

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)	電子通信学会論文誌, Vol. 55-A, No. 10, pp. 549-556
Citation(English)	, Vol. 55-A, No. 10, pp. 549-556
発行日 / Pub. date	1972, 10
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.

長時間平均スペクトルによる話者認識

正員 古井 貞熙† 正員 板倉 文忠†
正員 斉藤 収三†

Talker Recognition by the Longtime Averaged Speech Spectrum

Sadaaki FURUI†, Member, Fumitada ITAKURA†, Member
and Shuzo SAITO†, Member

あらまし 話者の特徴を示すパラメータとして、短文章の長時間平均(パワー)スペクトルをとり上げ、一年半以上の長期間にわたって多くの話者について、そのスペクトルパターンの変動を調べるとともに、その変動特性を考慮した話者認識を試みた。

その結果、そのパターンは2~3日から3週間程度までの間ではきわめて安定しているが、その後時期がたつにつれて、同じ発声者によるものでも種々に変化することが明らかになった。しかし、そのパターンの長期間にわたる変動を考慮した適切な距離尺度を設定することにより、3か月以上の長期間をおいた後でも、話者の認識が十分に可能であることが確認された。3か月おきに4時期にわたって測定した長時間平均スペクトルを既知サンプルに用いて、その変動を考慮した重みつき距離を設定することにより、その最終の時期の既知サンプルから3か月後の未知サンプルに関して、9名から1名を選ぶ話者識別の場合に平均91%、許称者が8名ないし26名の話者照合の場合に平均93~95%の認識率が得られた。

1. ま え が き

音声の個性を表現するパラメータとしてこれまで種々のものが検討されてきた。その中で比較的有望なパラメータの一つとして、音声を構成する音韻系列の影響が無視できると考えられる程度の長期間にわたった平均パワースペクトルを用いる試みがいくつか報告されている^{(1)~(7)}。

一般に話者認識の問題を取り扱う上で最も困難な点の一つは、ある程度長い期間をおくと、同じ話者が同じ言葉を発声しても特徴パラメータが変化してしまうところにある^{(1), (2)}。しかし、話者認識実験において、既知サンプルから未知サンプルまでの時間間隔の効果について積極的な検討を試みた例はまだきわめて少なく^{(3), (4)}、3か月程度の長期の時間間隔をおいた認識実験はまだ報告されていない。

本論文では、話者の特徴を示すパラメータとして短文章の長時間平均(パワー)スペクトルをとり上げ、そのスペクトルパターンの長期間にわたる変動に着目

して行なった実験的考察の結果について述べる。具体的方法としては、一年半以上の長期間にわたり多くの話者について長時間平均スペクトルパターンの変動を調べて、そのパターンの時期的変化に関する種々の特性を明らかにするとともに、その変動特性を考慮した話者認識を試み、既知サンプルから未知サンプルまでの時間間隔を3か月程度にまで長くにとって、その時間間隔と認識率との関係について検討する。

2. 長時間平均スペクトルの測定

2.1 測定方法

図1に示すように、しゃ音した部屋で発声された音声を3.4 kHz 低域通過フィルタに通し、8 kHz, 11 bit でサンプリングおよび量子化を行なって離散波形に変換する。これを順に256サンプル(32 msecに相当)ずつとり出して短時間パワースペクトル(密度)に変換する。この際、切り出しによるひずみの混入を防ぐため時間領域でハミングウィンドウをかけ合わせておく。音声波形全体について、順に32 msecの区分ごとに上記のようにして短時間パワースペクトルを求め、最後にそれを全体について平均して長時間平均スペクトルとする。

† 電電公社武蔵野電気通信研究所基礎研究部, 武蔵野市
Musashino Electrical Communication Laboratory, N.T.T.,
Musashino-shi, Japan 180
論文番号: 昭 47-410 [A-112]

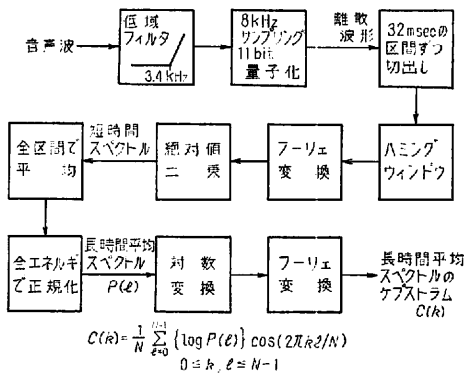


図1 長時間平均スペクトルの測定方法
Fig. 1—Block diagram of the measuring method.

この長時間平均スペクトルは全エネルギーで正規化し、対数変換した後にさらにフーリエ変換を行なってケプストラム⁽⁹⁾とする。定義から容易に知られるように、ケプストラムの低次の成分は対数スペクトルパターンの概形を示し、高次の成分は微細構造に関与する。したがってケプストラムに変換することにより、両者を分離して取り扱うことができる。またケプストラムは対数スペクトルに関して線形直交変換を施したものであるから、ケプストラム領域での距離は対数スペクトル領域での距離に比例する。

発声する内容は童話の一節(10秒程度)で、発声速度、ピッチ、発声レベルなどの条件については、発声者に対して特に指示を与えないこととする。測定は主に男性9名について数日~3か月の時間間隔をおき反復して行なう。

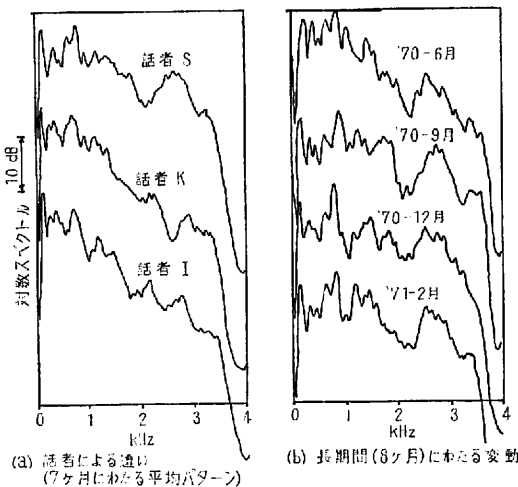


図2 長時間平均スペクトルの測定例
Fig. 2—Examples of the longtime averaged spectra.

2.2 測定結果

図2に長時間平均スペクトル(対数)の測定例を示す。(a)はS, K, Iの3名の話者のそれぞれについて、ほぼ半年にわたるサンプルの平均パターンを示したものであり、(b)は話者Sについて、長期間にわたるパターンの変動を示したものである⁽⁹⁾。

3. 話者認識

3.1 認識モデル

長時間平均スペクトルのケプストラムに関して、その変動を考慮した認識モデルを構成し、話者識別および話者照合を行なう。ここで話者識別とは、未知音声があらかじめ定められた m 人の登録話者のうちのいずれの話者によるものであるかを判定する(クラス数 = m) ものであり、話者照合とは、未知音声がある登録話者Aのものであるか否かを判定する(クラス数 = 2) ものである。判定はつぎのような種々の距離にもとづいて行なう。

(i) 既知サンプルの平均値との重みつき距離

$$S(C, \{D_{r,j}\}) = |V_r|^{-N} (C - \bar{D}_r) V_r^{-1} (C - \bar{D}_r)^t \quad (1)$$

ここで C : 未知サンプルのケプストラム (N 次元ベクトル)

$D_{r,j}$: 登録話者 r の j 番目の既知サンプルのケプストラム

$\bar{D}_r = \frac{1}{n} \sum_{j=1}^n D_{r,j}$: 既知サンプルの平均値

V_r : 登録話者 r の共分散行列 ($N \times N$ 行列)

n : 各登録話者の既知サンプルの数

この距離は、各登録話者に属するサンプルの分布を多次元空間における球に近づける線形変換を行なった後の空間でのユークリッド距離(の二乗)に相当している。また、この距離は、特徴ベクトルが多次元正規分布に従うことを仮定し、その共分散行列の各成分の値が比例的に変化できる条件のもとでの最大尤度によって認識を行なうことに相当する。

ここで式(1)が意味をもつためには V_r が正定値であることが必要であり、このためには特徴ベクトルの次元 N と既知サンプルの数 n に関して、 $N \leq n$ という条件が満たされなければならない。しかし実際には N がかなり大きい上、あまり多くの既知サンプルを得ることができないため、この条件を満たすのは困難になる。かりにこの条件が満足されても、既知サンプルの数が十分大きくないときには大きな誤差を含む形

で求まることが予想される。この対策として V_r のかわりにつぎのような行列を用いる。

[方法 I]

$$V_{1r} = \begin{bmatrix} v_{r11} & & 0 \\ & v_{r22} & \\ 0 & & v_{rNN} \end{bmatrix} \quad (2)$$

ここで v_{rji} は共分散行列 V_r の対角要素であり、 V_r が V_{1r} で置きかえられるためには、ケプストラムの各成分の間の相関が十分に小さいことが必要である。

[方法 II]

$$V_{II} = \frac{1}{m} \sum_{r=1}^m V_{1r} \quad (3)$$

m : 登録話者の数

[方法 III]

$$V_{III} = \frac{1}{m} \sum_{r=1}^m V_r \quad (4)$$

V_r が V_{II} や V_{III} で置きかえられるためには、各話者の共分散行列が十分に類似していることが必要である。

(ii) Nearest-Neighbor 法 [方法 IV]

個々の既知サンプルとの距離を求め、その最小値を用いる。

$$S(C, \{D_{rj}\}) = \min_j \|C - D_{rj}\| \quad (5)$$

この方法の特徴は、話者ごとの分布の違いを考慮できる点と、凸領域に位置していなくても取り扱える点にある。

(iii) 既知サンプルの平均値とのユークリッド距離の二乗 [重みなし]

$$S(C, \{D_{rj}\}) = \|C - \bar{D}_r\|^2 \quad (6)$$

話者識別の場合、未知サンプルからこれらの距離に関して最も近い標準パターン(既知サンプル)の話者をもって、その未知サンプルの話者と判定し、話者照合の場合は、標準パターンからの距離があるしきい値よりも小さいサンプルはその話者によるものとし、それ以外のサンプルはほかの話者によるものであるとして棄却するという方式を用いる。なおこのしきい値は、便宜上2種類の誤り率(受け入れるべきサンプルを棄却してしまう割合)と、棄却すべきサンプルを受け入れてしまう割合)が等しくなる値に事後的に設定し、このときの誤り率を認識誤り率とする。

3.2 既知サンプルと未知サンプルの設定

既知サンプルの数と期間、既知サンプルから未知サンプルまでの時間間隔などが認識率に及ぼす効果を調べるために、既知サンプルと未知サンプルを図3に示

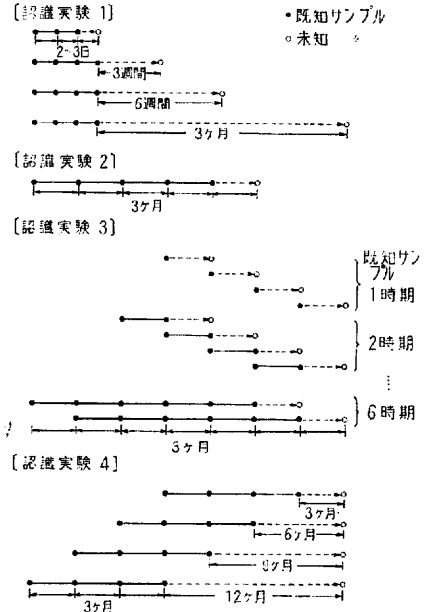


図3 既知サンプルと未知サンプルの設定
Fig. 3—Relations between reference samples and unknown ones in four recognition experiments.

すように設定して4種類の認識実験を行なう。

話者識別の場合の登録話者の数は9名とし、9名のうちから1名を選ぶ形をとる。各登録話者の既知サンプルとして各時期ごとに3回繰り返して発声したサンプルを用い、未知サンプルとして9名が3回ずつ発声した27サンプルを用いる。話者照合の場合は9名の登録話者それぞれについて、受け入れるべきサンプルとしてその話者が発声した3サンプルを用い、棄却すべきサンプルとして、その話者と異なる話者(詐称者)8名(または26名)がそれぞれ3回ずつ発声した24(または78)サンプルを用いる。

2.1で述べた方法およびパワースペクトルの偶関数性から明らかなように、求められるケプストラムの有効な次元は129次元になるが、その高次の成分はスペクトルパターンの微細構造に関係し、認識の上での重要度は次数が高くなるにつれて低くなる。そこで認識に用いるケプストラムの次元(N)を、予備実験の結果などをもとに、話者識別の場合は64(方法IIIを用いるときは31)、話者照合の場合は41(方法IIIを用いるときは21)に設定し高次の成分は無視する。

4. 認識結果

4.1 認識実験1

図4は既知サンプルから未知サンプルまでの時間間

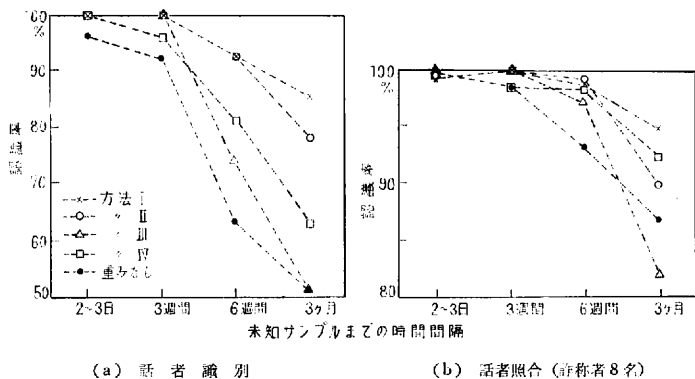


図4 認識実験1の結果
Fig. 4—Identification and verification accuracy of 1st recognition experiment.

隔を変えたときの、9名の話者に関する平均認識率を示したものである。この実験結果について各話者ごとの認識率を用いて、話者、認識方法I～IV、既知サンプルから未知サンプルまでの時間間隔、およびこれらの交互作用に関する分散分析を行なった。その結果話者識別と話者照合に共通して、時間間隔の効果および時間間隔と話者の交互作用の効果が危険率0.5%で有意差あり、話者および認識方法の効果が危険率1%で有意差ありと認められた。

これらの結果から、既知サンプルが10日間程度の短期間のものである場合、既知サンプルから未知サンプルまでの時間間隔が2～3日あるいは3週間程度に短いときには認識率が高いが、時間間隔が長くなるに従い単調に認識率が低下してしまい、3か月程度になるときわめて低い認識率しか得られないことがわかる。また分散分析の結果から、話者によって長時間平均スペクトルパターンの安定性に差のあることがわかる。

4.2 認識実験2

表1(a)に話者識別および話者照合の結果を示す⁽¹¹⁾。未知サンプルとして用いているサンプルは、認識実験1で時間間隔を3か月とした場合のサンプルと同じである。長時間にわたったデータを既知サンプルとして

表1 認識実験2の結果

	方法I	方法II	方法III	方法IV	重みなし
(a) 話者識別	74.1%	88.9%	96.3%	85.2%	55.6%
話者照合	92.6	93.5	98.2	96.8	92.1
(b) 話者識別	100%	92.6%	100%	(未知サンプルを含めて V_{I_1} , V_{II} , V_{III} を求めた場合)	
話者照合	99.5	96.3	100		

用いて、その変動を考慮した重みつき距離を設定することにより、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、よい認識率を達成することができることがわかる。

表1(b)は未知サンプルも含めて求めた V_{I_1} , V_{II} , V_{III} (\bar{D}_r の算出には未知サンプルを含めない)を用いて認識を行なった場合の結果である。理想的な重みづけが行なえれば100%程度の認識率を達成することが可能であること、す

なわち長時間平均スペクトルパターンには十分な個人性が含まれていることを示している。

距離尺度の設定の方法としては、表1(a), (b)の結果では方法IIIによる場合が比較的良好な認識率を示している。

4.3 認識実験3

図3に示したように、既知サンプルの数に関して同じ条件の実験を、異なる4時期の未知サンプルに対して行ない、その平均値として認識率を求めた。図5は9名の話者に関する平均認識率を、種々の距離尺度の場合について示したものである。話者識別の結果を(a)に、詐称者が8名の場合の話者照合の結果を(b)に、26名の場合の結果を(c)に示す^{(11), (12)}。

4.3.1 既知サンプルの数の効果について 図5の結果から話者識別と話者照合に共通して、距離尺度の設定法にかかわらず、既知サンプルが1時期のみのときには、実験1の3か月後の未知サンプルに関する結果と同程度の認識率しか得られず、よい認識を行なうためには、ある程度長期間にわたったサンプルが必要であることがわかる。

4.3.2 距離尺度の設定法について 方法I～IVによる認識率を比較すると、話者識別の場合は方法IIまたはIIIが比較的稳定した結果を示し、そのうちでは方法IIのほうがややよい結果を示している。話者照合の場合は詐称者が8名のときの種々の方法による差は小さいが、詐称者が26名にふえるとその差がやや大きくなり、方法IIIによる場合が比較的稳定したよい結果を示している。

方法IよりもIIあるいはIIIのほうがよい結果を示すのは、各話者ごとの共分散行列の相違がそれぞれの推定誤差に比較して小さいためと考えられる。

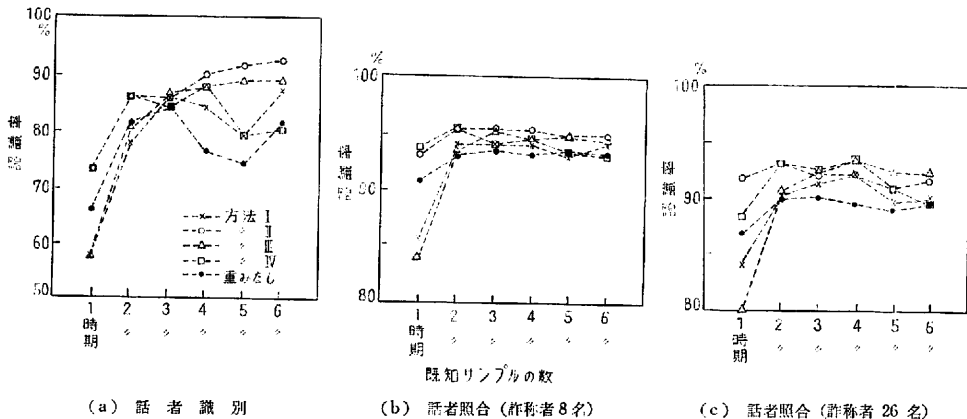


図 5 認識実験 3 の結果
Fig. 5—Identification and verification accuracy of 3rd recognition experiment

4.3.3 詐称者の数について 話者照合の詐称者の数を 8 名から 26 名にふやした場合に、1~3 期の認識率の低下が見られるが、4 時期の既知サンプルを用いて、方法 II および III によって行なった話者照合の結果について統計的検定を行なったところ、両者の間には差がないという仮説は、有意水準 5% で棄却するまでには至らなかった。詐称者の数の認識率に及ぼす影響については、今後さらに検討する必要がある。

4.3.4 話者による相違について

異なる 4 時期の未知サンプルのおのおのに対して、既知サンプルとして 3 か月おきに求めた 4 時期のデータを用い、距離尺度として方法 III を使って行なった認識実験の結果を、話者別に集計して図 6 に示した。この結果からわかるように、話者による誤り率（認識率）

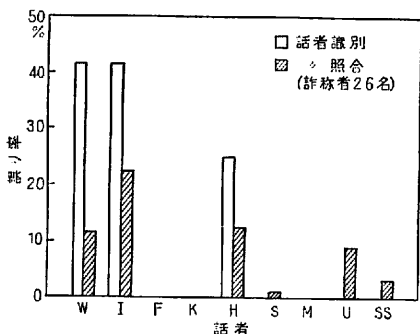


図 6 話者別平均認識誤り率
Fig. 6—Averaged error rate for individual talkers.

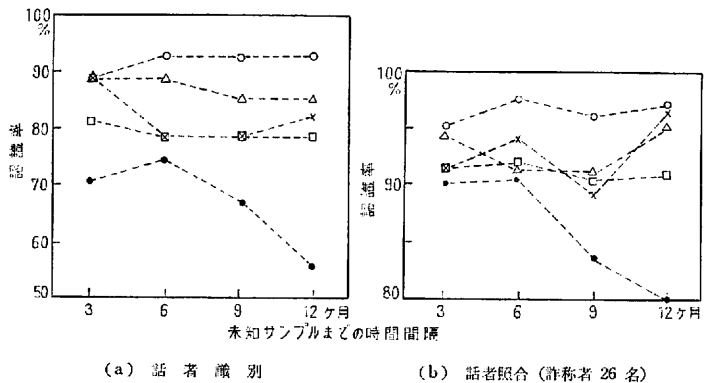


図 7 認識実験 4 の結果
Fig. 7—Identification and verification accuracy of 4th recognition experiment.

の相違は著しく、その分布の様子は話者識別と話者照合とできわめて類似している。

4.4 認識実験 4

結果を図 7 に示す。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 か月から 12 か月まで長くなった場合、重みつき距離を用いなかったときには時間間隔の増加に伴う認識率の低下が見られるが、方法 I~IV を用いた場合についてはそのような傾向は見られない。したがって認識実験 3 で得られた結果のうち、4 時期程度以上の長期間の既知サンプルによる重みつき距離を用いた場合の結果は、未知サンプルまでの時間間隔を 3 か月からさらに 12 か月程度まで延長した場合についても通用するものと考えられる。

5. 考察と検討

5.1 音声サンプルの長さ

話者の特徴パラメータとして長時間平均スペクトル

を用いる利点の一つは、発声内容にあまり影響されない点にある。しかし、そのように音韻系列の影響が無視できるためにはどの程度の長さがあればよいかという点は必ずしも明らかではない。これまでに用いているサンプルは 10 秒長程度の連続音声であるが、これが必要十分な長さであるかどうかを調べるため、つぎのような補足的実験を行なった。

既知サンプルにはこれまでと同じものを用い、未知サンプルとして異なる内容の 30 秒長程度と 3 秒長程度の 2 種の連続音声を用いて、既知サンプルと同じ内容の連続音声を用いた場合の結果と比較した。3 か月おきに 4 時期にわたって求めた既知サンプルを用いて、その最終のサンプルから 3 か月後のサンプルを認識する場合の結果を図 8 に示す。

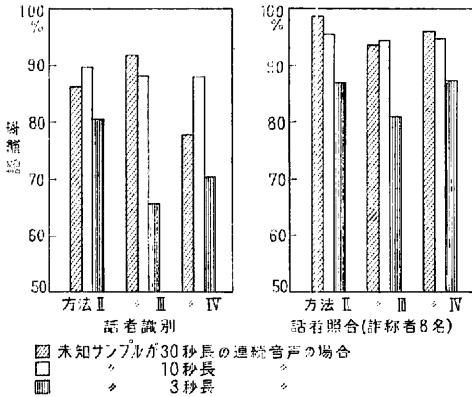


図 8 既知サンプルと未知サンプルの発声内容が異なる場合と等しい場合の比較
Fig. 8—Effects of the speech materials of unknown talker on the identification and verification accuracy.

30 秒長のサンプルの場合は、10 秒長のサンプルすなわち既知サンプルと同じ文章の場合と、ほぼ等しい認識率が得られているが、3 秒長のサンプルの場合は、それに比較してきわめて低い値しか得られず、10 秒以上のサンプルによるパターンに対して、3 秒程度のサンプルによるパターンが大きく異なっていることを示している。上の結果から、長時間平均スペクトルパターンを求めるサンプルとして必要な長さは、ほぼ 10 秒程度であると考えられ、本論文で用いているサンプルは、発声内容が無視できるという条件をほぼ満たしていると考えられる。

5.2 長時間平均スペクトルパターンの変動

5.2.1 ケプストラムの各成分の分散

長時間平均スペクトルのケプストラムにおける各成分の時期的変

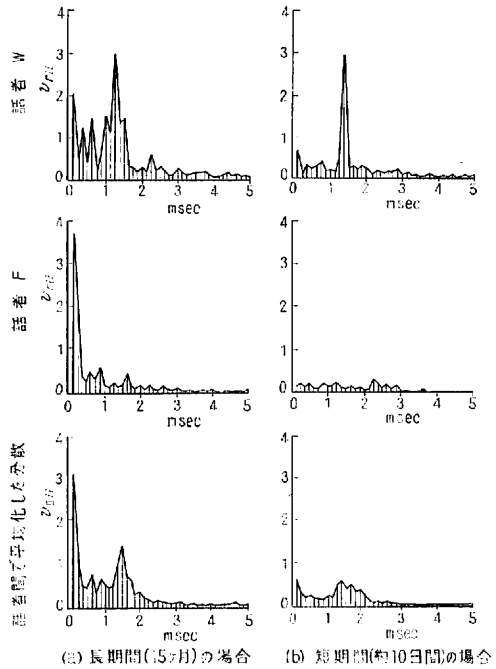


図 9 長時間平均スペクトルのケプストラムの分散
Fig. 9—Variance of the cepstrum of the longtime averaged spectrum.

動特性を調べるため、その短時間および長期間にわたる分散を求めた。短期間の変動を調べるサンプルとしては、2~3 日おきに 4 回測定したスペクトルデータを用い、長期間の変動を調べるサンプルとしては、3 か月おきに 6 時期にわたって測定したデータを用いた。図 9 は、2 名の話者のおのおの分散のパターンと、9 名の話者について平均化したパターンを示したものである。

この検討によって、たとえば、話者 W にみるように、短期間と長期間の分散においてほぼ等しい次数の付近に特異なピークを生ずる話者の多いことがわかった。また、すべての話者に共通して、長期間になるに従い低次の成分の分散が大きくなる傾向がみとめられた。これは、長期間になるに従い、長時間平均スペクトルパターンの全体的な形の変動が、比較的細かい構造の変動に比べて大きくなることを示している。

5.2.2 パターンの総合的な変動

長時間平均スペクトルパターンの総合的な変動の様子を明らかにするために、各サンプルをケプストラムに関する N 次元空間 ($N=64$) 内の一点と考え、ある時期のサンプルから、その後の各時期のサンプルまでの (ユークリッド) 距離の変化を調べた。比較的短期間の変動を示す

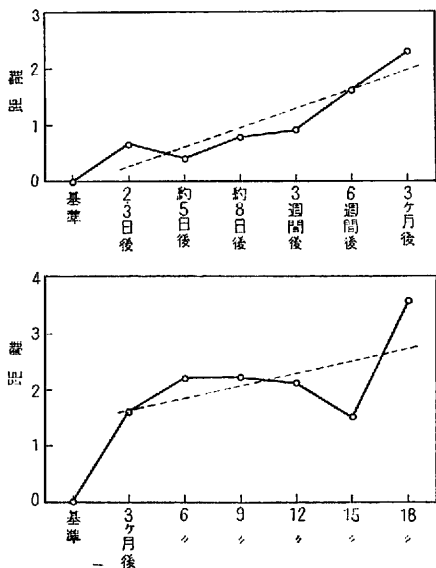


図 10 長時間平均スペクトルパターンの変動 (話者 I の場合)

Fig. 10—Temporal transition of the longtime averaged spectrum.

サンプルとして、基準の時期から 10 日後までの 3 時期および 3 週間後、6 週間後、3 か月後の各時期のサンプルを用い、比較的長期間の変動を示すサンプルとして、基準の時期を含めて 3 か月ごとに 7 時期にわたって求めたサンプルを用いた。距離の原点は基準の時期に繰り返して発声した 3 サンプルの平均値に設定し、その後の各時期の 3 サンプルの平均値までの距離を求めた。男性 9 名の各話者ごとに求めた結果のうち、話者 I の場合の結果を図 10 に示す。図中の破線は距離の変化を、基準の点を除いて最小二乗近似したもの(回帰直線)である⁽¹³⁾。

時期的変化に一方方向性があるかどうかを調べるために、上記の回帰直線の傾斜について検定を行なった。その結果、傾斜がないという仮説は、短期間のサンプルについては 9 名中 7 名、長期間のサンプルに関しては 2 名に関して有意水準 5% で棄却された。このことから、3 か月程度までの間ではパターンの一方方向的な変化がきわめて著しく、3 か月ごとの 18 か月にわたる変化に関しては、その傾向はかなり少なくなることがわかる。前者の傾向は、認識実験 1 の場合に、時間間隔が長くなる従って単調に認識率が低下した原因になっていると考えられる。また、短期間と長期間の傾向の相違は認識実験 1 と認識実験 4 の結果の差違に対応づけることができる。

4.3.4 で述べた話者別の誤り率と上の結果を比較したところ、話者識別と話者照合に共通して大きな誤りを示す 3 名の話者のうちの 2 名は、長期間にわたっても一方方向的な変化を示す話者と一致していることがわかった。本論文で用いている認識方法 I~IV では一方方向的な変化を完全に取り除くことは不可能であるから、このような話者に対しては、既知サンプルを求める周期を短くするとともに、一定期間にわたった既知サンプルを逐次更新して行く必要がある。

5.3 認識に用いるケプストラムの次元

長時間平均スペクトルのケプストラムについて、9 名の話者間で平均化した共分散行列を求めたところ、相関の強い成分は低次にほとんど集中していることがわかった。これまでに述べた認識実験では、ケプストラムの零次から 20 ないし 63 次 (7.875 msec) までを用いているが、このような高次の成分まで考慮することは必ずしも必要ではない。ここではケプストラムの各成分間の相関は低次の成分についてのみ考慮し、それ以上の高次の成分については、共分散行列の対角要素すなわち分散のみを考慮する形を採用して認識実験を行ない、認識に必要な十分な次元の数を推定した。

その結果、話者識別については 0~10 次について相関を考慮し 20 次 (2.5 msec) までについて分散を考慮すればほぼ十分で、4 時期の既知サンプルを用いて 3 か月後の未知サンプルを認識する場合の認識率は平均 91% であった。話者照合の場合は 0~10 次について相関を考慮し 30 次 (3.75 msec) までについて分散を考慮すればほぼ十分で、上と同じケースについて平均 93% (詐称者が 26 名のとき) の認識率が得られた。

5.4 話者照合の判定のしきい値

話者照合における判定のしきい値を、あらかじめ全登録話者に共通に一定の値に設定した場合について、補助的に検討した。その結果、これまでのように事後に最適値に設定する場合に比べて認識率は若干低下するが、その値は 3% 程度におさえられる見通しが得られた。

6. む す び

これまでの実験により、長時間平均スペクトルパターンは、2~3 日から 3 週間程度までの間ではきわめて安定しているが、その後、時期がたつにつれて、同じ発声者によるものでも種々に変化することが明らかになった。しかし、そのパターンの長期間にわたる変

動を考慮した適切な距離尺度を設定することにより、3か月以上の長期間をおいた後でも、話者の認識が十分に可能であることが確認された。

3か月おきに4時期にわたって測定した長時間平均スペクトルを既知サンプルに用いて、その変動を考慮した重みつき距離を設定することにより、その最終の時期の既知サンプルから3か月後の未知サンプルに関して、9名から1名を選ぶ話者識別の場合に平均91%、詐称者が8名ないし26名の話者照合の場合に平均93~95%の認識率が得られた。各登録話者に関して、既知サンプルの測定周期を1か月程度にまで短縮し、半年程度にわたったデータの集積をその周期で逐次更新してゆくような方法を採用すれば、認識率はさらに良くなるであろう。

今後、長期間にわたる長時間平均スペクトルパターンの変動特性の解明、その変動をさらに適確にとり入れた距離尺度の設定法、作り声による変化、女声によるパターンとの比較、長時間平均スペクトルパターンを決定している要因の追求をはじめ、その他の種々の問題点について検討を進めて行きたい。

謝辞 日ごろ、ご指導いただく当研究所鵜淵基礎研究部長、ならびに研究室の諸氏に感謝の意を表す。

文 献

- (1) W.A. Hargreaves and J.A. Starkweather: "Recognition of speaker identity", *Language & Speech*, **6**, p. 63 (1963).
- (2) P.D. Bricker, R. Gnanadesikan, M.V. Mathews, S. Pruzansky, P.A. Tukey, K.W. Wachter and

- J.L. Warner: "Statistical techniques for talker identification", *Bell Syst. tech. J.*, **50**, 4, p. 1427 (1971).
- (3) S.K. Das and W.S. Mohn: "A scheme for speech processing in automatic speaker verification", *IEEE Trans. AU-19*, 1, p. 32 (1971).
- (4) J.E. Luck: "Automatic speaker verification using cepstral measurements", *J. Acoust. Soc. Amer.*, **46**, 4, 2, p. 1026 (1969).
- (5) 落合, 福村: "音声通信における基本品質について", *研実報*, **8**, 12, p. 1469 (昭 34).
- (6) H. Mangold: "Die individuellen unterschiede bei langzeitmessungen am sprachsignal", *NTZ*, **6**, p. 364 (1969).
- (7) G. Ramishvili: "Some problems in automatic identification of spoken utterances and speakers", *STL-QPSR*, **2**, 3, p. 36 (1970).
- (8) B.P. Bogert, M.J.R. Healy and J.W. Tukey: "The quefrency analysis of time series for echoes", *Proc. of the Symposium on Time Series Analysis*. John Wiley & Sons, New York, p. 209 (1963).
- (9) 古井, 板倉, 齋藤: "長時間平均スペクトルによる話者照合", *音響学会講演論文集*, 2-1-2 (昭 46-05).
- (10) 古井, 板倉, 齋藤: "長時間平均スペクトルの変動を考慮した話者認識", *音響学会講演論文集*, 3-1-18 (昭 46-11).
- (11) 古井, 板倉, 齋藤: "音声の長時間平均スペクトルによる話者認識", *音響学会音声研究会資料* (昭 47-01).
- (12) 古井, 板倉, 齋藤: "長時間平均スペクトルの経時変化を考慮した話者認識", *昭 47 信学全大*, 56.
- (13) 古井, 板倉: "長時間平均スペクトルの経時変化—話者認識におよぼす効果—", *音響学会講演論文集*, 3-2-5 (昭 47-05).

(昭和 47 年 4 月 1 日受付)