一般化 Hough 変換と連続 DP マッチングを用いた ワード スポットティング

非会員 杉 野 浩 司 (京都工芸繊維大学)

非会員 新 美 康 永 (京都工芸繊維大学)

Word Spotting based on the Generalized Hough Transform and continuous DP matching

Sugino Kouji, Non-member, Niimi Yasuhisa, Non-member (Kyoto Institute of Technology)

This paper describes a method to detect particular speech segments as a keyword from speech database. The Generalized Hough Transform (GHT) has been used to find arbitrary complex shapes on the image plane. We extended the GHT to be able to deal with three-dimensional shapes and applied it to the keyword spotting in which a speech signal is represented as a spectral sequence. Based on experimental results on the keyword spotting in which 25 keywords were tried to extract from 750 utterances produced by 30 male speakers, we propose an approach for improving the keyword spotting performance. This approach combines GHT and a word spotting method based on DP matching.

キーワード：スポットティング, 一般化 Hough 変換, 連続 DP マッチング, スペクトル時系列

1. はじめに

近年、動画像や音声などのデータを扱うマルチメディアと呼ばれる分野が注目されており、記憶媒体の大容量化とともにこれらのデータをコンピュータの記憶装置にも多量に蓄積することができるようになり、巨大なサイズの画像・音声のデータベースが増えていくことが考えられる。このような巨大なデータベースの応用として、例えば、ニュースを朗読した音声を多量に蓄積しておき、必要に応じて、特定のキーワードを含むニュースを取り出す事などがある⁽⁹⁾。そのためには、通常キーボードから検索したい単語あるいは句を入力し、これをもとに検索するワード スポットティング (word spotting) と呼ばれる方法が使われる。このようなシステムでは、検索情報に相当する隠れマルコフモデルを自動的に生成して、入力音声とマッチングをとることでよりキーワードを検出することができる⁽¹⁾⁻⁽⁴⁾。

本論文では、検索情報をキーボードからではなく音声で入力する場合を考える。このような状況では、ワード スポットティングを行うために連続 DP マッチング法が用いられてきた⁽⁵⁾。しかし、DP マッチング法は音声パターンの時間軸方向の非線形な変動をうまく吸収し、特定話者の音声認識では強力な方法であるが、異なる話者の音声のマッチングにはそれほど有効ではない。異なる話者間では時間軸方向よりも、周波数軸方向の変動、すなわち、スペクトル形状の変動の方が大きなウェイトを占め、単純な DP マッチ

ング法では効果が上がらない。音声スペクトルの周波数軸方向の変動をもたらす最も大きな要因は、話者の声道長の差である。声道の形状が長さ方向に相似的に変形した場合、音声スペクトルのピークの位置は声道長の逆数に比例して変化する。音声スペクトルの周波数軸を対数に変換すれば、声道長に起因するスペクトルの変動は、対数周波数軸方向の平行移動として表現される。

パターンの平行移動に対して不変なマッチング能力をもつ方法に Hough 変換がある。我々は任意の二次元形状を検出できる一般化 Hough 変換を拡張し、三次元の特徴量を持つ音声の時系列スペクトルに対してワード スポットティングを行なう手法⁽¹¹⁾を提案した。そして、予備実験として音声スペクトルの周波数軸が、線形の場合とメルスケール (近似的に 1000Hz 以下は線形、1000Hz 以上は対数周波数) の場合とでワード スポットティングの性能を比較した。その結果メルスケールの場合にその性能が大幅に改善されることを確かめた⁽¹²⁾。もちろん、話者間のスペクトルの変動は、声道長の変化によるものだけではないことは明らかであるが、この予備実験により、Hough 変換が話者間の周波数軸方向の変動をある程度吸収する能力をもっていることが確かめられた。

そこで本稿では音声の時間軸方向の変動を吸収する能力の高い連続 DP マッチング法によるワード スポットティングを併用することにより、提案法の精度の向上を図った。すなわち、一般化 Hough 変換によりキーワードの存在する

候補区間を定め、その区間内で連続 DP マッチングを行ない、キーワードの存在区間と長さを求める。この情報をもとに再度一般化 Hough 変換を適用することによってスポッティング精度を向上することができた。

以下に本論文の構成を述べる。2. では一般化 Hough 変換の説明を行なう。3. では一般化 Hough 変換を 3 次元空間に拡張する。さらに 3 次元空間として音声のスペクトル時系列を考え、一般化 Hough 変換を用いたワードスポッティングの原理を説明する。4. では一般化 Hough 変換と連続 DP マッチング法を併用して、ワードスポッティングの精度を改善する方法を説明する。5. では実験の結果と考察を述べる。

2. 一般化 Hough 変換

Hough 変換は、画像中の特徴点が与えられた時、その特徴点を含むような図形に関する仮説を立て (投票)、多くの特徴点から同一の仮説が生成される時 (多数決)、その仮説を正しいものと考えて図形の検出を行なう手法である。直線や楕円のように、代数方程式によって表現される解析的対象を検出する Hough 変換では、方程式中のパラメータによって張られる空間上で投票と多数決を行なえばよい⁽⁶⁾。しかし、パラメータによって表現することができない任意形状の幾何学的対象を検出する場合には、このような方法を用いることができない。Ballard は、「一般化 Hough 変換」を提案し、この問題を解決した⁽⁷⁾。これは「形状をデータとして与え、ポーズ (位置・姿勢・大きさ) に対応する座標変換パラメータによって対象物を表す」という対象の表現法である。この場合には形状のパラメータ表現は不要であり、任意形状の対象を表現することが可能になる。このように「形状を固定」し、座標変換パラメータ空間上で投票と多数決を行なう Hough 変換が一般化 Hough 変換である。具体的な変換方法を以下に示す。

[パラメータ空間]

図形の変形には、平行移動、線形な伸縮、回転などが考えられるが、ここではワードスポッティングへの応用で用いる平行移動だけを考え、平行移動 (u, v) によって張られる 2 次元空間をパラメータ空間とする。

[形状モデルの表現]

基準点 O を任意に定め、検出対象図形の境界線上の点 (X_i, Y_i) において、 O に向かうベクトル $\vec{r}_i$ と、その点における画像の濃度勾配方向 ϕ_i を求める。(図 1)

2π を n 個に量子化した $\phi^{(1)} \sim \phi^{(n)}$ をインデックスとして、 ϕ_i が含まれるインデックスのエントリに $\vec{r}_i$ を登録する。このようにして作成された図 2 のような表を R-table という。

[投票と図形の検出]

対象画像上の特徴点 (X_i, Y_i) およびその点における ϕ_i を求める。 ϕ_i に対応するインデックスに登録されているエントリベクトル $\vec{r}$ すべてに対して、以下の式によって平行移動 (u, v) を求め、点 (u, v) の投票度数を 1 増やす。(図 3)

図 1 形状の記述

Fig. 1. Description of a figure

$\phi^{(1)}$	$\vec{r}_{1,1}, \vec{r}_{1,2}, \vec{r}_{1,3}(=\vec{r}_i), \vec{r}_{1,4}, \dots$
$\phi^{(2)}$	$\vec{r}_{2,1}, \vec{r}_{2,2}, \dots$
$\dots$	
$\phi^{(n-1)}$	$\vec{r}_{n-1,1}, \vec{r}_{n-1,2}, \dots$
$\phi^{(n)}$	$\vec{r}_{n,1}, \vec{r}_{n,2}, \dots$

図 2 濃度勾配方向をインデックスとした R-table

Fig. 2. R-table indexed by the direction of gradient

図 3 特徴点毎の投票とピーク検出

Fig. 3. Vote and peak detection

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} X_i \\ Y_i \end{pmatrix} + \vec{r} \dots \dots \dots (1)$$

すべての特徴点に対して上の投票が終われば、投票度数の高いパラメータ (u, v) を抽出する。このパラメータの位置に形状モデルの基準点を重ねるようにして写像したとき、その位置に形状モデルのインスタンスが存在する可能性が高い。

3. ワードスポッティングへの適用

上で述べた一般化 Hough 変換をワードスポッティングに適用するには、変換の際にキーワードを形状モデルとして R-table を作成し、対象音声データに対して投票・多数決を取ることににより、実現することができる。

本研究ではデータとして与える形状として、図 4 に示し

図4 スペクトル時系列
Fig. 4. A spectral pattern

たような音声から得られるスペクトル時系列を扱うことにする。したがって形状データによって張られるパラメータ空間は、時間軸・周波数軸・パワー軸の3次元となる。そのため一般化 Hough 変換を3次元に拡張する。

以下にその具体的な方法を説明する。

[パラメータ空間]

時間、周波数、パワーによって張られる3次元空間における平行移動ベクトルをパラメータ空間とする。

[形状モデルの表現]

検出したい音声区間のスペクトル時系列(参照データ)を求める。次に参照データの重心を基準点 O として選ぶ。そして参照データの各点 (X_i, Y_i, Z_i) から基準点に向かうベクトル $\vec{r}_i$ を求め、R-table に登録すればよい。

ここで、パラメータ空間が2次元の場合では R-table のインデクスに画像の濃度勾配方向を用いていたのに対し、パラメータ空間が3次元であることを考慮して、R-table のインデクスとして特徴点に接する平面の法線方向を用いる。

参照データ上の点 (X_i, Y_i, Z_i) に接する平面 α_i を考える。接平面 α_i を $z = f(x, y)$ で表せば

$$z = Z_i + (x - X_i) \frac{\partial f}{\partial X_i} + (y - Y_i) \frac{\partial f}{\partial Y_i} \quad \dots\dots (2)$$

となる。よって点 (X_i, Y_i, Z_i) における接平面の法線方向は

$$\left(\frac{\partial f}{\partial X_i}, \frac{\partial f}{\partial Y_i}, -1 \right) \quad \dots\dots (3)$$

で表される。これから以下の式によって得られる法線方向の方位角 θ と仰角 ϕ を R-table のインデクスとして用いる。

$$\theta = \tan^{-1} \left(\frac{\partial f}{\partial Y_i} / \frac{\partial f}{\partial X_i} \right), \quad 0 \leq \theta < 2\pi \quad \dots\dots (4)$$

$$\phi = \tan^{-1} \left(\frac{1}{\sqrt{\left(\frac{\partial f}{\partial X_i}\right)^2 + \left(\frac{\partial f}{\partial Y_i}\right)^2}} \right), \quad 0 \leq \phi < \frac{\pi}{2} \quad \dots\dots (5)$$

なおスペクトル時系列は離散量なので、 $\partial f / \partial X_i, \partial f / \partial Y_i$ の計算には図5に示すオペレータを用いる。

以上から各点 (X_i, Y_i, Z_i) の $\vec{r}_i$ を、方位角と仰角をそれぞれ m, n 個に量子化したインデクス $\theta^{(1)} \sim \theta^{(m)}, \phi^{(1)} \sim \phi^{(n)}$ に対応するエントリに登録することによって R-table が作成される。(図6, 7参照)

[投票と検出]

検出される対象となる音声のスペクトル時系列(対象データ)を求める。対象データの各点 (X_i, Y_i, Z_i) における接平面の法線方向 (θ_i, ϕ_i) を求める。 (θ_i, ϕ_i) に対応するインデクスに登録されているエントリベクトル $\vec{r}$ すべてに対して、パラメータ (x, y, z) を以下の式によって求め、点 (x, y, z) の投票度数を1増やす。

-1	0	1
-2	0	2
-1	0	1

(a) x軸方向

-1	-2	-1
0	0	0
1	2	1

(b) y軸方向

図5 勾配計算に用いるオペレータ
Fig. 5. Operators for gradient calculation

図6 3次元形状の記述
Fig. 6. Description of a 3-dimensional figure

	$\phi^{(1)}$	...	$\phi^{(n)}$
$\theta^{(1)}$	$\vec{r}_{1,1,1}, \vec{r}_{1,1,2}, \dots$		$\vec{r}_{1,n,1}, \vec{r}_{1,n,2}, \dots$
$\theta^{(2)}$	$\vec{r}_{2,1,1}(=\vec{r}_i), \dots$		$\vec{r}_{2,n,1}, \vec{r}_{2,n,2}, \dots$
...			
$\theta^{(m)}$	$\vec{r}_{m,1,1}, \vec{r}_{m,1,2}, \dots$		$\vec{r}_{m,n,1}, \dots$

図7 θ と ϕ をインデクスとした R-table
Fig. 7. R-table indexed by θ and ϕ

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} + \vec{r} \dots\dots\dots (6)$$

すべての対象データ点に対して上の投票が終われば、投票度数の高いパラメータを抽出する。このパラメータの位置に、検出したい音声区間のインスタンスが存在する可能性が高い。

4. スポッティング精度の向上

以上に述べた一般化 Hough 変換を用いたワードスポッティング法による実験を行ない、その結果からその問題について検討した⁽¹¹⁾。

キーワードの検出に失敗する原因の一つとして、参照データとして R-table の作成に用いた発話の時間長と、対象データに存在するキーワードの時間長が異なっていることが挙げられる。つまり同じキーワードを発話する場合でも、毎回その発話時間長は変化する。3. で述べた方法ではキーワードから作られる R-table を基に投票が行なわれるという性質から、パラメータ空間の投票度数はキーワード候補が参照データの時間長で発話されているという仮定のもとで得られた結果である。よってこの参照データと対象データのキーワード発話時間長の違いが投票結果に誤差として出てくる。

ここでは上に述べた投票誤差を少なくするために、DP マッチング法を利用し一般化 Hough 変換を二段で行なう。つまり一段目の一般化 Hough 変換でキーワード候補領域を求め、そして各候補領域内で参照データとの DP マッチングを行なう。そしてその結果を二段目の一般化 Hough 変換に反映させる。以下にその方法について述べる。

〈4・1〉 一段目の一般化 Hough 変換 ここでは 3. に述べた方法で一段目の一般化 Hough 変換を行なう。そして設定された閾値よりも大きい投票度数をもつパラメータを抽出する。このパラメータはキーワードである可能性が高い領域を中心に分布するため、その局所的なピークを各分布内で求める。これら投票度数のピークの時間成分を中心に、前後に参照データの時間長の幅を持つ時間領域をキーワード候補領域として保存する。これに基づき、次に示す DP マッチングを用いた処理を行なう。

〈4・2〉 連続 DP マッチングによるキーワード候補区間の推定 上で求められた各キーワード候補領域に対して、連続 DP マッチングによってキーワード候補が存在する時刻とその時間長を求める。参照データとして入力されたキーワードを標準パターン、キーワード候補領域を入力音声として連続 DP を行なう。

キーワードすなわち標準パターンのフレーム数を N 、キーワード候補領域すなわち入力音声のフレーム数を L 、 $d(i, j)$ を入力音声の第 i フレームと標準パターンの第 j フレームのスペクトルのユークリッド距離、入力音声のあるフレーム(以下に述べる方法により自動的に定まる)から第 i フレ-

図 8 DP マッチングにおける傾斜制限
Fig. 8. Slant constraint for DP matching

ムまでと標準パターンの第 1 フレームから第 j フレームまでの累積距離の最小値を $g(i, j)$ とする。図 8 に示す傾斜制限を用いると、 $g(i, j)$ は次の漸化式によって計算できる。

$$g(i, j) = \min \left\{ \begin{array}{l} g(i-2, j-1) \\ g(i-1, j-1) \\ g(i-1, j-2) + d(i, j-1) \end{array} \right\} + d(i, j) \quad (2 \leq i \leq L; 2 \leq j \leq N) \dots\dots\dots (7)$$

ただし境界条件として次式を用いる。

$$\left. \begin{array}{l} g(0, j) = \infty \quad (1 \leq j \leq N) \\ g(1, j) = \infty \quad (2 \leq j \leq N) \\ g(i, 0) = \infty \quad (0 \leq i \leq L) \\ g(i, 1) = d(i, 1) \quad (1 \leq i \leq L) \end{array} \right\} \dots\dots\dots (8)$$

さらに式 (7) の最小値が右辺の第 k 行目 ($k = 1, 2, 3$) の式によって達成されたとき、これを $backpointer(i, j)$ として記憶しておく。

式 (7) の漸化式と式 (8) の境界条件を用いて $g(i, N)$ ($g \leq i \leq L$) を計算し、それらの最小値を与える i の値を I とする。 $backpointer(I, N)$ の値から順次 $backpointer$ の値をさかのぼることにより、 $g(I, N)$ を与えるマッチングの経路が判明し、マッチングの開始フレーム I' が求まる。これによりキーワード候補の時間長 M は $M = I - I'$ となり、対象データのフレーム I, I' の区間に参照データの M/N 倍の長さでキーワード候補が存在することになる。

〈4・3〉 二段目の一般化 Hough 変換 上で求められたキーワードの候補区間に対して再び一般化 Hough 変換によってパラメータ空間に投票を行なう。各キーワード候補は参照データに対して M/N 倍の長さを持つことが分かっている。R-table を構成する各エントリベクトル $\vec{r}$ が

$$\vec{r} = \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix}, \quad \begin{array}{l} r_x: \text{時間成分} \\ r_y: \text{周波数成分} \\ r_z: \text{パワー成分} \end{array} \dots\dots\dots (9)$$

であるとする、対象データの各点 (X_i, Y_i, Z_i) に関して投票する際に、パラメータ (x, y, z) を以下の式によって求め、点 (x, y, z) の投票度数を 1 増やす。

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} + \begin{pmatrix} \frac{M}{N} \cdot r_x \\ r_y \\ r_z \end{pmatrix} \dots\dots\dots (10)$$

すべての対象データ点に対して上の投票が終われば、3.と同様にして投票度数の高いパラメータを抽出する。

これにより、入力されたキーワードの時間長と検索対象データに存在するキーワードの時間長の違いによる投票誤差を減らすことができる。

5. 実 験

〈5・1〉 実験条件 4.で述べた方法を用いてワードスポッティングの実験を行なった。検索したいキーワードとしては男性話者1名が発話した25単語の発話音声を用いた。これらの単語の平均発話時間長は619.3msecである。スポッティングの対象となる音声には音響学会連続音声データベースを使用し、男性話者30名(キーワードの発声者とは異なる)が発話した750文を対象文とした。付録にキーワードと検索対象文を示した。この対象文の継続時間長の合計は0.911時間である。また、キーワードの発声者と対象文の発声者が同一の場合の性能を調べるための実験も行なった。この場合はキーワードの発話者が発声した25文、継続時間長73.5secを使用した。

これらの音声から得られるスペクトル時系列の分析条件を表1に示す。そしてこのスペクトルをメルスケール128次元に変換したものを実験で使用した。

〈5・2〉 認識条件 二段目の一般化Hough変換を行なった後、パラメータ空間の中で閾値よりも大きい投票度数を持つパラメータの時間成分を、キーワードが存在する時刻として検出した。

また参照データの時間長に比例してR-tableのエントリベクトル数が増加するため、各投票セルの投票度数もそれに比例して増加する。したがってキーワードの時間長による閾値の変動を避けるため、閾値処理は投票度数をキーワードの時間長で正規化した値に対して行った。

検出時刻が正解時刻の400msec以内に存在する場合を正解検出、それ以外を誤検出とした。

また音声データベースに出現するキーワードの正解時刻の決定は、音声波形の視察と聴取により行なった。

本手法では音声データから1回の処理につき1種類のキーワード(1つの形状モデル)を検出するため、検出率(%), 湧き出し誤り率(個/hour)は以下ようになる。

$$\text{検出率} = \frac{\text{正解検出数}}{\text{キーワード出現数}} \dots\dots\dots (11)$$

$$\text{湧き出し誤り率} = \frac{\text{検出数} - \text{正解検出数}}{\text{発話総時間}} \dots\dots (12)$$

〈5・3〉 実験結果とその考察 以上の条件でワードスポッティングを行なったときの25単語の平均の検出率と湧

表1 分析条件

Table 1. Parameters for speech analysis

標準化周波数	16 kHz
分析区間窓	ハミング窓
分析区間長	25.6 msec
FFT 長	512
分析周期	5 msec
高域強調	0.98
特徴量	20 次のLPC分析から求めたスペクトル

き出し誤り率の関係を図9に示す。ただし図9において(a), (b), (c), (d)はそれぞれ,

(a)周波数軸がメルスケール128次元であるスペクトル時系列に対して一般化Hough変換を行なった結果。

(b)キーワードと同一話者が発話したデータに対して一般化Hough変換を行なった結果。

(c)(a)で用いたスペクトル時系列に対してDPマッチングと二段の一般化Hough変換を適用した結果。

(d)(a)で用いたスペクトル時系列に対してDPマッチングのみでワードスポッティングを行なった結果。

を表している。

(a), (b)を比較してみると、キーワードと対象データの話者が同じ場合には湧き出し誤りが50%以上抑えられていることが分かる。これは本手法が音声のスペクトルの形状的な特徴を用いていることによると考えられる。また(c)から、DPマッチングを組み合わせることによってスポッティング精度が向上していることが分かる。

検出結果について観察してみると、正解の検出失敗については、キーワード長に比例して失敗する傾向が見られる。これは参照データが短くなるとR-tableを構成するエントリベクトル数が少なくなり、キーワード長の正規化に対する投票誤差の影響が大きくなり十分な投票度数が得られなかったためと考えられる。またキーワードの発話長の違い

図9 正解検出率と湧き出し誤り率
Fig. 9. Detect rate and false alarm

は実際には非線形の伸縮であるのに対し、今回の手法では、DP マッチングの結果からキーワードの発話長の違いが一般的な伸縮として投票を行なったというのも原因の一つと考えられる。次に湧き出し誤りについて考えると、実験結果の分析から、母音部分のスペクトルが投票結果に与える影響が大きく、同じような母音列をもつ区間に湧き出し誤りが多く見られることが分かった。

6. おわりに

一般化 Hough 変換と連続 DP マッチングを用いてキーワードスポッティングを行なう手法を提案し、キーワードスポッティングの精度の向上を図った。

今後の課題として、DP マッチングの結果からの伸縮情報を非線形に考慮した R-table の作成と投票による正解検出率の向上と、母音部分のスペクトルの類似から起こる湧き出し誤りの抑制、そしてキーワード長による投票誤差を考慮した投票方式の検討などがある。

(平成 10 年 3 月 25 日受付, 同 10 年 8 月 12 日再受付)

文 献

- (1) Wilpon, J.G., Lee, C.H. and Rabiner, L.R., "Application of hidden Markov models for recognition of a limited set of words in unconstrained speech", Proc. of ICASSP-89, pp.254-257 (1989)
- (2) Rohlicek, J.R., Russel, W., Roukos, S. and Gish, H., "Continuous hidden Markov modeling for speaker-independent word spotting", ibid, pp.627-630 (1989)
- (3) Kawahara, T., Munetsugu, T., Kitaoka, N. and Doshita, S., "Keyword and phrase spotting with heuristic language model", Proc. of ICSLP-94, pp.815-819 (1994)
- (4) Kogure, T. and Katagiri, S., "A novel spotting-based approach to continuous speech recognition: Minimum error classification of keyword-sequences", J. Acoust. Soc. Jpn., Vol. E16, No. 3, pp.147-157 (1995)
- (5) Christiansen, R.W. and Rushforth, C.K., "Detecting and locating keywords in continuous speech using linear predictive coding", IEEE Trans. on Acoust., Speech, and Signal Processing, Vol. ASSP-25, No. 5, pp.361-367 (1977)
- (6) R.O. Duda and P.E. Hart, "Use of the Hough transformation to detect lines and curves in pictures", Communications of ACM, Vol. 15, pp.11-15 (1972)
- (7) D.H. Ballard, "Generalizing Hough Transformation to Detect Arbitrary Shapes", Pattern Recognition, Vol. 13, No. 2, pp.111-122 (1981)
- (8) 和田, 松山, "Hough 変換に基づく図形検出法の新展開", 情報処理, Vol. 36, No. 3, pp.253-263 (1995-3)
- (9) 櫻井, 有木, "キーワードスポッティングによるニュース音声の索引付けと分類", 信学技報, SP96-66, pp.37-44 (1996-11)
- (10) 岩橋, 山下, 溝口, "キーワードスポッティングにおけるピッチバタンの利用", 信学技報, SP96-106, pp.79-85 (1997-01)
- (11) 杉野, 小林, 新美, "一般化 Hough 変換を用いたワードスポッティング", 信学技報, SP97-23, pp.39-44 (1997-06)
- (12) 杉野, 小林, 新美, "一般化 Hough 変換を用いたワードスポッティングの精度向上", 信学技報, SP97-58, pp.7-12 (1997-11)
- (13) 新美 康永, 「音声認識」, 共立出版 (1979)

付 録

キーワード

現実, ニューヨーク, パソコン, 給付水準, 救助作業, 百歩, 代々木駅, 健康器具, ファミリーレストラン, ゆっくり, 研究者, 謙虚, ヨーロッパ, プロ野球ブーム, お客さん, アフリカ人, 個別文化, 一軒家, 厳しい, フィリピン,

非常口, 落ち度, 石組み, 完璧主義, 生き物

検索対象文

- あらゆる現実をすべて自分のほうへねじ曲げたのだ。
- 一週間ばかりニューヨーク取材した。
- テレビゲームやパソコンでゲームをして遊ぶ。
- 物価の変動を考慮して給付水準を決める必要がある。
- 救急車が十分に動けず救助作業が遅れている。
- 言論の自由は一步譲れば百歩も千歩も攻めこまれる。
- 会場の周辺には原宿駅や代々木駅もあるしちょっと歩けば新宿御苑駅もある。
- 老人ホームの場合は健康器具やひざ掛けだ。
- ちょっと遅い昼食をとるためファミリーレストランに入ったのです。
- 嬉しいはずがゆっくり寝てもらえない。
- 自然の研究者は自然をねじ伏せようとしてはいけない。
- おごりを捨て謙虚な姿勢を取り戻さねば冬は過ごせない。
- 先だつてごく短期間だが久方ぶりにヨーロッパへ行った。
- しかしこのプロ野球ブームも永久に続くとは限らぬ。
- お客さんはそれじゃ練習さえすればだれにでもできるんじゃないかなって考え始めるよ。
- アフリカ人は実に巧みにびゅんとつばを吐く。
- 前者を普遍文化と呼び後者を個別文化と呼ぶことにする。
- 叔父さんは岬の一軒家に独りぼっちで住んでいた。
- 立春がすぎても厳しい寒さの日々が続く。
- 大昔のフィリピンには豊かな土地があった。
- 旅館やホテルに着くと非常口を尋ねる。
- やるべきことはやっておらんら落ち度はない。
- 私は上着を脱ぎ石組みの上に両手をついてうつおせになった。
- 外人サンは完璧主義である。
- 人間とは微妙で複雑な生き物である。

杉野 浩 司 (非会員) 1970 年生。1993 年京都工芸繊維大学電子情報工学科卒業。1995 年同大学院博士前期課程修了。1998 年 3 月同大学院博士後期課程単位取得退学。1998 年大阪メディアポート (株) に入社。

新美 康 永 (非会員) 1938 年生。1962 年京都大学工学部電子工学科卒業。1964 年同大学院修士課程修了。同年同大学工学部助手。1970 年京都工芸繊維大学工学部助教授。現在、同大学教授。工学博士。この間、シミュレーション言語、音声情報処理、自然言語処理、人工知能などの研究に従事。