HOKKAIDO UNIVERSITY

Title	計算機を用いた母音識別の一方法
Author(s)	伊土, 誠一; 斎川, 勝男; 栃内, 香次 他
Citation	北海道大學工學部研究報告, 59, 75-84
Issue Date	1971-03-10
Doc URL	https://hdl.handle.net/2115/41022
Type	departmental bulletin paper
File Information	59_75-84.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

計算機を用いた母音識別の一方法

伊土誠一* 斎川勝男*
栃内香次* 仲丸由正*

(昭和45年9月30日受理)

A Method of Discrimination between Japanese Vowels Based on Spectral Analysis

Seiichi IDO Katsuo SAIKAWA
Koji TOCHINAI Yoshimasa NAKAMARU

(Received September 30, 1970)

Abstract

In this paper we deal with a method of discrimination between Japanese vowels pronounced separately by means of spectral distributions. It is said that it is necessary to process a large amount of data for the recognition of vowels by means of spectral distribution, but the amount of data processed are so large that it is impossible to execute real-time operations. According to decrease processing data, we have attempted a metod of discrimination using about ten maximum values of time-varying amplitudes of the outputs from each band pass filter. We could determine any type of Japanese vowels in this manner, with the aid of a digital computer in a very short time.

In this experiment we used average values of vowels pronounced separately by several speakers as reference patterns, and recognized an input vowel by comparing its pattern against these references.

The results obtained in this study are as follows.

1. Vowels could be discriminated from others sufficiently by this method.
2. There is a correlation between the results of this experiment and Kullback's divergence.
3. It seems to be suitable to introduce this method in processing on real time.

1. 序　　言

音声の研究は相当以前から音響学，言語学，医学等多くの分野で行なわれてきたが，機械による識別の決定的な方法が見つからないうちに電子計算機の出現をみ，音声合成，文字の認識とともに音声識別も電子計算機を使用する傾向にある。

人間が音声を識別する際には，音声のスペクトル配置や強さとそれらの時間的変化等を考慮して，自分があらかじめ持っている音声パターンとの比較を基にして識別を行なっていると考えられる。そこで音声のスペクトル分布に着目し，あらかじめ何人かの音声を標準パターンとして作り，それで識別する方法を実験することにした。しかし従来より音声をスペクトル分析によって識別するには一つの音声に対し多くの識別パターンを作ることが必要であると考えられてきたが，この方法だと電子計算機は，膨大な情報を処理する必要があり，実時間処理はほとんど不可

* 電子工学科　電子機器工学講座

能である。

ここでは各帯域中心周波数につき振幅の時間変化中の最大値1点をサンプル点にとり，サンプル点が1点でも音声の識別が十分可能であるかどうかを調べた。この方法の最大の利点はサンプル点が各中心周波数につき1点のため判定が非常に高速度で行なえることと，記憶すべきパターンの数が少なくてすむことである。なおこの実験では連続音声あるいは動的な話し言葉の識別ではなく，音声識別としては初歩であるが一つ一つ切りはなして発声した離散的な母音の識別を行なった。

2. 識 別 実 験

2.1 実 験 方 法

本来は Fig. 1 のブロック図に示すような配置で識別実験を行なうつもりであったが，設備の都合で実際には Fig. 2 のブロック図によって行なった。これは入力音声を一時，テープレコーダーに録音しておきこれをフィルターにより順次周波数分析する方法である。音声は同一人物でも時間や日により，また環境や体の具合によりピッチや強さ，音色が変わり，したがってスペクトル成分の配置にも大きな変化を生じる。しかしここでは不特定多数の人の音声を識別することを一応の目標とした。また音声サンプルをとる場合には以前とった音声サンプルを発声者に聞かせただけで，特別な発声練習は行なわなかった。そのためか発声者のほとんどが無響室での発声が初めてで，反響がないため全体的にのどに力が入った感じの声が多かった。音声サンプルの中には他人が聞くと何という母音か識別しにくいものもあった。

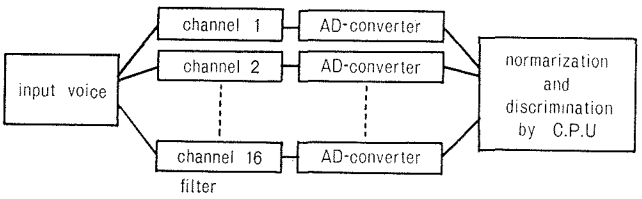

Fig. 1 Block diagram of experiment.

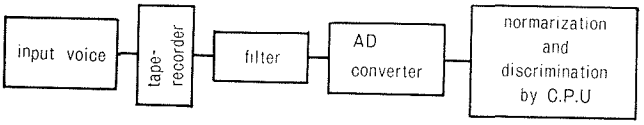

Fig. 2 Block diagram of experiment.

2.2 母音の周波数分析

予備実験により必要な周波数帯域は 125～4000 Hz あればよいことがわかったので，その間を 1/3 オクターブ間隔に 16 チャンネルに分け，これをフィルターの中心周波数にとった。フィルターとしては Brüel & Kjœr 社製で中心周波数から1オクターブ離れた点で約 50 dB (50 dB /octave) の減衰のあるものを使用した。

チャンネル n のフィルターの出力振幅を $f_n(t)$ とするとその最大値 $f_{n\max}$ を実験でのサンプル点とし，t は1母音を発声する継続時間にとる。

前述の方法で成人男性8人と成人女性8人の声を基にして，男声について標準パターンを二つ，女声について標準パターンを一つ作成し，これを識別標準パターンとした。音声は発音するたびに強さが大きく異なるので音声サンプルよりそのまま平均値を算出しても無意味であり，平均値を計算するためには音声サンプルを規格化する必要がある。本実験では下記のような手順で規格化を行なった。

次に示すように $n=1\sim3$ を除いた $f_{n\max}$ の中で最大の $f_{n\max}$ を求める。

$$F_{n\max} = \max(f_{4\max}, f_{5\max}, \cdots\cdots, f_{16\max}) \tag{2.1}$$

$f_{1\max}$, $f_{2\max}$, $f_{3\max}$ は男声のピッチ周波数であるが，人間はピッチ周波数が聞こえなくとも音声の識別に影響がない。このため本実験でも $f_{1\max}$～$f_{3\max}$ を抜き $f_{4\max}$ 以降について $F_{\max}$ を求めた。

規格化の最大値を N dB とすると，

$$\mathrm{D} = N - F_{\max} \tag{2.2}$$

$$F_{n\max} = f_{n\max} + \mathrm{D} \quad (n = 1\sim16) \tag{2.3}$$

この $F_{n\max}$ が規格化された値で，この値の平均をパターンの平均値とし，Fig. 3～Fig. 5 に示す。この図より各母音の平均値はそれぞれ周波数特性に大きな特徴があらわれており，特に500 Hz～3150 Hz で平均値の差が大きい。

女声の場合，ピッチ周波数と第一フォルマントが一致する場合がある。それゆえ規格化の欠点

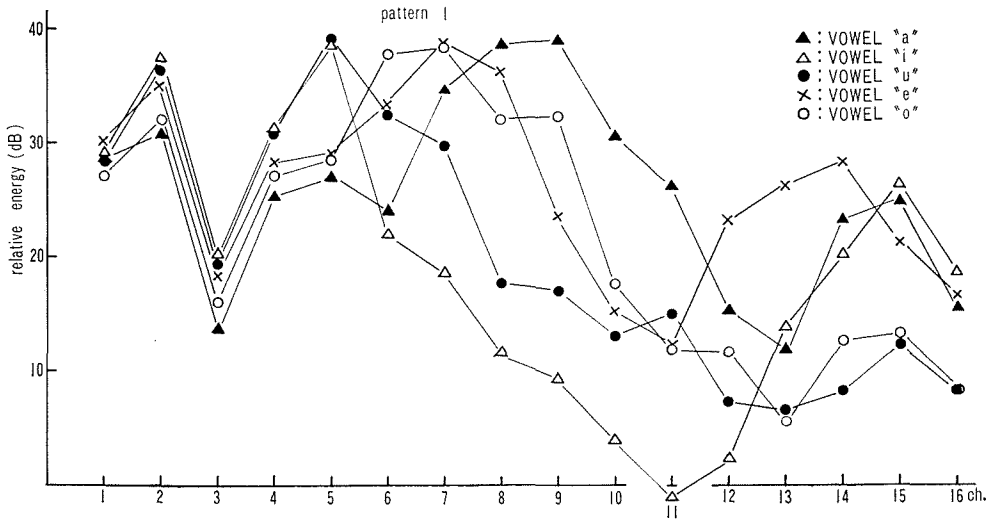

Fig. 3 Energy density spectra of Japanese vowels.

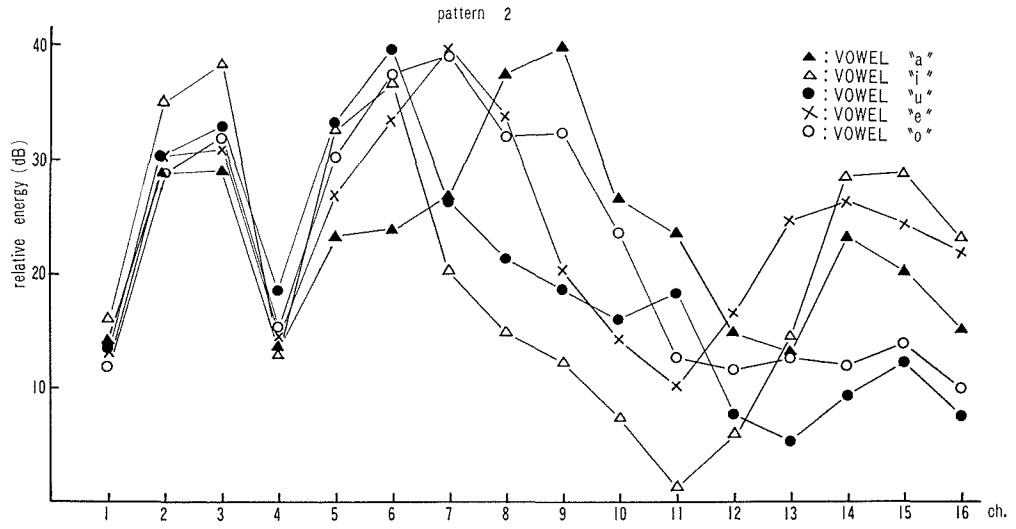

Fig. 4 Energy density spectra of Japanese vowels.

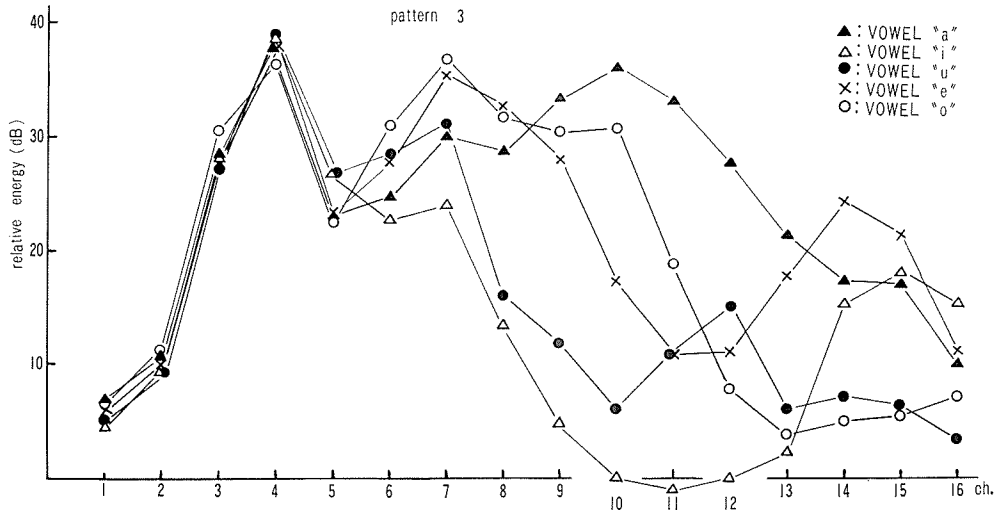

Fig. 5 Energy density spectra of Japanese vowels.

として，ピッチ周波数での $f_{n\max}$ が他の $f_{n\max}$ に比べ非常に大きいときピッチ周波数以外の $f_{n\max}$ が全体的に低くなって識別できなくなることがあげられる。この欠点を防ぐにはピッチ周波数に当る $f_{n\max}(f_{1\max}\sim f_{4\max})$ をあらかじめ小さくしておくとよいと思う。

$$\left.\begin{aligned} f'_{1\max} &= f_{1\max}-M \\ f'_{2\max} &= f_{2\max}-M \\ f'_{3\max} &= f_{3\max}-M \\ f'_{4\max} &= f_{4\max}-M \end{aligned}\right\} \tag{2.4}$$

ここで M の値を決めることが大きな問題となるが，これは標準パターン作成のための平均値と分散から決めるとよい。実験ではサンプル数が少なかったので，M のはっきりした値を決めることができなかったが，各母音によって若干違いはあるがほぼ3～5 dB ぐらいが適当であると思われる。

2.3 パターンの決定と母音の識別法

(1) 母音の識別 $f_{n\max}$ は母音 m の平均値 $a_{vn}(m)$ より離れれば離れるほど母音 m である可能性が小さくなる。実験ではそのことに着目し，$f_{n\max}$ が平均値 $a_{vn}(m)$ からどのくらい離れているかによって式 (2.5) に示すように 1, 0, −1 の3段階の重みをつけた。パターン l において，入力音声が母音 m である可能性は，

$$\left.\begin{aligned} &a_{3n}(l,m)\leqq f_{n\max}\leqq a_{2n}(l,m) \quad \text{のとき}\ g_n(m)=1 \\ &a_{2n}(l,m)<f_{n\max}\leqq a_{1n}(l,m) \\ &\quad \text{または} \\ &a_{4n}(l,m)\leqq f_{n\max}<a_{3n}(l,m) \quad \text{のとき}\ g_n(m)=0 \\ &f_{n\max}<a_{4n}(l,m) \\ &\quad \text{または} \\ &a_{1n}(l,m)<f_{n\max} \quad \text{のとき}\ g_n(m)=-1 \end{aligned}\right\} \tag{2.5}$$

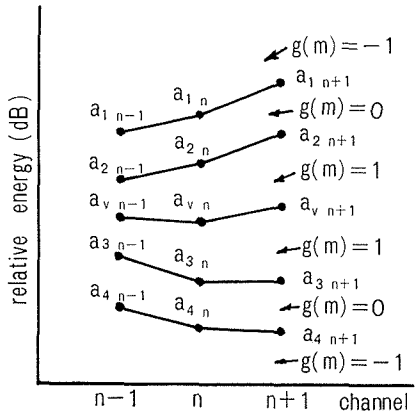

Fig. 6 Relations between a_{1n}, a_{2n}, a_{vn}, a_{3n}, a_{4n} and $g(m)$.

式 (2.5) よりチャンネル n における母音 m の可能性がわかり，これを全チャンネルにわたって加えあわせた式 (2.6) より入力音声パターン l が母音 m である可能性が求められる。

$$G(m)=\sum_{n=5}^{16} g_n(m) \tag{2.6}$$

次に $G(m)$ の最大のものを求め $G_{\max}(m_1)$ にする m_1 を入力音声と判定する。

$$G_{\max}(m)=\max\,(G(a), G(i), G(u), G(e), G(o))$$

ここで a_{1n}, a_{2n}, a_{3n}, a_{4n} を判定することが識別パターンを決定することであり，a_{1n}, a_{2n}, a_{3n}, a_{4n} の値が識別率に大きく影響を与える。本実験では標準パターン作成のための予備実験の平均と分散よりこの値を定めた。識別パターンの決定に使用した各パターンの分散を Table 1 に示す。

Table 1 Variance for vowels in each points

pattern	vowel	315Hz	400Hz	600Hz	630Hz	800Hz	1000Hz	1250Hz	1600Hz	2000Hz	2500Hz	3150Hz	4000Hz
pattern 1	a	12.5	7.2	6.0	6.1	3.1	22.5	18.9	17.9	20.7	18.0	17.4	16.0
	i	5.1	23.8	19.3	18.5	13.5	12.3	2.0	6.3	8.7	40.2	35.0	39.0
	u	4.3	20.7	21.9	19.4	12.1	23.7	23.1	28.3	15.2	35.7	26.8	58.2
	e	19.6	14.9	7.2	31.6	24.8	21.1	13.0	20.3	29.4	24.3	30.3	31.5
	o	13.2	8.3	7.5	15.4	31.9	33.6	21.6	29.6	34.2	42.9	43.7	55.1
pattern 2	a	3.9	6.2	12.6	4.4	0.2	16.6	17.9	5.5	9.8	3.9	6.7	19.1
	i	16.9	11.3	17.0	28.7	29.5	35.2	14.9	33.2	13.0	7.0	7.3	14.3
	u	18.3	0.9	10.5	8.3	9.6	9.3	13.3	18.4	23.5	10.7	10.8	16.5
	e	9.9	35.3	1.8	21.4	17.2	37.9	38.7	23.1	9.0	6.4	9.4	9.2
	o	15.3	21.1	5.0	9.7	10.3	10.2	40.2	68.4	62.8	10.8	12.1	18.7
pattern 3	a	26.6	38.6	14.6	29.0	18.2	15.5	21.7	21.2	18.8	29.1	13.1	22.1
	i	41.1	44.7	23.9	36.2	17.9	4.7	2.4	6.0	16.0	54.7	54.0	41.6
	u	24.4	41.6	18.4	26.2	16.9	18.5	25.5	10.7	7.0	21.6	15.3	23.7
	e	26.2	35.0	18.5	30.3	23.1	7.9	12.1	13.0	11.9	16.4	8.6	11.1
	o	14.6	55.5	15.9	35.3	16.3	21.8	17.7	7.9	8.3	10.6	9.5	12.3

(2) 識別パターンの決定　パターン l において，チャンネル n での母音 m の分散を $V_n(l, m)$ とすると a_{1n}, a_{2n}, a_{3n}, a_{4n} の値は式 (2.8)～(2.11) に従って決定できる。

$$a_{2n}=a_{vn}(m)+\frac{1}{2}\sqrt{V_n(l, m)}+5 \quad \text{(dB)} \tag{2.8}$$

$$a_{3n}=a_{vn}(m)-\frac{1}{2}\sqrt{V_n(l, m)}-5 \quad \text{(dB)} \tag{2.9}$$

$$a_{1n}=a_{2n}+6 \quad \text{(dB)} \tag{2.10}$$

$$a_{4n}=a_{3n}-6 \quad \text{(dB)} \tag{2.11}$$

母音の識別においてチャンネル 1～チャンネル 4 (中心周波数 125 Hz～250 Hz) はピッチ周波数の高低，すなわち入力音声がどのパターンに入るかを定めるために使用し，$G_{\max}(m_1)$ の決定には使用していない。また，チャンネル 5 (中心周波数 315 Hz) も各母音の平均値に差がないので使用していない。このようにどのチャンネルが母音識別に重要であるかを求めることは識別率や識

別時間に大きな影響を及ぼす。この判定に関しては第三章で実験に用いた Kullback's divergence を使用する方法について述べる。

以上述べた方法にしたがって，北大大型計算機センターの電子計算機 FACOM 230-60 を用いて前節で述べた規格化と本節の識別部分のプログラムを作成し，実験をおこなった。なお，このプログラムは FORTRAN で組んである。

3. 実験結果および考察

前述のような方法により識別実験を行なった結果を Table2, 3 に示す。その結果，男声については，パターン作成に用いた人の声に対する正答率は 95.8%，誤り率は 3.2%，判定不能は 1.0% であった。またパターン作成に用いなかった人の声の場合は，正答率 96.0 %，誤答率 1.8 %，判定不能 2.2 %であり，識別率はパターン作成の使用の有無にはほとんど関係しないという結果が得られた。女声については正答率が 98.7%，誤答率1.3%，判定不能が 0.0% であった。このことより母音の場合には離散的であればサンプル点が 1 点でも十分識別が可能であると思われる。また今後のプログラムの改良，フィルターの特性の改良，フィルターの中心周波数の移動等により正答率を上げることも期待できる。

次に判定誤りの原因を調べてみると，誤りはほぼ次のように分類できる。

(1) 母音イとエの特性的に高い調子は口むろの共鳴域でできる第二フォルマント (2000～4000Hz) によるが，人によって第 1 フォルマントと第二フォルマントのスペクトルエネルギーの比が多少異なるため，この比が大きすぎると，第二フォルマントの特徴が抽出されにくくなり，判定誤りが生ずる。この原因による誤りは，第一フォルマントが一致する「イ」と「ウ」，「エ」と「オ」の間で起きる。

(2) 「イ」と「エ」，「ウ」と「オ」は周波数が 2000～4000 Hz でほぼ $f_{n\max}$ の値が同じであるから第一フォルマント付近で $f_{n\max}$ の多少の変動が生ずると識別誤りを起こす。

Table 2, 3 における判定不能は $G_{\max}(m)$ にする m が二つ以上ある場合で，実験ではすべてその中に正しい母音が含まれていた。また識別誤りの場合も $G_{\max}$ と正しい母音の $G(m)$ の差がほとんど 1～2 であった。

識別率と同様，実時間処理においては，判定に要する時間が重要であり本実験の判定に必要な時間 T はほぼ

$$T = T_F + T_{AD} + T_N + T_P \tag{3.1}$$

ここで T_F : 周波数分析に要する時間

T_{AD} : AD 変換に要する時間　　T_N : 規格化に要する時間

T_P : 判定に要する時間

Table 2 Results of experiment for 16 males

		number of judgment						total
		a	i	u	e	o	reject	
input vowels	a	123				1	2	126
	i		119		4		2	125
	u	1	2	115		5	3	126
	e				122		2	122
	o				4	122		126

Table 3 Results of experimet for 9 females

		number of judgment						total
		a	i	u	e	o	reject	
input vowels	a	42				1		43
	i		44					44
	u			46				46
	e			2	44			46
	o					46		46

この実験では $T_N+T_P \fallingdotseq 100\,ms$ ぐらいであった*。人が音声を識別する場合，音声の前後に，forward masking と backward masking が合せてほぼ $100\,ms$[1] ぐらいあるので，識別時間はほぼ ($100\,ms$＋次の音声の発音継続時間) を目安にするとよく，したがって母音についてはかなりよい値である。子音を含む音声識別ではこれより判定時間を短かくする必要があると思われる。

4. Kullback's divergence による識別難易度の評価

本章では識別の尺度として Kullback の定義した確率密度関数間の距離 (Kullback's divergence) を用いて異なる母音の識別の難易度と各周波数における識別の寄与率を求めた結果について述べる。

母音 m_1 と母音 m_2 間の Kullback's divergence $D_{m_1 m_2}(f_1, f_2, \cdots\cdots, f_n)$ は式 (4.1) で定義されている。

$$D_{m_1 m_2}(f_1, f_2, \cdots\cdots, f_n) = \int_{-\infty}^{+\infty} P_{m_1}(\boldsymbol{F}) \ln (\boldsymbol{F})\, d\boldsymbol{F} - \int_{-\infty}^{+\infty} P_{m_1}(\boldsymbol{F}) \ln L(\boldsymbol{F})\, d\boldsymbol{F} \qquad (4.1)^{3)}$$

ただし $L(\boldsymbol{F}) = \dfrac{P_{m_1}(\boldsymbol{F})}{P_{m_2}(\boldsymbol{F})}$

$P_{m_1}(\boldsymbol{F}) = P(f_1, f_2, \cdots\cdots, f_n)$: 母音 m_1 の n 次元確率密度関数

$D_{m_1 m_2}$ は情報量やあいまい度によく似た性質があるので，$D_{m_1 m_2}$ を識別難易度を調べる尺度と考えてもよいと思われる。

ここで各周波数における $F_{n\max}$ が正規分布し，$F_{n\max}$ の分散と平均値は各母音によって異なると考えると，$P_{m_1}(F_{1\max}, F_{2\max}, \cdots\cdots, F_{n\max})$ は

$$P_{m_1}(\boldsymbol{F}) = \frac{1}{(2\pi)^{1/2} |\boldsymbol{V}|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{F}-\boldsymbol{U}_{m_1})' \boldsymbol{V}'_{m_1} (\boldsymbol{F}-\boldsymbol{U}_{m_1})\right] \qquad (4.2)$$

ただし，$\boldsymbol{V}_{m_1}$: m_1 の共分散行列
$\boldsymbol{U}_{m_1}$: m_1 の平均値ベクトル
$\boldsymbol{V}^{-1}$: 逆行列　$\boldsymbol{U}'$: 転置行列

ゆえに

$$\int_{-\infty}^{+\infty} P_{m_1}(\boldsymbol{F}) \ln L(\boldsymbol{F}) d\boldsymbol{F} = \frac{1}{2} \ln \frac{|\boldsymbol{V}_{m_2}|}{|\boldsymbol{V}_{m_1}|} + \frac{1}{2} t_r\, \boldsymbol{V}_{m_1} (\boldsymbol{V}_{m_2}^{-1} - \boldsymbol{V}_{m_1}^{-1}) + \frac{1}{2} t_r\, \boldsymbol{V}_{m_1}^{-1} (\boldsymbol{U}_{m_1} - \boldsymbol{U}_{m_2})(\boldsymbol{U}_{m_1} - \boldsymbol{U}_{m_2})' \qquad (4.3)$$

ただし，t_r : 対角和

同様に $\int_{-\infty}^{+\infty} P_{m_1}(\boldsymbol{F}) \ln L(\boldsymbol{F}) d\boldsymbol{F}$ を求め (4.3) 式とともに式 (4.1) に代入すると $D_{m_1 m_2}(f_1, f_2, \cdots\cdots, f_n)$ は

$$D_{m_1 m_2}(f_1, f_2, \cdots\cdots, f_n) = \frac{1}{2} t_r (\boldsymbol{V}_{m_1} - \boldsymbol{V}_{m_2})(\boldsymbol{V}_{m_2}^{-1} - \boldsymbol{V}_{m_1}^{-1}) + \frac{1}{2} t_r (\boldsymbol{V}_{m_1}^{-1} + \boldsymbol{V}_{m_2}^{-1})(\boldsymbol{U}_{m_1} - \boldsymbol{U}_{m_2})(\boldsymbol{U}_{m_1} - \boldsymbol{U}_{m_2})' \qquad (4.4)$$

となる。

実験では各母音につき 40 サンプルをとり，その共分散ベクトルと平均値ベクトルより $D_{m_1 m_2}$

* 計算機の CPU 使用時間による。

($\boldsymbol{F}$) を計算した。その結果を Table 4～Table 6 に示す。

ここでパターン2の母音アと他の母音の間の $D_{m_1 m_2}(\boldsymbol{F})$ の値が大きいが，これはパターン2に当る人が実験者中で1人のため，パターンの作成のとき 800 Hz での分散が非常に小さい値であったためである。Table 4～6 より，パターン1にあっては母音ア—オ，ウ—エ，ウ—オ，エ—オ間の判定が難しく，またパターン2においては母音イ—ウ，イ—エ，ウ—オ，エ—オ間で誤りが起り易い。これを Table 2 と比べると Kullback's divergence による判定の難易度と Table 2 の識別結果が非常によく一致していることがわかる。同様なことがパターン3の場合にも Table 3 と Table 6 よりいえる。これは Kullback's divergence の値を計算することが音声識別の難易度を決める尺度になることを示している。

各周波数については，パターン1では，315 Hz, 3500 Hz, 4000 Hz が，パターン2では 315 Hz, 1600 Hz が，パターン3では 315 Hz, 400 Hz, 500 Hz, 4000 Hz があまり判定に寄与しないことがわかる。最後に Kullback's divergence と識別率の関係が求められているので参考として Fig. 7 に示す。[3)]

Table 4 Kullback's divergence for individual test points in the pattern 1

	test points												
	315 Hz	300 Hz	500 Hz	630 Hz	800 Hz	1000 Hz	1250 Hz	1600 Hz	2000 Hz	2500 Hz	3150 Hz	4000 Hz	11 points
Dai($\boldsymbol{F}$)	19.9	1.2	9.8	80.8	177.2	45.5	211.3	19.2	0.7	0.7	2.6	1.0	1,383.9
Dau($\boldsymbol{F}$)	23.2	6.6	1.4	47.6	102.0	13.7	6.1	3.1	1.7	9.9	2.2	3.1	574.7
Dae($\boldsymbol{F}$)	2.1	9.1	18.4	2.3	47.5	11.0	12.6	3.3	8.4	1.2	0.4	0.3	274.4
Dao($\boldsymbol{F}$)	1.9	23.5	16.8	5.6	12.6	6.5	10.3	0.7	1.7	5.0	1.7	3.1	215.1
Diu($\boldsymbol{F}$)	0.0	4.7	6.0	2.1	4.4	5.3	75.1	3.8	5.0	3.8	6.7	2.6	329.3
Die($\boldsymbol{F}$)	5.9	7.3	38.9	26.3	11.7	8.5	54.2	47.2	12.2	2.2	0.8	0.2	434.9
Dio($\boldsymbol{F}$)	7.0	20.5	36.0	24.9	28.2	11.0	49.9	10.4	6.0	1.4	4.4	2.7	471.9
Due($\boldsymbol{F}$)	7.3	0.2	8.1	14.3	3.0	0.3	0.6	11.0	19.5	13.9	2.9	1.9	205.0
Duo($\boldsymbol{F}$)	8.5	3.0	7.1	11.7	14.1	0.8	0.4	0.8	0.4	0.5	0.2	0.0	159.4
Deo($\boldsymbol{F}$)	0.1	1.8	0.0	1.2	2.8	0.3	0.1	5.6	13.4	8.1	1.8	1.9	149.3

Table 5 Kullback's divergence for individual test points in the pattern 2

	test points												
	315 Hz	400 Hz	500 Hz	630 OH	800 Hz	1000 Hz	1250 Hz	1600 Hz	2000 Hz	2500 Hz	3150 Hz	4000 Hz	11 points
Dai($\boldsymbol{F}$)	15.0	19.8	7.1	68.3	2,190.0	17.3	30.3	11.0	0.2	6.2	11.2	4.0	3,469.5
Dau($\boldsymbol{F}$)	16.4	149.5	0.0	45.6	1,282.2	9.8	1.9	7.0	5.0	35.9	7.6	3.1	2,569.5
Dae($\boldsymbol{F}$)	2.9	9.8	57.0	3.3	1.130.9	7.2	7.5	1.4	14.1	2.1	2.2	4.1	2,035.4
Dao($\boldsymbol{F}$)	8.8	18.5	21.4	5.4	188.5	1.0	5.2	6.3	2.3	2.8	0.9	1.5	530.2
Diu($\boldsymbol{F}$)	0.0	10.2	2.9	3.9	3.5	6.4	20.2	0.3	5.3	45.1	31.9	15.5	274.8
Die($\boldsymbol{F}$)	2.7	1.4	117.4	14.5	3.0	1.4	4.2	4.2	9.4	0.8	2.8	0.2	270.4
Dio($\boldsymbol{F}$)	0.4	1.5	44.5	20.1	26.8	17.5	6.4	1.1	1.7	33.1	25.5	11.1	396.1
Due($\boldsymbol{F}$)	3.2	40.6	58.7	12.1	0.3	1.4	3.8	3.7	29.4	36.8	14.0	17.2	339.1
Duo($\boldsymbol{F}$)	0.5	13.7	22.8	12.6	18.8	5.6	2.2	1.5	2.2	0.1	0.2	0.2	189.2
Deo($\boldsymbol{F}$)	1.0	0.7	0.8	0.6	11.6	6.2	0.1	1.3	11.6	25.9	10.3	12.3	171.6

Table 6 Kullback's divergence for individual test points in the pattern 3

	test points												
	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1000 Hz	1250 Hz	1600 Hz	2000 Hz	3150 Hz	2500 Hz	4000 Hz	11 points
Dai($\boldsymbol{F}$)	0.5	0.1	2.0	7.3	45.4	182.5	272.4	104.3	20.9	0.3	1.2	1.1	864.7

Dau(***F***)	0.6	0.3	0.1	5.7	26.6	53.7	21.0	18.3	23.2	4.1	8.1	1.9	441.7
Dae(***F***)	0.0	0.3	1.8	0.5	1.4	33.7	31.7	25.3	0.9	2.6	2.1	0.3	332.5
Dao(***F***)	0.2	1.0	2.9	0.3	0.5	1.7	10.7	47.6	26.6	10.1	12.2	0.8	141.7
Diu(***F***)	0.1	0.7	2.4	0.3	2.8	6.0	36.1	29.8	1.9	2.6	6.6	4.7	303.0
Die(***F***)	0.4	0.7	5.9	11.2	27.2	51.6	36.5	15.0	17.9	4.1	3.0	2.0	320.9
Dio(***F***)	1.3	1.5	8.2	9.5	38.6	123.6	94.1	8.9	0.5	7.6	11.7	4.4	415.2
Due(***F***)	0.4	0.0	0.9	9.6	13.8	12.0	0.3	1.5	15.8	16.0	21.0	4.1	267.7
Duo(***F***)	1.0	0.2	1.7	8.3	20.9	30.3	3.0	5.9	0.7	0.6	0.2	1.0	221.3
Deo(***F***)	0.2	0.3	0.1	0.0	0.3	15.7	4.3	1.2	20.0	29.3	28.6	1.5	323.4

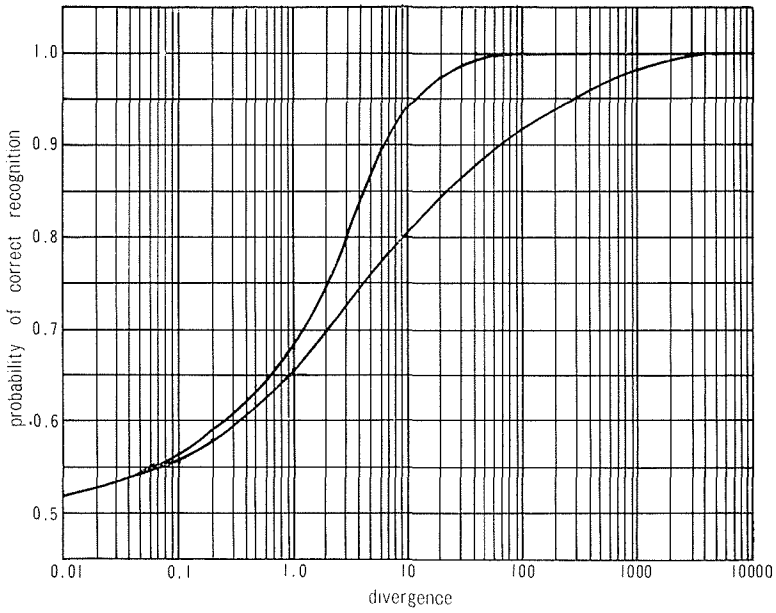

Fig. 7 Upper and lower bounds on probabdliy of correc t recognition as a function of divergence, for normal variables with unequal covariance matrixes.

5. 結　　言

実験の最終目的は電子計算機を用いた子音をも含む音声の実時間識別であり，著者らが述べてきた母音の識別は，その第一歩であるが，このことをも考慮した本実験の判定法の特徴として次のことがいえる。

(1) サンプル点が各中心周波数につき1点のため電子計算機による処理時間が非常に短かくてよく，実時間処理に適していると思われる。

(2) 基本となるパターンの数が増しても各パターンに対し並列にプログラムが動作するため，処理時間がほとんど増加しない。

(3) すべての母音について $G(m)$ が計算されるので識別誤りが生じても，$G(m)$ の2番目に大きい母音が正しい音声である場合がほとんどであり，今後音声識別に学習，または単語単位の識別を導入する場合に大きな利点となる。

(4) 識別率に大きく影響するのは a_{1n}, a_{2n}, a_{3n}, a_{4n} であるが，これは式 (2.8)～(2.11) をあらかじめプログラムに組んでおくと，同一人物についての識別は非常に容易である。

最後に本講座の牧野君はじめ実験に協力くださった方々に深く感謝します。

参 考 文 献

1) 伊福部達：電子通信学会医用電子と生体工学専門委員会資料，13 (昭45)
2) T. Marill, D. M. Green: IEE, IT-9, (1963), p.11～p.17.
3) S. Kullback: Information theory and statistics, 1952, Wiley.
4) 坂井利之：パターン認識の理論，(昭42)，共立出版