. 3. Speaker verification errors vs. comparison threshold