

論文 / 著書情報  
Article / Book Information

論題(和文)	話者認識
Title	
著者(和文)	古井貞熙
Author	SADAOKI FURUI
出典(和文)	日本音響学会誌37巻5号, Vol. 37, No. 5, pp. 232-238
Journal/Book name	, Vol. 37, No. 5, pp. 232-238
発行日 / Issue date	1981,

# 話 者 認 識\*

古井貞熙\*\*

(日本電信電話公社武蔵野電気通信研究所)

## 1. 話者認識とは

音声波の中には、伝えようとしている言葉の意味内容に関連した音韻性情報、誰が話しているかという個人性情報のほか、発声者の喜怒哀楽に関する情報など、種々の情報が含まれている。このうち、音韻性情報を自動的に抽出し認識するのが、いわゆる、音声認識 (speech recognition) であり、個人性情報を抽出して認識するのが話者認識 (speaker recognition) である<sup>1), 2)</sup>。広義にはこの両者を合わせて音声認識ということもある。話者認識の研究を歴史的に見ると、人間が聴覚によって話者を認識するときに、音声波あるいはそのスペクトルのどのような手がかりを用いているかという研究<sup>3)</sup>や、スペクトルの時間的変化を図形に表したスペクトログラム——いわゆる声紋 (voice print) ——を人間が目で見ても話者を認識する可能性に関する研究<sup>4)</sup>も、1950年代から行われている。しかし近年、電子計算機や電子回路の発達とともに、自動的な話者認識の研究が急速に発展して来ており、その能力は、一部人間の能力を凌駕するほどにまで進んでいる。

話者の自動 (機械) 認識の研究は、(狭義の) 音声認識の研究に比べて歴史が浅く、一般の関心もまだあまり高くないが、犯罪の捜査のみならず、情報化社会、データ通信サービス等の発達に伴って、需要が増して来るものと考えられる。ここでは、以後話者認識という言葉で、このような自動的な認識に限定して用いることとする。

## 2. 話者認識の分類

### 2.1 話者識別と話者照合

話者認識の形態は、話者識別 (speaker identification) と話者照合 (speaker verification) に分けることができる<sup>5)-7)</sup>。図-1に示すように、話者識別とは、未知音声<sup>(Unknown)</sup>を、あらかじめ登録してある種々の話者の音声 (標準パターン) と比較して、最も類似している標準パターンを選び、その登録話者が発声した音声と判定するもので、

複数の候補カテゴリーのうちの一つと判定するという意味では、単語音声認識等の考え方と共通している。一方話者照合とは、未知音声<sup>(Unknown)</sup>が、その発声者が名乗った本人のものであるか否かを、名乗った本人の標準パターンと未知音声との類似の度合いによって判定するものである。話者識別では、正しい話者が選ばれる割合 (認識率) は、登録されている話者の数に依存し、話者の数が大きくなるに従って単調に低下するが、話者照合では本人の音声として受理するか棄却するか<sup>(Accept or Reject)</sup>の二値の判定が行われるので、登録話者数と未知入力音声の話者数がある程度以上に大きければ、認識率は登録話者数には依存しない<sup>8), 9)</sup>。

話者照合の例としては、電話による買物、電話による送金 (振替)・残高照会・振込通知等のバンキングサービス、個人の秘密に関する情報へのアクセス、重要機密の保管場所への入室者のチェック等において、本人か否かの鍵として音声を用いることが考えられる。話者識別の例としては、犯罪捜査のようなときに、犯行時に録音された音声<sup>(Stored)</sup>が、複数の容疑者の内の誰の声であるかを判定するような場合等がある。ただしこの場合は、容疑者の中に真犯人が含まれていないこともありうるので、実際には話者照合と話者識別の組み合わせの形態をとることが多いと考えられる。音声認識や音声応答の実用化によって、今後電話機を通じた種々のサービスが実現されてくると、それに伴って上述のように、音声によって個人を識別する必要性が増して来ると考えられる。

### 2.2 発声内容依存形と不依存形

話者認識の方法には、認識に用いる音声の発声内容、すなわち言葉をあらかじめ決めておく方法 (発声内容依

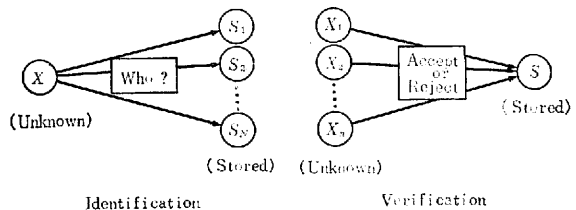


図-1 話者識別 (speaker identification) と話者照合 (speaker verification) の原理

\* Speaker recognition

\*\* Sadaoki Furui (Musashino Electrical Communication Laboratory, NTT, Musashino, 180)

存形, text-dependent の方法)と、どんな言葉を発声してもよいとする方法(発声内容不依存形, text-independent の方法)がある。音声波に含まれる音韻性情報と個人性情報を明確に分離する技術はまだ確立されていない上、個人性情報は音声波に第一義的にふくまれる音韻性情報を完全に乱すほど大きくはなり得ないという意味で、一ランク細かい情報であるため、発声内容不依存形の方法で、高い認識率を得るのはかなり難しい。音声スペクトルに含まれる音韻性情報と個人性情報の関連を統計的に調べた結果によれば、<sup>10),11),1)</sup> 両者には有意な交互作用があることが分かっており、発声内容依存形の方法の場合はこの交互作用の情報も用いることができるので、使える情報がそれだけ豊富になる。

発声内容依存形の方法の中には、ある特定の音素(例えば鼻音/n/)のスペクトルに含まれる情報を用いる方法のように、すべての話者がある特殊な決まった言葉を発声しなければならないという場合もあるが、現在研究されている方法は、各話者がそれぞれに決めたキーワードを一貫して発声すればよい、というものがほとんどである。この場合は、実用的にはそのキーワード(自分の名前等を用いてもよい)の違いも手がかりとして用いることができるので、それだけ高い認識率を得るのが容易となる。研究段階の認識実験では、すべての話者に共通のキーワードを用いて評価を行う場合が多い。

話者照合の応用の多くの場合は、キーワードをそれぞれの人について固定することが可能であるが、犯罪捜査の場合等では、必ずしも同じ言葉どうしを比較することができない場合も生ずる。このような目的のためには、発声内容不依存性の方法を開発する必要がある。このような技術が確立すれば、音声認識においてシステムを発声者の声に適応させる際に、あるきまった学習サンプルを用いずに、任意の音声をもとにしてその発声者の声の特徴をとらえ、その発声者に合った標準パターンを読み出して用いたりすることができるようになる。

### 3. 話者認識に用いる特徴

話者認識においても音声認識の場合と同様に、音声波からスペクトルあるいはそれを表現する特徴パラメータの形で情報を抽出し、それを用いて認識を行う。音声波に含まれる発声者個人に関する情報は、声道長、声帯等の先天的な発声器官の個人差に起因するものと、なま

り、アクセント等の後天的な発話習慣(話し方のくせ)に起因するものがあるが、音声波から両者を明確に分離して抽出するのは難しい。このため、両者の特徴が同時に含まれる特徴が用いられることが多い。

実際の話者認識の方法としては、スペクトル包絡、基本周波数、パワー等の時間パターンをそのまま用いる場合と、その時間パターンから抽出した、時間軸を有しない統計的特徴を用いる場合がある。発声内容依存形の話者認識の場合は前者の方法を、発声内容不依存形の場合は後者の方法を用いることが多いが、後述のようにその逆の例もある。スペクトル包絡の時間パターンを直接的に用いる方法は、単語音声認識でしばしば用いられる方法と類似しているが、次章で述べるように、いくつかの話者認識に特有の配慮が必要である。

基本周波数やパワーの時間パターンは発話習慣に直接的に関連しており、声帯模写では一般にこのような特徴を真似していると考えられる。自動的な話者認識においても、このような特徴を基礎とする方法は作り声に対して弱く、スペクトル包絡の特徴を組合せれば、作り声に対して強くすることができることが示されている<sup>12)</sup>。

## 4. 話者認識系の構成と問題点

### 4.1 認識系の構成

話者認識系の基本的構成は図-2に示す通りで、音声波から特徴抽出をしたのち、あらかじめ登録されている各登録話者の標準パターンとの比較を行って、その類似の度合いにより、認識の判定が行われる。通常、類似度のかわりに、値が小さくなるほど類似性の度合いが大きいことを示す“距離尺度”を用いることが多い。

話者識別の場合は、入力音声との距離が最も小さい標準パターンの話者を選択し、話者照合の場合は、入力音声と標準パターンの距離と、あらかじめ定めたしきい値との大小関係によって判定を行う。このしきい値と二種類の誤り率、すなわち本人の音声を棄却してしまう誤り率と、他人の音声を受理してしまう誤り率との間には、図-3に示すような関係がある。実用的には、二種類の誤り率の重要性に応じて、しきい値を適切な値に設定するのが望ましい。認識実験の場合は、二種類の誤り率が等しくなるような値(図中のP)にしきい値を事後的に設定し、このときの誤り率によって評価を行う場合が多い。

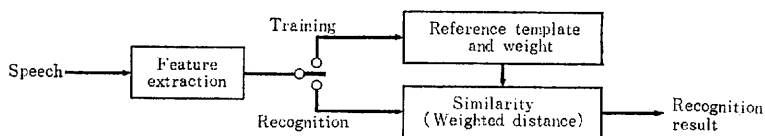


図-2 話者認識系の基本的構成

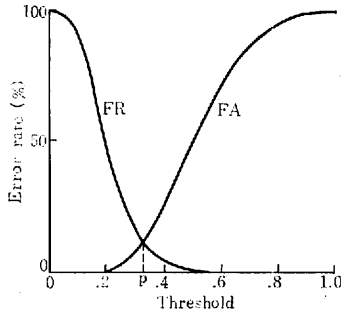


図-3 話者照合における判定のしきい値と二種類の誤り率の関係。  
FR: 本人の音声を受却してしまう誤り  
FA: 他人の音声を受理してしまう誤り

#### 4.2 特徴パラメータの長期的変動

話者認識系を構成するときの重要な問題点として、特徴パラメータの長期間にわたる話者内変動の影響がある。すなわち、同じ人が同じ言葉を繰返し発声しても、ある程度時期をおくと、その音声のスペクトルが変動し、それが認識系の性能に影響を及ぼす。これが音声認識の場合にはあまり問題にならなかった<sup>13)</sup> 背景には、音声認識と話者認識の二つの相違点がある。その一つは、話者認識を実際に行う場合には、各話者の標準パターンを作るための学習サンプルと未知入力音声が同時期に発声されることはまれで、この両者の間には通常時期差があるということである。二番目の理由は、2.2節に述べたように、スペクトルパターンに含まれる個人性情報は、音韻性情報よりも一ランク細かい特徴であるために時期差による変動の影響を受けやすいということである。

このように、本質的に時期差によって変化する性質を有する点が、音声と指紋の相違点であり、音声のこの性質は、聴覚による話者認識にも影響を及ぼすことが、実験的に認められている。特徴パラメータの時期的変動に対する対策としては、次のような方法をとるのが望ましい<sup>14)</sup>。

- ①長期間にわたって比較的安定な特徴パラメータを用いる。(スペクトル等化処理—後述)
- ②各話者の標準パターンと距離尺度を定めるために、できるだけ長期間にわたる学習サンプルを用いる。
- ③標準パターンを適切な周期で更新する。

#### 4.3 特徴パラメータの評価

特徴パラメータの有効性を評価したり、有効なパラメータを選択するには、直接的にそれらのパラメータを用いた認識実験を行うほかに、パラメータごとの $F$ 比と呼ばれる話者間/話者内分散比を用いる方法<sup>15)</sup>、それを多次元に拡張した divergence を用いる方法<sup>16)</sup>、認識誤り率

による knock out 法<sup>17)</sup>などが提案されている。判別分析を行って、 $F$ 比を最大にする空間へ射影する方法は、情報圧縮の手段として、しばしば用いられる。

### 5. 最近の話者認識研究の動向

話者認識システムはまだ実用化されるには至っていないが、最近の5年程度の間には技術的に大きな進歩が見られ、米国 Texas Instruments 社における、男女計300名の話者による2年半以上にわたる入室管理システムの試行実験<sup>16)</sup>、Bell 研究所における男女計100名の電話音声による5か月にわたる話者照合実験<sup>18)</sup>など、大規模な実験も行われるようになって来た。基礎的な研究も国内外で活発に行われている。

最近の主な研究例を表-1にまとめて示した。このうちのいくつかの例については、後に簡単に説明するが、詳しく知りたい読者のために、研究者名の欄に文献番号を示した。認識率の欄には、報告されている値をそのまま載せたが、表に示されている各条件のほか、学習サンプルの収集期間、入力音声までの時期差等の条件も、それぞれの実験によって異なっているので、この数値だけで単純にその方式の優劣を比較することはできない。一般的には、発声内容不依存形のシステムはまだ基礎研究段階にあるが、発声内容依存形の話者照合は、実用化に近づいていると言えよう。実用化のために今後更に検討しなければならない課題については、7章で述べる。

### 6. 話者認識系の具体例

#### 6.1 統計的特徴を用いる発声内容依存形の認識系

##### 6.1.1 認識方法

発声内容依存形、すなわち発声する言葉をあらかじめ決めておく認識系の一つとして、スペクトルパラメータの統計的特徴を用いる方法を紹介する<sup>20)</sup>。この方法では、各話者があらかじめ決められた複数個の単語を区切って発声し、それぞれの単語から統計的特徴を抽出して認識に用いる。

図-4に示すように、音声波から10ms間隔で相関係数を抽出し、各区間ごとに有声/無声の判定を行って、有声音区間のみの時系列に変換する。1次の相関係数を、単語の全有声音区間にわたって平均化し、この値を用いて、一次系の逆フィルタを構成する。伝送歪や時期差によるスペクトルの話者内変動に対処するために、もとの相関係数に畳み込みを行って逆フィルタ処理(スペクトル等化処理)を行い、こののちに基本周波数と10次までのLPCケプストラムからなる時系列に変換する。この時系列から、各パラメータの全有声音区間における平均値、標準偏差、低次のフーリエ展開係数、およびパラメータ相互間の相関係数を計算する。これらの統計量のうち、話者の分離に有効な成分を選択し、認識のための特

表-1 最近の話者認識の主な研究例

研究機関	研究者 (年代)	発声内容 依存(TD) 不依存(TID)	発声内容	特徴パラメータ・方法	識別(I) 照合(V)	登録話者数 ( ): 照合 の詐称者数	認識率 (%)
電電公社 武蔵野通研	古井・板倉・ 森藤 <sup>19)</sup> (1972)	{ TD TID	短文章 (10~30秒)	長時間平均スペクトルのケプストラムの重みつき距離, 時期差の効果的分析	I	男 9	90.8
					V	男 9(26)	93.5
	古井・板倉 <sup>20)</sup> (1973)	TD	1~4単語	対数断面積比と基本周波数の統計的特徴(平均値・標準偏差・相互相関係数)の重みつき距離	I	男 9	96.8~99.1
					V	男 9(111)	97.3~99.3
	古井 <sup>21)</sup> (1974)	TD	1~2単語	上の認識系においてスペクトル等化処理(逆フィルタ)による時期差の影響の除去, 短期間学習	I	男 9	92.3~99.4
					V	男 9(8)	97.2~98.9
	古井 <sup>22)</sup> (1980)	TD	4単語	LPC ケプストラムと基本周波数の統計的特徴(平均値・標準偏差・相互相関係数・フーリエ展開係数)の重みつき距離	I	男 9	100
					V	男 9(111)	99.9
東北大	松本・曾根・ 二村 <sup>23)</sup> (1977)	TID	短文章 (0.5秒~)	ケプストラムと基本周波数の断片的 正準判別分析	I	男10	96
日立中研	市川・中島・ 中田 <sup>24)</sup> (1979)	TD	1単語	Cosh 尺度を用いた DP マッチングと平均基本周波数, 帯域選択自己平均逆フィルタ法による伝送歪の除去, 電話音声	I	男 5	91
Bell Labora- tories (米国)	Rosenberg <sup>25)</sup> Sambur (1975)	TD	短文章 (2秒)	基本周波数, パワー, 線形予測係数(4次, 8次)の DP マッチングの後, 種々の距離尺度を計算	V	男22 (55)	99 96(作り声)
	Posenberg <sup>18)</sup> (1976)	TD	同上	上の認識系において線形予測係数を除き, 電話音声に適用	V	男女100 (同)	適応後 FR 4% FA 5%
	Furui <sup>26)</sup> Rosenberg (1980)	TD	同上	LPC ケプストラムとその多項式展開係数の DP マッチング, スペクトル等化処理	V	男10(49) 女10(49) 男21(75)	99.8 } 電話 99.6 } 音声 99.2
	Atal <sup>28)</sup> (1974)	{ TD TID	短文章 (0.5秒~)	種々の線形予測パラメータを直交線形変換し有効性を評価, LPC ケプストラムが最も有効	V	男10(9)	98(1秒 TD)
					I	男10	98(0.5秒 TD) 93(2秒 TID)
	Sambur <sup>27)</sup> (1976)	{ TD TID	短文章 (2秒)	種々の線形予測パラメータを直交線形変換し, 固有値の小さい成分を使用, 音韻性と個人性を分離	V	男21(20)	99(TD)
I					男21	99(TD) 94(TID)	
Texas Instruments (米国)	Doddington <sup>29)</sup> (1976)	TD	CVC 形 1~4単語	母音中央部(100ms)のスペクトルマッチング, スペクトル等化処理, 計算機室の入室管理に適用	V	男180 (同)	4単語 FR 0.3% FA 1%
SCRL (米国)	Markel <sup>29)</sup> Davis (1979)	TID	会話音声 (39秒~)	PARCOR 係数と基本周波数の統計量(平均値・標準偏差)の重みつき距離	V	男11 女6 (同)	96
					I	男11 女6	98
Philips 研 究所 (西独)	Bunge <sup>30)</sup> (1977)	{ TD TID	短文章 (11秒)	長時間平均スペクトルに対して種々の距離尺度を適用	{ V I	男41 女9 (同)	98~99

SCRL: Speech Communications Research Laboratory FR: 本人拒否率 FA: 詐称者受率

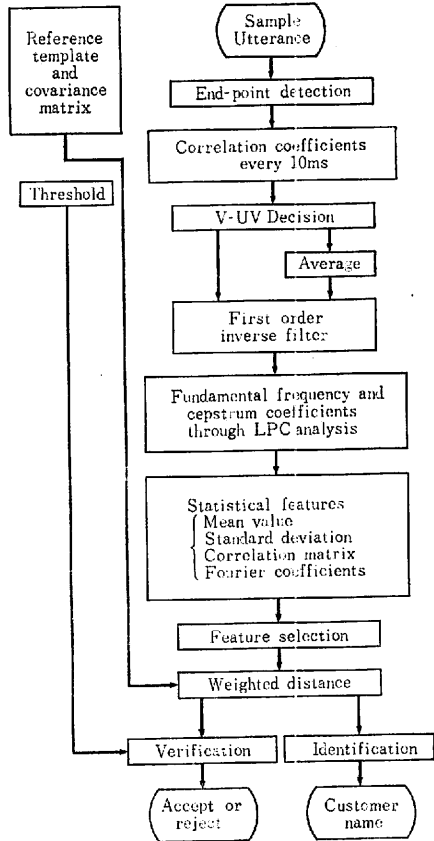


図-4 単語音声スペクトルの統計的特徴を用いた話者認識系のブロック図

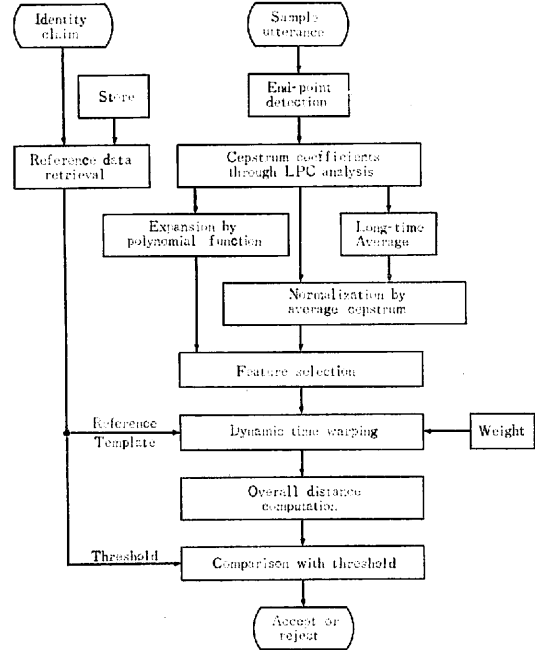


図-5 短文章音声のケプストラムと多項式展開係数の時系列を用いた話者照合系のブロック図

徴ベクトルとする。選択する成分は、多数話者の学習サンプルを用いた分散分析によって決めておく。認識の判定は、入力音声の特徴ベクトルと各話者の標準パターンとの、各成分の安定性を考慮した重みつき距離によって行う。

上述のスペクトル等化処理は、近似的に音源と声道のスペクトル包絡特性を分離し、前者を取除いて後者のみを残していると考えられることができる。このことによって、音声の長期間にわたる変動の影響を少なくし、高い認識率を保つことができることが、確かめられている<sup>21)</sup>。

6.1.2 認識実験

実験には、男性9名が3年間の長期間にわたって発声した日本語4単語の音声サンプルを用い、4単語の特徴を組合わせて認識の判定を行った。

まず、マイクロホンで収録した音声を用い、各話者ごとに、3か月おきの4時期に収録した12サンプルを学習サンプルとして、その3か月後の音声を確認する実験を、多時期の音声サンプルを用いて行った。話者照合の詐称者(本人以外の話者)の音声としては、103名の男性話者が1回ずつ発声したものと、登録話者9名のうちの本人

人以外の話者が多時期に発声したのものを用いた。実験結果は表-1の上から4番目の欄に示す通りで、極めて高い認識率が得られている。

次にこの方法を、同時に録音した電話音声とマイクロホン音声にそれぞれ適用して認識実験を行ったところ、電話音声に対しても、マイクロホン音声とほぼ同程度の性能を示すことが確かめられた。現在この方法を用いた、実際の電話音声によるオンライン実験が進められている。

6.2 スペクトル時系列を用いる発声内容依存形の認識系

6.2.1 認識方法

発声内容依存形の二番目の例として、スペクトル時系列を直接的に用いる方法の一つを紹介する<sup>25)</sup>。この方法では、単に時系列の非線形時間伸縮マッチングを行うだけでなく、スペクトルの動的な特徴を積極的に抽出して用いている。

音声サンプルとしては短文章音声を用い、図-5に示すように、10msごとに10次までのLPCケプストラムを抽出する。音声区間全体にわたって平均化したケプストラムの値を、各時点のケプストラムの値から減ずることによって、スペクトル等化処理を行う。一方、各ケプストラムの時系列を、10msごとに90msの区間を用いて直交多項式展開し、その1次と2次の係数の時系列を得る。これらの展開係数と、平均値を減じて正規化され

たケプストラムのうち、話者を分離する上で有効性の高い18種類の成分の時系列をとり出し、標準パターン時系列との非線形時間伸縮マッチングを行う。最適なマッチングが行われたときの標準パターンと入力音声の距離を、あらかじめ学習サンプルを用いて各話者ごとに設定してあるしきい値と比較して、話者照合の判定を行う。

### 6.2.2 認識実験

実験には、表-1の上から9番目の欄に示す通り、電話音声とマイクロホン音声の3種類の音声サンプルセットを用いた。各話者は約2か月にわたって、平均ほぼ1日おきの割合で同じ文章を繰返し発声し、詐称者の音声としては、登録話者のうちの本人以外が繰返し発声したものと、他の40~45名が1回ずつ発声したものをを用いた。各話者の標準パターンは、初めの5サンプルの平均化によって作成し、その後7回発声するごとに、その直前の5サンプルを用いて更新した。話者照合の判定のしきい値も同時に更新した。

実験結果は表-1に示す通りで、3種類の音声サンプルセットのいずれに対しても、99%以上の認識率が得られた。次に上記の電話音声を、24 kbit ADPCMおよびLPC ボコーダの伝送系に通し、元の電話音声を含めて、学習サンプルと入力音声異なる伝送系を通った場合についての実験を行ったが、この場合も99%以上の認識率が得られることが分かった。その後、このシステムを用いて、男女計120名の実際の電話音声によるオンラインの実験が6か月にわたって行われ、良好な結果が得られている。

### 6.2.3 統計的特徴と動的特徴の比較

発声内容依存形の認識系における、スペクトル時系列を直接的に用いる方法と、統計的特徴を用いる方法の優劣をつけるのは難しく、今後の実験的検討に待たなければならないが、認識性能が同程度であれば、一般的に前節で述べたような統計的特徴を用いる方が、記憶容量や計算量が少なくすみ、有利であると考えられる。表-1の下から3番目の欄に示した Doddington の方法<sup>29)</sup>は、母音中央部のみの時系列を用いることによって情報量を削減し、しかも多数話者の長期間にわたる音声に対して、良い認識結果を得ている。

### 6.3 統計的特徴を用いる発声内容不依存形の認識系

発声内容を限定しない認識方法の代表的なものは、長時間平均スペクトルを用いる方法である。表-1の一番上の欄に示した方法<sup>19)</sup>は、10~30秒の短文章音声の長時間平均スペクトルのケプストラムを用いるもので、ケプストラムを用いてスペクトルパターンを展開することにより、長期間にわたって比較的安定した特徴を重視することができるという特徴を有している。

表-1の下から4番目の欄に示した Sambur の方法<sup>27)</sup>は、LPC 分析パラメータを直交変換した特徴パラメー

タの平均値と分散を用いるものである。このとき、LPC 分析パラメータの共分散行列の比較的大きい固有値を有する固有ベクトルへの射影は音韻性に、比較的小さい固有値を有する固有ベクトルへの射影は個人性に関与するという大胆な仮定に立ち、後者のみを話者認識に用いている。

表-1の下から2番目の欄に示した Markel らの方法<sup>29)</sup>は、長時間の音声に対して、PARCOR 係数と基本周波数の平均値と標準偏差を抽出し、判別分析によって個人差を強調する空間に射影して認識を行うものである。

表-1の上から5番目の欄に示した松本らの方法<sup>23)</sup>は、ケプストラムの空間をいくつかの音韻に対応する部分空間に分割し、各部分空間において正準判別分析を行って個人差を強調する空間に射影して認識を行うもので、今後さらに改良を加えれば、有望な方法となる可能性がある。

## 7. 話者認識の今後の課題

話者認識の具体的な応用としては、2.1 節で述べたように種々の例を上げることができる。実用化に際しては、当面は ID カード等と組合わせて用いられることが多いと考えられるが、米国防空軍の委託を受けた Mitre 社の調査<sup>31)</sup>でも、音声は指紋や筆跡よりも誤り率、処理時間のいずれの点でも秀れていることが示されており、将来は、音声を印鑑、ID カード、クレジットカード、サイン等のかわりに用いることも考えられる。音声を用いる大きなメリットの一つは、電話で本人の確認ができるという点であろう。

このような将来のニーズに答えるために、今後研究を進めなければならない課題としては、次のようなものがある。

- (1) 長期間たっても安定で、しかも作り声などに強い特徴パラメータの開発。
- (2) 電話機、伝送路等による歪や雑音を除去する対策の開発。
- (3) 標準パターンのための記憶容量、分析・認識のための計算量の削減。
- (4) 数千あるいは数万人の長期間の音声によるテスト。
- (5) 発声内容不依存形の認識方法の開発。

音声波に含まれる個人性情報と音韻性情報の分離は、不特定話者用の音声認識システムの開発のためにも重要な課題であり、今後本格的な研究の必要性が増してくると思われる。

## 謝 辞

日頃御指導御鞭撻を頂く、当所畔柳基礎研究部長ならびに板倉第四研究室長に感謝する。日頃熱心に討論して頂く研究室の諸氏に感謝する。

## 文 献

- 1) 古井貞照, “音声波に含まれる個人性情報の研究,” 学位論文, 東京大学 (1978).
- 2) 新美康永, 音声認識, 情報科学講座 E19・3. (共立出版, 東京, 1979).
- 3) 伊藤憲三, 斎藤収三, “個人性の知覚に寄与する音響パラメータの分析,” 信学会電気音響研究, EA78-71 (1979).
- 4) O. Tosi, H. Oyer, W. Lashbrook, C. Pedrey, J. Nicol and E. Nash, “Experiment on voice identification,” J. Acoust. Soc. Am., 51, 6(pt. 2), 2030-2043 (1972).
- 5) 古井貞照, “話者認識,” 昭和55年連合大会, 33-6 (1980).
- 6) B. S. Atal, “Automatic recognition of speakers from their voices,” Proc. IEEE, 64, 4, 460-475 (1976).
- 7) A. E. Rosenberg, “Automatic speaker verification: A review,” *ibid.*, 475-487.
- 8) G. R. Doddington, “Speaker verification,” Rome Air Development Center, Tech. Rep. RADC 74-179 (1974).
- 9) 古井貞照, “話者認識における登録話者数の効果,” 昭和52年信学会情報部門企大, 228 (1977).
- 10) 坂井利之, 田畑孝一, “VCV 音節の多変量解析,” 信学会論文誌, 56-D, 1, 63-70 (1973).
- 11) 古井貞照, “音声スペクトルパターンに含まれる音韻性と個人性の解析,” 昭和51年信学会総合企大, S11-10 (1976).
- 12) A. E. Rosenberg and M. R. Sambur, “New techniques for automatic speaker verification,” IEEE Trans. Acoust., Speech, Signal Processing, ASSP-23, 2, 169-176 (1975).
- 13) 古井貞照, “話者認識と音声認識におけるスペクトル変動の影響の比較,” 音講論集, 361 (1977, 4).
- 14) B. S. Atal, “Automatic speaker recognition based on pitch contours,” J. Acoust. Soc. Am., 52, 6(pt. 2), 1687-1697 (1972).
- 15) M. R. Sambur, “Selection of acoustic features for speaker identification,” IEEE Trans. Acoust., Speech, Signal Processing, ASSP-23, 2, 176-182 (1975).
- 16) G. R. Doddington and B. M. Hydrick, “User adaptation in operational speaker verification,” J. Acoust. Soc. Am., 61, Suppl. 1, S89(A) (1977).
- 17) B. M. Hydrick and G. R. Doddington, “Performance evaluation of speaker verification in entry control,” J. Acoust. Soc. Am., 64, Suppl. 1, S182 (1978).
- 18) A. E. Rosenberg, “Evaluation of an automatic speaker verification system over telephone lines,” Bell Syst. Tech. J., 55, 6, 723-744 (1976).
- 19) 古井貞照, 板倉文忠, 斎藤収三, “長時間平均スペクトルによる話者認識,” 信学会論文誌, 55-A, 10, 549-556 (1972).
- 20) 古井貞照, 板倉文忠, “単語の統計的パラメータによる話者認識,” 信学会論文誌 56-A, 11, 717-724 (1973).
- 21) 古井貞照, “音声の個人性パラメータの時期的変動と話者認識,” 信学会論文誌 57-A, 12, 880-887 (1974).
- 22) 古井貞照, “統計的特徴と動的特徴による話者認識の比較,” 音学会音声研究, S80-23 (1980).
- 23) 松本弘, 二村忠元, “正確判別分析による Text を限定しない話者認識,” 信学会電気音響研究, EA77-33 (1977).
- 24) 市川熾, 中島晃, 中田和男, “電話音声を対象とした話者照合,” 音響学会誌, 35, 2, 63-69 (1979).
- 25) S. Furui and A. E. Rosenberg, “Experimental studies in a new automatic speaker verification system using telephone speech,” 1980 IEEE Int. Conf. Acoust., Speech, Signal Processing, 1060-1062 (1980).
- 26) B. S. Atal, “Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification,” J. Acoust. Soc. Am., 55, 6, 1304-1312 (1974).
- 27) M. R. Sambur, “Speaker recognition using orthogonal linear prediction,” IEEE Trans. Acoust., Speech, Signal Processing, ASSP-24, 4, 283-289 (1976).
- 28) G. R. Doddington, “Personal identity verification using voice,” Proc. ELECTRO-76, 22-4, 1-5 (1976).
- 29) J. D. Markel and S. B. Davis, “Text-independent speaker recognition from a large linguistically unconstrained time-spaced data base,” IEEE Trans. Acoust., Speech, Signal Processing, ASSP-27, 1, 74-82 (1979).
- 30) E. Bunge, “Automatic speaker recognition by computers,” 1977 Conf. Pattern Recog., Image Processing, 88-93 (1977).
- 31) “Test results-Advanced development models of BISS identity verification equipment,” MITRE Technical Rep. 1, Rev. 1, MTR-3442 (1977).