

2005年度卒業論文

動画像の時間軸領域における圧縮と
その評価方法に関する考察

法政大学

工学部 情報電気電子工学科

学籍番号：02d3006

氏名：安沢 昭伸

指導教授：斎藤 兆古

目次

第1章 序論	3
第2章 フーリエ変換の基礎	4
2.1 概要	4
2.2 関数の直交性と線形性	6
2.3 離散値系フーリエ変換	7
2.4 まとめ	9
第3章 K-L 情報量	10
3.1 概要	10
3.2 K-L 情報量の定義	10
3.3 K-L 情報量の性質	11
3.4 画像への適用	11
第4章 静止画像の圧縮	13
4.1 概要	13
4.2 領域法を用いた圧縮	13
4.3 まとめ	15

第 5 章	動画像の圧縮	1 6
5.1	概要	1 6
5.2	三次元変換による圧縮	1 6
5.3	各画素による圧縮	1 7
5.4	まとめ	1 8
第 6 章	まとめ	1 9
第 7 章	参考文献	2 0
付録	mathmatica によるコード	2 1

第1章 序論

今日、コンピュータグラフィクスは極めて豊富な情報を有するため、高品質なグラフィックデータをハンドリングするには必然的に比較的高度なハードウェアを必要とする。このような現状を踏まえ、コンピュータグラフィクスのもつ本質的な情報を抽出する手段として様々な映像情報処理技術が提案されている。

ここではフーリエ変換を用いてコンピュータグラフィクスの豊富な情報から本質的な情報を抽出し、如何にしてハードウェアに対して負担の少ない信号量へ圧縮するかについて検討する。また現在、人間の視覚に対して説得力のある画像評価方法が存在しない。よって本論文では新しい画像評価方法を導入している。具体的には情報量の概念を導入し、画像を確率分布として扱い、その線形空間的意味の距離を計算することによって評価している。

第2章 フーリエ変換の基礎

2.1 概要

古典的な関数変換の目的は、解析的に扱いにくい関数系を解析的に扱いやすい関数系へ変換することである。例えば、ラプラス(Laplace)変換は微積分演算を単純な掛け算や割り算へ化す変換である。また、関数系の変換とは、ある関数、例えば時間変化する関数を解析が容易な周期関数の和で表現することにも使われる。具体的な例としてフーリエ(Fourier)変換を考えよう。フーリエ変換は解析的に扱えない関数を解析的に扱える角周波数の異なる正弦波と余弦波の和で表現する変換である。換言すれば、フーリエ変換は解析的に扱いにくい関数系を解析的に扱いやすい関数系へ分解する変換と考えてもよい。

いま、ある任意の時間 t をパラメータとする関数 $f(t)$ を一定値 a_0 、正弦波および余弦波の和で表現できるとしよう。すなわち、 ω を角周波数として、

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin(i\omega t) + \sum_{i=1}^{\infty} b_i \cos(i\omega t) \quad (1)$$

と仮定する。

問題は(1)式の係数、 a_0 、 a_i 、 b_i を決める方法である。いま、関数 $f(t)$ が時間 $t=0$ から $t=T$ の区間で(1)式の係数を決めることを考えれば、(1)式は、 $\omega = 2\pi/T$ であるから、

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i \frac{2\pi}{T} t\right) \quad (2)$$

とも書ける。

(2)式の両辺を時間 $t=0$ から $t=T$ の区間について積分すると、

$$\begin{aligned}
& \int_0^T f(t) dt \\
&= \int_0^T \left\{ a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i \frac{2\pi}{T} t\right) \right\} dt \\
&= a_0
\end{aligned} \tag{3}$$

が成り立つ。

したがって、定数項 a_0 は次のようになる。

$$a_0 = \frac{1}{T} \int_0^T f(t) dt \tag{4}$$

次に、(2)式の両辺に正弦波 $\sin[j(2\pi/T)t]$ 、 $j=1,2,3,\dots$ を掛け算し、時間 $t=0$ から $t=T$ の区間について積分する。

$$\begin{aligned}
& \int_0^T f(t) \sin\left(j \frac{2\pi}{T} t\right) dt \\
&= \int_0^T \left\{ a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i \frac{2\pi}{T} t\right) \right\} \times \sin\left(j \frac{2\pi}{T} t\right) dt \\
&= a_i \frac{2}{T}, i = j \text{ or } 0, i \neq j
\end{aligned} \tag{5}$$

(5)式から係数 a_i は

$$a_i = \frac{2}{T} \int_0^T f(t) \sin\left(i \frac{2\pi}{T} t\right) dt \tag{6}$$

として得られる。

同様に、(2)式の両辺に余弦波 $\cos[j(2\pi/T)t]$ 、 $j=1,2,3,\dots$ を掛け算し、時間 $t=0$ から $t=T$ の区間について積分することで

$$b_i = \frac{2}{T} \int_0^T f(t) \cos\left(i \frac{2\pi}{T} t\right) dt \tag{7}$$

として係数 b_i が得られる。

したがって、関数 $f(t)$ は、区間 $t=0$ から $t=T$ で、

$$f(t) = \frac{1}{T} \int_0^T f(t) dt + \sum_{i=1}^{\infty} \left\{ \frac{2}{T} \int_0^T f(t) \sin\left(i \frac{2\pi}{T} t\right) dt \right\} \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} \left\{ \frac{2}{T} \int_0^T f(t) \cos\left(i \frac{2\pi}{T} t\right) dt \right\} \cos\left(i \frac{2\pi}{T} t\right) \quad (8)$$

と書ける。これがいわゆるフーリエ変換の原型となるフーリエ級数であり、左辺の関数 $f(t)$ を右辺の計算が簡単な定数項と三角関数の和に変換している。

2.2 関数の直交性と線形性

フーリエ級数の考え方の中に重要な関数間で成り立つ性質、すなわち、関数の直交性(orthogonality)が使われている。

まず(4)式の係数 a が計算される過程を考えよう。(3)式は(2)式の両辺に定数値1を掛け算し積分する演算である。このとき、

$$\int_0^T 1 \cdot \sin\left(i \frac{2\pi}{T} t\right) dt = 0, \int_0^T 1 \cdot \cos\left(i \frac{2\pi}{T} t\right) dt = 0, \quad (9)$$

$i = 1, 2, \dots$

の関係が成り立つために、(4)式の係数 a_0 がうまく計算できた。この関係を、定数値1と正弦波 $\sin[j(2\pi/T)t]$ 、および余弦波 $\cos[j(2\pi/T)t]$ 間の直交性と呼ぶ。

同様に、(6),(7)式で計算される a_i 、 b_i は

$$\int_0^T \sin\left(i \frac{2\pi}{T} t\right) \sin\left(j \frac{2\pi}{T} t\right) dt = 0,$$
$$\int_0^T \cos\left(i \frac{2\pi}{T} t\right) \cos\left(j \frac{2\pi}{T} t\right) dt = 0, \quad (10)$$

$i \neq j$

なる直交性が成り立つことに基づいている。

(9), (10)式から、直交性とは与えられた関数を他の関数の和で表現しようとする場合、和となる関数の大きさ(係数)を一意的に決める条件であることがわかる。言い換えれば、与えられた関数を他の関数の和で表現できる条件である。

さて、ある任意の時間 t をパラメータとする関数 $f(t)$ は、区間 $t=0$ から $t=T$ で、一定値、正弦波および余弦波の和で表現できることがわかった。この変換は、一定値、正弦波および余弦波間で直交性が成り立つことが条件であった。この結果に至る過程を考えてみると、まず、展開される関数の和でもとの関数が表現されるとする大前提があることに気づくであろう。ある関数が他の関数の和で表現できる性質を線形性と呼ぶ。では、(6)式で計算される正弦波の係数 a_i が

定数 c_i の n 個の和で表現されるとしよう。すなわち、

$$a_i = c_i + c_i + \dots = nc_i \quad (11)$$

である。(11)式から、和を前提とする系では比例関係が成り立つことを意味することがわかる。すなわち、線形系とは比例関係が成立する系である。

2.3 離散値系フーリエ変換

離散値系(discrete)とはどのような系であろうか。離散値なる用語が生まれたのは計数型計算機を用いて数値計算を行う場合、連続関数をサンプリングして得られる数値で代表したことに起因する。例えば、関数 $f(t)$ を時間 $t=0$ から $t=T$ の区間で、 t ごとに n 個サンプリングして離散値系で表すと、

$$F = [f(\Delta t), f(2\Delta t), f(3\Delta t), \dots, f(n\Delta t)]^T \quad (12)$$

なるベクトルとなる。

同様に、(1)式を離散値系で書けば、

$$\begin{pmatrix} f(\Delta t) \\ f(2\Delta t) \\ f(3\Delta t) \\ \vdots \\ f(n\Delta t) \end{pmatrix} = a_0 \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \begin{pmatrix} \sin\left(\frac{2n}{T}\Delta t\right) & \sin\left(2\frac{2n}{T}\Delta t\right) & \sin\left(3\frac{2n}{T}\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}\Delta t\right) \\ \sin\left(\frac{2n}{T}2\Delta t\right) & \sin\left(2\frac{2n}{T}2\Delta t\right) & \sin\left(3\frac{2n}{T}2\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}2\Delta t\right) \\ \sin\left(\frac{2n}{T}3\Delta t\right) & \sin\left(2\frac{2n}{T}3\Delta t\right) & \sin\left(3\frac{2n}{T}3\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}3\Delta t\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sin\left(\frac{2n}{T}n\Delta t\right) & \sin\left(2\frac{2n}{T}n\Delta t\right) & \sin\left(3\frac{2n}{T}n\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}n\Delta t\right) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} \cos\left(\frac{2n}{T}\Delta t\right) & \cos\left(2\frac{2n}{T}\Delta t\right) & \cos\left(3\frac{2n}{T}\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}\Delta t\right) \\ \cos\left(\frac{2n}{T}2\Delta t\right) & \cos\left(2\frac{2n}{T}2\Delta t\right) & \cos\left(3\frac{2n}{T}2\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}2\Delta t\right) \\ \cos\left(\frac{2n}{T}3\Delta t\right) & \cos\left(2\frac{2n}{T}3\Delta t\right) & \cos\left(3\frac{2n}{T}3\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}3\Delta t\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \cos\left(\frac{2n}{T}n\Delta t\right) & \cos\left(2\frac{2n}{T}n\Delta t\right) & \cos\left(3\frac{2n}{T}n\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}n\Delta t\right) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \quad (13)$$

or

$$\mathbf{F} = a_0 \mathbf{I} + \mathbf{A}\mathbf{S} + \mathbf{B}\mathbf{C}$$

となる。ただし、 \mathbf{I} は n 次の単位列ベクトルである。

また、(13)式右辺の係数 a 、ベクトル \mathbf{S} と \mathbf{C} は、

$$a_0 = \frac{1}{n} \mathbf{I}^T \cdot \mathbf{F}, \mathbf{S} = \frac{2}{n} \mathbf{A}^T \cdot \mathbf{F}, \mathbf{C} = \frac{2}{n} \mathbf{B}^T \cdot \mathbf{F} \quad (14)$$

で与えられる。

さて、離散値系でフーリエ係数を計算する過程で、(13)式を

$$\begin{aligned}\mathbf{F} &= a_0 \frac{1}{\sqrt{n}} \mathbf{I} + \sqrt{\frac{2}{n}} \mathbf{A} \mathbf{S} + \sqrt{\frac{2}{n}} \mathbf{B} \mathbf{C} \\ &= a_0 \mathbf{I}^T + \mathbf{A}' \mathbf{S} + \mathbf{B}' \mathbf{C}\end{aligned}\tag{15}$$

と書き直すと、

$$a_0 = \frac{1}{\sqrt{n}} \mathbf{I}' \mathbf{T} \cdot \mathbf{F}, \mathbf{S} = \sqrt{\frac{2}{n}} \mathbf{A}'^T \cdot \mathbf{F}, \mathbf{C} = \sqrt{\frac{2}{n}} \mathbf{B}'^T \mathbf{F}\tag{16}$$

の関係が成り立つ。 \mathbf{I} は n 次の単位行列である。この結果は、離散値系で正弦波や余弦波のフーリエ係数を求める場合、(16)式の係数行列 \mathbf{A}' 、 \mathbf{B}' の逆行列がそれぞれの転置行列で与えられることを意味する。言い換えれば、変換行列の逆行列が変換行列の転置行列で与えられることでフーリエ係数の直交性が満足される。

2.4 まとめ

本章では、フーリエ解析の考え方から出発した。フーリエ級数から、関数の変換、直交性、さらに線形性の概念を説明した。つぎに、連続関数をサンプリングして得られる離散値系について、フーリエ級数の離散値形を導き、離散値系における直交性や線形変換の概念を紹介した。

3章 K-L 情報量

3.1 概要

現在画像の評価として主に相関係数による方法が一般的である。しかし相関係数による評価は本当に正しいのだろうか、という疑問から新しい評価方法として K-L 情報量を提案する。K-L 情報量とはカルバック・ライブラー情報量のことであり Shannon が提案した情報量を一般化した概念である。情報量は事象の起こる確率によって表現される情報の量であり、確率分布と密接な関係がある。また情報理論における相互情報量に近い概念である。つまり、二つの確率分布の独立性、従属性の指標となる。ここでは画像を確率分布とみなし K-L 情報量を計算する。

3.2 K-L 情報量の定義

未知の確率分布関数 $G(x)$ に従って観測された n 個のデータを $X_n = \{x_1, x_2, \dots, x_n\}$ とする。データを発生するこの確率分布関数 $G(x)$ を以下では真の分布と呼ぶことにする。これに対して我々が想定したモデルを $F(x)$ とする。確率分布関数 $G(x)$ および $F(x)$ が、それぞれ密度関数 $g(x)$ および $f(x)$ を持つ場合は連続モデル (連続分布モデル) という。一方、 $g(x)$ および $f(x)$ が有限または可算無限個の離散点 $\{x_1, x_2, \dots, x_k, \dots\}$ に対して、次のような事象 $\{\omega; X(\omega) = x_i\}$

の確率

$$\begin{aligned} g_i &= g(x_i) = \Pr(\{\omega; X(\omega) = x_i\}), \\ f_i &= f(x_i) = \Pr(\{\omega; X(\omega) = x_i\}), \quad i = 1, 2, \dots \end{aligned} \quad (3.1)$$

で表される場合は、離散モデル (離散分布モデル) という。

このときモデル $f(x)$ のよさを真のモデル $g(x)$ との確率分布としての近さによって評価するものとする。この近さを測る尺度として、次のカルバック・ライブラー情報量 (Kullback・Leibler) K-L 情報量を用いることを提案する。

$$I(G; F) = E_G \left[\log \frac{G(x)}{F(x)} \right] \quad (3.2)$$

ここで E_G は確率分布 G に関する期待値を示す。

確率分布関数が密度関数 $g(x)$ と $f(x)$ をもつ連続モデルの場合には K - L 情報量は

$$I(g; f) = \int_{-\infty}^{\infty} \log\left\{\frac{g(x)}{f(x)}\right\} g(x) dx \quad (3.3)$$

と表される。一方、確率が $\{g(x_i); i = 1, 2, \dots\}, \{f(x_i); i = 1, 2, \dots\}$ で与えられる離散モデルの場合には

$$I\{g; f\} = \sum_{i=1}^{\infty} g(x_i) \log \frac{g(x_i)}{f(x_i)} \quad (3.4)$$

と表される。以上をまとめると次のようになる。

$$I(g; f) = \int \log\left\{\frac{g(x)}{f(x)}\right\} dG(x) = \begin{cases} \int_{-\infty}^{\infty} \log\left\{\frac{g(x)}{f(x)}\right\} g(x) dx & \text{連続モデル} \\ \sum_{i=1}^{\infty} g(x_i) \log \frac{g(x_i)}{f(x_i)} & \text{離散モデル} \end{cases}$$

この式において連続モデルの $g(x)$ は密度関数だが、離散モデルの $g(x_i)$ は確率であることに注意してほしい。

3.3 K-L 情報量の性質

K-L 情報量に関しては、次のような性質がある。

- (1) $I(g; f) \geq 0$
- (2) $I(g; f) = 0 \Leftrightarrow g(x) = f(x)$

この性質から、K-L 情報量の値が小さいほど、モデル $f(x)$ は $g(x)$ に近いと考えることができる。

3.4 画像への適用

離散モデルにおける $g(x)$ や $f(x)$ は確率分布である。確率分布には次のような性質がある。

$$g(x) \neq 0,$$

$$\sum_{i=1}^n g(x_i) = 1,$$

また、このような性質をもつ分布は確率分布とすることができる。したがって画像の数値データをこの性質に適応するように正規化する。この処理を行うことによって画像データを確率分布とし、K-L 情報量で評価可能とするのである。しかし画像データには0の値が存在する。この問題に対して0の部分に他の数値に比べて非常に小さい値を加えることにした。よって K-L 情報量は正確な値ではない。だが、加える数値をある一定値以下にすると K-L 情報量の値が収束するため実用上問題ないと思われる。

4章 静止画像の圧縮

4.1 概要

フーリエ変換によるデータ圧縮の原理は、もとのデータの特徴を表す絶対値の大きなスペクトラムのみを残し他をゼロにすることにある。サンプルデータは、 256×256 ビットイメージの画像を使う。

イメージデータの圧縮は、絶対値の大きいスペクトラムを残す場合、ある閾値より大きいスペクトラムのみを残す閾値法と、スペクトラムの特定部分（ふつう、1行1列の要素から順番に）を残す領域法が考えられる。

前者はスペクトラムが拡散しても良好な再現性が期待できる。しかし、順番がランダムになるから要素位置を記憶しなければならないため、圧縮効果が低いのが欠点である。さらに演算に時間がかかるためあまり実用的ではない。

後者は、必ずしも絶対値の大きなスペクトラムが特定部分に集中するとは限らないため、すべてのデータには適用できないが、圧縮効果はきわめて大きくなる。また、変化率の少ない低周波成分で平均的に原データを再現するため、高周波ノイズが低減される。ほとんどの画像は低周波領域にスペクトラムが集まるため非常に有効な手段である。他にも前者と比べて演算に時間がかからない、というメリットもある。

本論文では動画も扱うためデータの量が膨大である。したがって領域法のみ扱うこととする。

4.2 領域法による圧縮

まず、本実験で用いるサンプルイメージを図1に示す。イメージデータは 256×256 の画像を使う。



図1 サンプルイメージデータ

次にサンプルイメージを赤 (R)、緑 (G)、青 (B) の3つの要素に分け、それぞれに対しフーリエスペクトラムを計算する。その結果を図2に示す。

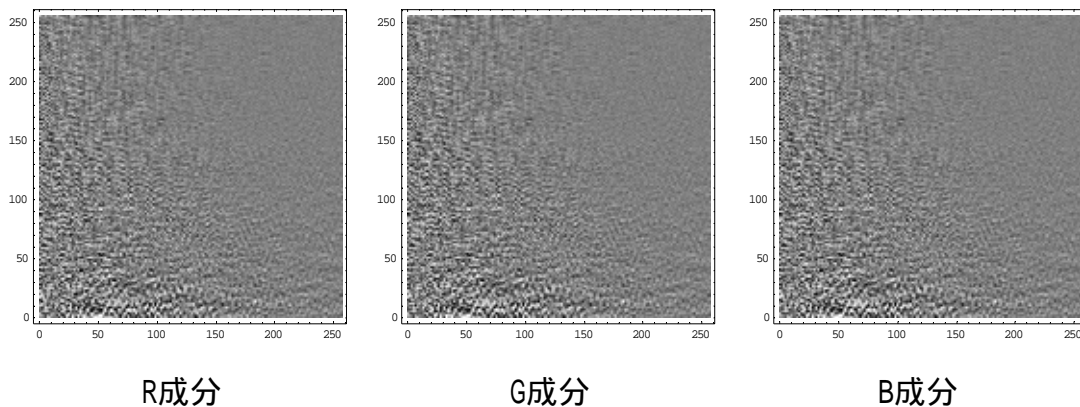


図2 各色成分におけるフーリエスペクトラム

このスペクトラムでは原点が最も絶対値が大きく、原点から距離が離れるにつれ小さくなっている。領域法を適用して $26 \times 26 = 676$ 個の領域だけを残し、他をゼロとすることで約1%に圧縮する。同じ方法を用いて5%、25%に圧縮する。再現画像を図3に示す。



1 %



5 %



2 5 %

図3 再現画像

このときの相関係数とK-L情報量は次のようになる。

圧縮率	相関係数	K-L情報量 ($\times 10^{-3}$)
1 %	0.92567	28.5194
5 %	0.961681	14.8038
2 5 %	0.985726	7.00131

4.3 まとめ

本章では領域法によるイメージデータの圧縮法について述べ、相関係数とK-L情報量の数値データを示した。相関係数ではすべて90%以上の再現性があるとなっているが、そのようには見えない。しかしK-L情報量についてみると、25%の画像は5%より2倍の再現性がある、となっている。これならば視覚的に考えると、相関係数よりも納得し得るであろう。また25%のデータを残すだけでもある程度の再現性は保たれることがわかる。

第5章 動画像の圧縮

5.1 概要

現在一般的に使用されている動画像の圧縮方法は次のようである。1フレームごとに前章のようにフレーム単位で圧縮を実行する。さらに基準画像を設け、その画像との差分（誤差）を保存する。場面が切り替わらなければ似たような画像の連続となるため差分は少なく、この方法は有効である。

本論文ではこれとは異なる2つの動画像の圧縮方法について述べる。基本的に前章で述べたフーリエ変換による領域法を動画に適用している。

2つの圧縮方法とはフーリエ変換を実行する対象の選択方法が異なる。縦、横、時間（フレーム数）軸について三次元フーリエ変換し集まったスペクトラムの特定部分を残す方法と、各画素について時間軸方向にフーリエ変換し時間軸の始点部分に集まったスペクトラムを残す方法である。前者は演算に時間がかかるが、フレーム数に対する依存性が低い。後者は演算時間は少なく、フレーム数が多いときは効率よく圧縮できる。しかしフレーム数が少ないとき、どの程度領域を保存するかを選択が難しい。

5.2 三次元フーリエ変換による圧縮

まず本実験で使用する画像を図1に示す。サンプルイメージデータの縦・横・フレーム数は120×160×64である。



図1 サンプルイメージデータ

次に三次元に対してフーリエ変換し、スペクトラムの縦 1/2、横 1/2、時間軸 1/2 の領域だけ残す。同じ処理をフレーム数に変化をつけて実行する。このときの相関係数と K-L 情報量は次のようになる。

フレーム数	相関係数	K-L 情報量 ($\times 10^{-3}$)
3 2	0.967521	17.7053
6 4	0.966132	18.2293
1 2 8	0.96519	19.3412

次に $120 \times 160 \times 64$ の画像を約 5 % に圧縮した画像を図 2 に示す。このとき相関係数 0.963081、K-L 情報量 23.077×10^{-3} である。

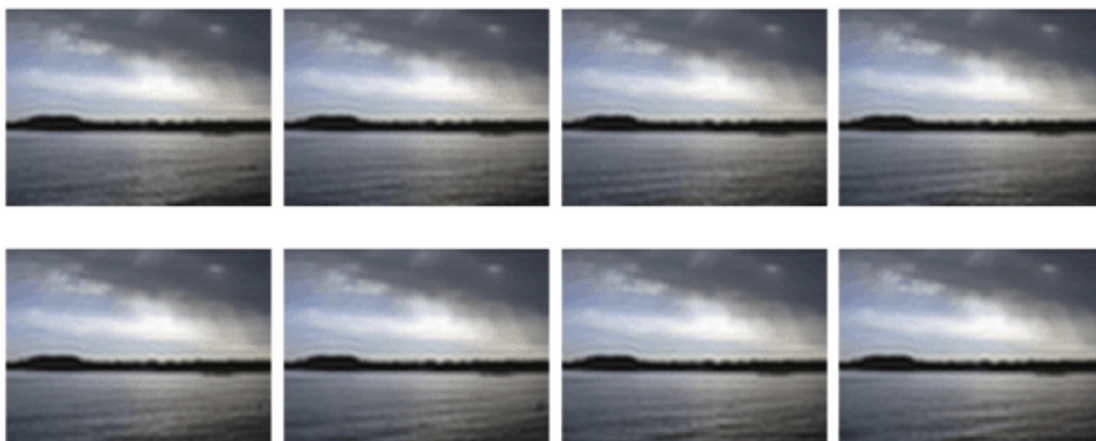


図 2 約 5 % に圧縮した画像

5.3 各画素に対するフーリエ変換における圧縮

サンプルイメージデータは 5 . 2 と同じものを使う。各画素に対してフーリエ変換し、スペクトラムの時間軸を 1/8 だけ残す。このとき縦と横のスペクトラムはそのまま残している。同じ処理をフレーム数に変化をつけて実行する。このときの相関係数と K-L 情報量は次のようになる。

フレーム数	相関係数	K-L 情報量 ($\times 10^{-3}$)
3 2	0.994151	4.91018
6 4	0.994739	5.55013
1 2 8	0.9952	7.48906

次に $120 \times 160 \times 64$ の画像を各画素について約 5% に圧縮した画像を図 3 に示す。このときの相関係数は 0.991681、K-L 情報量は 4.4292×10^{-3} である。

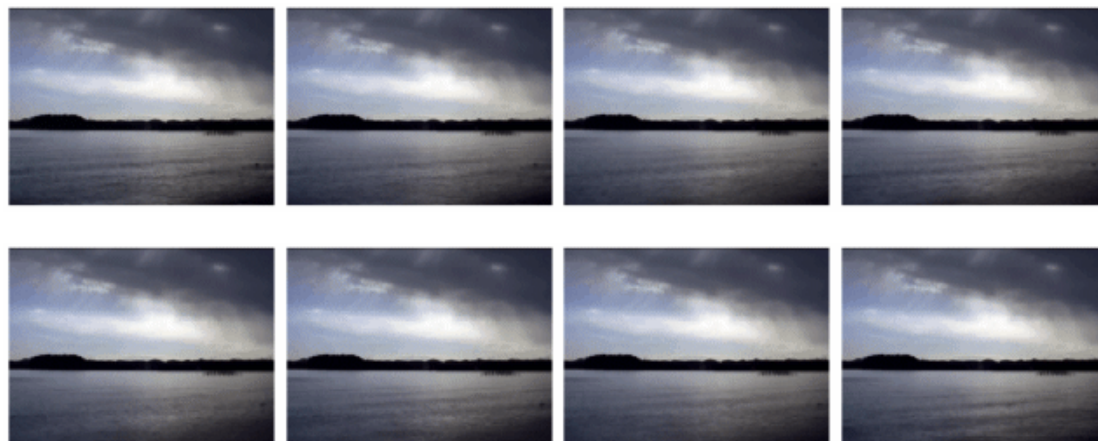


図 3 約 5% に圧縮した画像

5.4 まとめ

本章では動画の圧縮方法について述べた。三次元フーリエ変換における圧縮は相関係数または K-L 情報量どちらで評価しても、安定した圧縮を行うことができるといえる。他方、各画素に対するフーリエ変換における圧縮は相関係数で評価するとフレーム数が多いほうが良いことになるが、K-L 情報量で評価するとフレーム数が少ないほうが良いことになる。残すスペクトラムに対して適切な領域を設定する必要がある。しかしフレーム数を少なくして圧縮できるというメリットがあるため利用できる余地があると考えられる。

圧縮後について三次元フーリエ変換における圧縮は、本例ではわかりにくいですが、輪郭がぼやけているが動きはなめらかである。一方、各画素に対するフーリエ変換における圧縮は輪郭がはっきりと残っているが、動きが悪い。それぞれに特徴があり、上手く組み合わせれば最適な圧縮が可能だと考えられる。

第6章 まとめ

本論文では新しい評価方法として K-L 情報量について述べた。評価静止画において、相関係数による評価と K-L 情報量による評価において傾向による違いはなかった。しかし動画では評価方法による違いが存在した。相関係数では確認できない情報を K-L 情報量は抽出できることを示した。

また静止画圧縮の拡張として動画圧縮を述べたが、フーリエ変換の対象によって異なる圧縮方法を示した。三次元フーリエ変換は安定した圧縮が可能であるが、各画素フーリエ変換はフレーム数によって再現性の変化が存在する。したがって再現性において未だ開発・検討の余地がある。

一例としてウェーブレット変換を使用することが考えられる。縦軸、横軸方向はウェーブレット変換により圧縮し、時間軸方向はフーリエ変換を使用する、など組み合わせてみると圧縮効率や再現性が向上するかもしれない。これについては今後の研究に期待したいと思う。

第7章 参考文献

- [1] Kullback,S. and Leibler,R.A. (1951):"On information and sufficiency",
Ann.Math.Statics., 22, 79-86
- [2] 小西 貞則、北川 源四郎著、"情報量規準"、朝倉書店
- [3] 斎藤兆古著、"Mathematicaによるウェーブレット変換"、朝倉書店
- [4] 斎藤兆古著、"ウェーブレット変換の基礎と応用 Mathematicaで学ぶ"、
朝倉書店

付録 mathematica によるコード

本論文で使用したコードをここに示す。静止画は 1/2、三次元圧縮と各画素圧縮は 1/8 に圧縮している。

K-L 情報量

```
kl[x_,y_]:=Module[{dim,aa,bb},
  aa=x;bb=y;
  aa=(aa-Min[aa])/(Max[aa]-Min[aa]);bb=(bb-Min[bb])/(Max[bb]-Min[bb]);
  aa=aa+0.00001;
  bb=bb+0.00001;
  aa=aa/Apply[Plus,aa];
  bb=bb/Apply[Plus,bb];
  dim=Dimensions[aa];
  Apply[Plus,Table[aa[[i]]*Log[aa[[i]]/bb[[i]]],{i,dim[[1]]}]]*1000;
```

静止画の圧縮

```
assyuku[x_]:=Module[{aa,bb,dim},
  dim=Dimensions[x];
  aa=Table[FourierCos[x[[All,All,n]]],{n,dim[[3]]};
  bb=Table[0.,{i,dim[[3]]},{j,dim[[1]]},{k,dim[[2]]};
  Do[bb[[i,j,k]]=aa[[i,j,k]],{i,dim[[3]],{j,Round[0.5*dim[[1]]]},{k,Round[0.5*dim[[2]]]};
  aa=Table[FourierCos[bb[[n,All,All]]],{n,dim[[3]]};
  Transpose[aa,{3,1,2}];
```

三次元フーリエ変換による圧縮

```
dim3[x_]:=Module[{aa,dim,bb},
  dim=Dimensions[x];
  aa=Table[FourierCos[x[[All,n,All,All]]],{n,3}];
  bb=Table[0.,{k,dim[[2]]},{n,dim[[1]]},{i,dim[[3]]},{j
    ,dim[[4]]}];
  Do[bb[[k,n,i,j]]=aa[[k,n,i,j]],{{n,Round[0.5*dim[[1]]
    ]},{i,Round[0.5*dim[[3]]]},{j,Round[0.5*dim[[4]]]},{k
    ,dim[[2]]}];
  aa=Table[FourierCos[bb[[n,All,All,All]]],{n,3}];
  Transpose[aa,{2,1,3,4}];
```

各画素に対するフーリエ変換による圧縮

```
kakugaso[x_]:=Module[{aa,bb,dim},
  dim=Dimensions[x];
  aa=Table[FourierCos[x[[All,i,j,k]]],{i,3},{j,dim[[3]]}
    ,{k,dim[[4]]}];
  Do[aa[[All,All,All,n]]=0.,
    {n,Round[0.125*dim[[1]]],dim[[1]]}];
  bb=Table[FourierCos[aa[[i,j,k,All]]],{i,3},{j,dim[[3]]
    },{k,dim[[4]]}];
  Transpose[bb,{2,3,4,1}];
```

2005 年度
卒業論文

逆問題的手法による画像認識と
その動画像認識への応用

法政大学
工学部 情報電気電子工学科

学籍番号 02d3001

氏名 青木 亮一

指導教授 斎藤 兆古

目次

1. 序論

2. 静止画像の固有パターン

- 2.1 画像の色成分
- 2.2 カラー画像の固有パターン
- 2.3 固有パターン

3. 固有パターンの一意性

- 3.1 データベース
- 3.2 線形システム方程式
- 3.3 解ベクトル
- 3.4 静止画像認識
- 3.5 GVSPM 法の導入
- 3.6 GVSPM 法の静止画像認識への応用

4. 合成画像法による動画像認識

- 4.1 合成画像法

5. 動画像認識の手話への応用

- 5.1 同一人物による手話の識別
- 5.2 数人による手話の識別
- 5.3 GVSPM 法の動画像認識への適用
- 5.4 クラスタ分類を利用した GVSPM 法の動画像認識への適用

6. まとめ

7. 参考文献

1. 序論

現在，人間の評価能力を前提として映像を計算機によって処理・加工する技術は比較的多く実用化されている。しかし、映像を認識する技術は完成度が低く、最終的には人間の視覚情報処理能力へ依存せざるを得ない。圧倒的な演算処理能力をもつ最新の計算機をもってしても、画像認識の分野においては人間の視覚情報処理能力には到底及ばない。映像を計算機によって認識させることは、計算機に人間と同等な視覚情報処理能力を与えることになり、現状の様々な手法では人間の視覚情報処理能力には匹敵し難い。

このような現状をふまえて、本論文では、電子的に計算機スクリーン上に可視化された映像から、画素数，配置，変形等に無関係となる映像の普遍量を抽出する概念を導入し、人間の視覚情報処理能力を計算機に与え、計算機自身によって映像を識別させることを試みる。

ここで我々が認識の対象とする映像は、人間の視覚によって認識し得る形態で表現された情報をいい、それらを総称して本論文では可視化情報と呼ぶ。可視化情報には、人間の視覚で捉えることが可能なカメラやビデオで撮影された画像等の映像がある。また、通常、人間の視覚情報処理機能で捉えることができない情報を映像化した赤外線映像や電子顕微鏡による映像等がある。さらに、人間の視覚で認識可能な形態へ変形された信号情報、例えば音声等を電子的にオシロスコープ等の機器で映像化した情報も可視化情報である。これらの様々な可視化情報の中で本論文では、カメラやビデオで撮影した画像を認識させることを試みる。

人間の脳の視聴覚情報機能は、眼球を通して得られた映像を処理することによって外界の情報を認識するが、その対象となる視覚情報には、一定の構成ルールに従ってコード化された文字や言語等のコード化視聴覚情報と、コード化されていない可視化情報、すなわち非コード化視聴覚情報がある。前述の可視化情報は、非コード化視覚情報を処理する人間の視覚情報処理能力を前提としたデータであり、人間は人間の持つ視覚情報処理能力によって何らかの普遍量を感じ認識しているものと考えられる。従って、計算機によって人間の視覚情報処理能力を実現し認識するためには、可視化情報から計算機で処理可能な何らかの普遍量を抽出する必要がある。

計算機のスクリーン上の静止画像は2次元平面上の画素(Pixel)で構成されている。各画素は可視光の波長によってそれぞれ赤、緑、及び青(以下それぞれ R 、 G 、 B と略記)成分の情報を持ち、数値として表される。さらに動画は静止画像を時系列上に一定の間隔で静止画像を表示させているものに他ならない。このため、画像はスクリーン上における画素のRGB成分の幾何学的配置で表現されるため、画像を構成する画素数や画像の2次元平面上の位置・角度などのハードウェアに依存する成分を持つ。このような性質を削除した画像の普遍量を本論文では画像の固有パターン(Eigen pattern)と呼び、この画像の固有パターンを用いて画像認識を試みる。

本論文で採用する画像認識手法は線形システム方程式の解を導くことに基づく方法である。しかしながら、得られた線形システム方程式は不適切なシステム方程式であるため、最小自乗法を用いて近似解を導く。換言すれば、本論文で採用する画像認識手法は、あらかじめ用意した画像データをデータベースとし、その中から誤差ノルムが最小となる解要素を求める方法である。しかし、最小自乗法を用いることにより画像認識の処理速度や画像認識率の向上という問題がある。その問題を解決するため、本論文では逆問題的手法による画像認識手法について述べる。

本論文の構成は、以下の通りである。第2節では、静止画像から固有パターンを抽出する方法について述べ、第3節では、抽出された固有パターンの一致評価法について述べ、最小自乗法を用いて、画像認識を行う。また固有パターンの一致性または一意性評価法に一般化ベクトルサンプルパターンマッチング法(GVSPM 法)を導入し、認識を行

い、最小自乗法と GVSPM 法の固有パターン的一致評価法における認識比較について述べる。第 4 節では、静止画像認識手法を動画認識手法に拡張する方法について述べる。第 5 節では、動画認識について述べ、静止画像認識と同様に、最小自乗法を用いて、動画認識を行う。その後、GVSPM 法を用いて、合成画像法による動画認識を行う。さらに、固有パターン間的一致評価方法にクラスター分類による考え方を導入し、合成画像法による動画認識を行い、最小自乗法と GVSPM 法とクラスター分類の固有パターン的一致評価法における認識比較について述べる。

2 . 静止画像の固有パターン

2.1 画像の色成分

計算機のスクリーン上の画像情報は2次元平面、すなわち x-y 直交座標上の画素で構成されており、また各画素は R, G, B 色成分で構成されている。そこで、m, n をそれぞれ x, y 直行座標系の x, y 方向の画素数とすれば、画像データを m×n 次のベクトルとして扱うことができる。ベクトルとして表現されたカラー画像データにおける第 k 番目の画素の R, G, B 成分を R_k, G_k, B_k とすると、画素は式(1)で表される。

$$\begin{aligned} image &\in pixel_{i,j} \\ pixel_{i,j} &\in R_k, G_k, B_k \\ i &= 1,2,3,..,m \quad j = 1,2,3,..,n \quad k = 1,2,3,..,m \times n \end{aligned} \tag{1}$$

2.2 カラー画像の固有ベクトル

カラー画像の固有パターンは、画像中の R 成分の強度値、G 成分の強度値、B 成分の強度値で与えられ、1 画像に対して 3 組得られる。R 成分分布を I_R とすると、

$$I_R = R_k \quad k = 1,2,3,..,m \times n \tag{2}$$

で与えられる。次に R 成分分布 I_R をダイナミックレンジ D の分布に正規化した正規化 R 成分分布 I_R^D を式(3)で定義する。 $Round[*]$ は括弧内の値を整数化する演算、 $Max[*]$ は括弧内の集合における最大値の要素を求める演算を示す。

$$\begin{aligned} I_R^D &\in Round \left[D \times \frac{R_k}{Max[I_R]} \right] \\ k &= 1,2,3,..,m \times n \end{aligned} \tag{3}$$

次に、式(3)で求められたダイナミックレンジ D の正規化 R 成分分布 I_R^D において 1 から D までの数値の出現頻度を求めヒストグラム化する。 $Count[* , p]$ は括弧内の集合における数値 p の個数を数える演算である。

$$E_R \in Count [I_R^D, p] \quad p = 1,2,3,..,D \tag{4}$$

式(4)で求められた E_R を R 成分固有ベクトルと言う。G 成分、B 成分についても同様にして固有ベクトルを算出する。

2.3 固有パターン

固有パターンを E とすると固有パターン E は R 成分固有パターン、G 成分固有パターン、B 成分固有パターンの 3 個の固有パターンの集合からなり、式(5)で与えられる。Fig.1 に静止画像の固有パターンの一例を示す。Fig.1 から、画像はそれぞれ独自の固有パターンを持つことが分かる。

$$E \in E_R, E_G, E_B \quad (5)$$

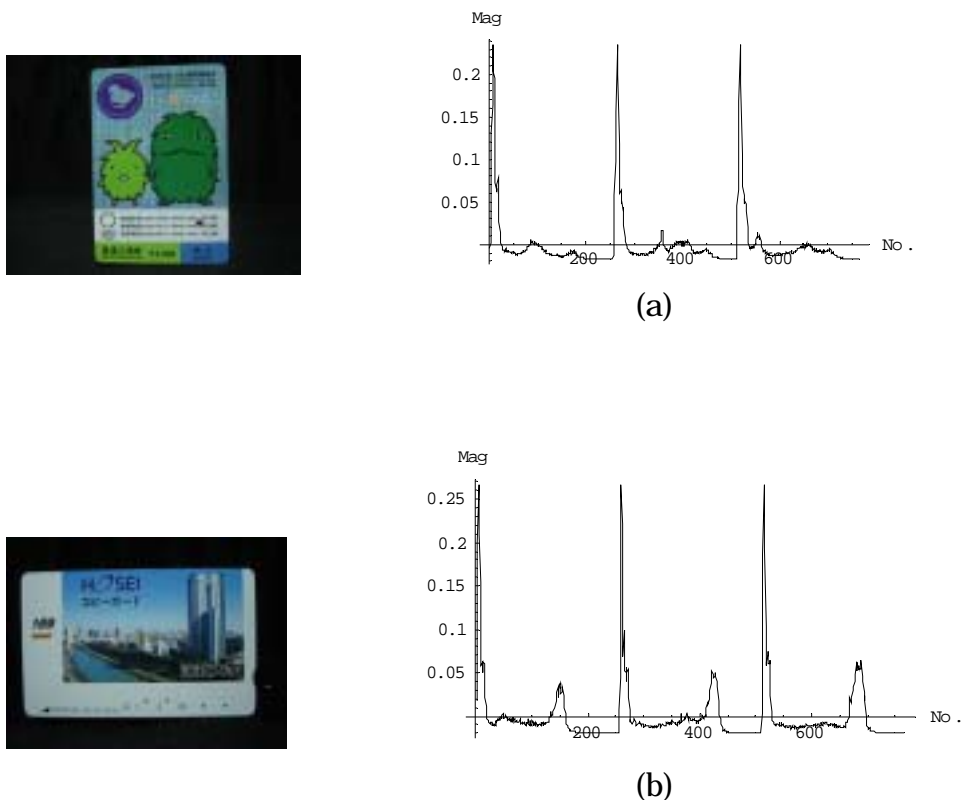


Fig.1 (a),(b) Sample Images and their RGB Eigen Patterns

3. 固有パターンの一意性

3.1 データベース

例として n 個の画像データから固有パターンを抽出しデータベースを構築する。

それぞれの画像データは R 成分固有パターン, G 成分固有パターン, B 成分固有パターンの 3 個の固有パターンを持つ。よってこれらの 3 個の固有パターンは式(6)のシステム行列を構築可能とする。

$$\begin{aligned} C_R &= [E_{R1}, E_{R2}, E_{R3}, \dots, E_{Rn}] \\ C_G &= [E_{G1}, E_{G2}, E_{G3}, \dots, E_{Gn}] \\ C_B &= [E_{B1}, E_{B2}, E_{B3}, \dots, E_{Bn}] \end{aligned} \quad (6)$$

3.2 線形システム方程式

入力画像データの R, G, B 成分固有パターンをそれぞれ E_{YR} , E_{YG} , E_{YB} とすると R, G, B 成分の線形システム方程式は式(7)で表すことができる。 X_R , X_G , X_B は解ベクトルであり, 式(6)を構成する固有パターンの重みに対応し, 解ベクトルの要素で最大値をもつ要素が残りの要素に対して相対的に値が大きいほど, 入力画像データと式(6)を構成する特定の固有パターンが等しいことを意味する。

$$\begin{aligned} E_{YR} &= C_R \cdot X_R \\ E_{YG} &= C_G \cdot X_G \\ E_{YB} &= C_B \cdot X_B \end{aligned} \quad (7)$$

3.3 解ベクトル

固有パターンの要素数はダイナミックレンジ D によって決まり, 本論文では 0 から 255 の 256 段階の範囲で構成している。データベース数 n よりも固有パターンの要素数の方が大きいと仮定する。すなわち, 式(7)のシステム方程式から解ベクトル X_R , X_G , X_B を求めることは不適切問題を解くことに帰する。この不適切問題の解法として、最小自乗法を採用する。解ベクトル X_R , X_G , X_B の導出は式(8)で与えられる。

$$\begin{aligned} X_R &= [C_R^T C_R]^{-1} C_R^T E_{YR} \\ X_G &= [C_G^T C_G]^{-1} C_G^T E_{YG} \\ X_B &= [C_B^T C_B]^{-1} C_B^T E_{YB} \end{aligned} \quad (8)$$

式(8)より得られた R, G, B 解ベクトル X_R , X_G , X_B を畳み込み演算し, 解ベクトル X を得る。* は畳み込み演算を示す。

$$X = X_R * X_G * X_B \quad (9)$$

式(9)で最大値を取る要素から静止画像は認識する。

3.4 静止画像認識

3.1、3.2、3.3 に示した固有パターンと線形システム方程式の静止画像認識手法を用いてデータベースと入力画像それぞれの固有パターン間の一意性を評価し画像認識を行い、精度の確認を行う。

3.4.1 データベース、入力画像

背景の色情報が認識に影響しないように、白熱電球の下で、暗室にて撮影を行った画像を使用する。Fig.2 にデータベース画像、Fig.3 に入力画像を示す。



No.1



No.2



No.3



No.4



No.5



No.6



No.7



No.8



No.9



No.10



No.11



No.12



No.13



No.14



No.15



No.16



No.17



No.18



No.19



No.20



No.21



No.22



No.23



No.24



No.25



No.26



No.27



No.28



No.29



No.30



No.31



No.32



No.33



No.34



No.35



No.36



No.37



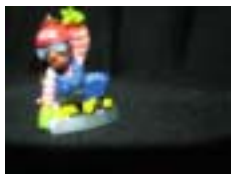
No.38



No.39



No.40



No.41



No.42



No.43



No.44



No.45



No.46



No.47



No.48



No.49



No.50

Fig.2 Database Images



No.1



No.2



No.3



No.4



No.5



No.6



No.7



No.8



No.9



No.10



No.11



No.12



No.13



No.14



No.15



No.16



No.17



No.18



No.19



No.20



No.21



No.22



No.23



No.24



No.25



No.26



No.27



No.28



No.29



No.30



No.31



No.32



No.33



No.34



No.35



Fig.3 Test Images

3.4.2 認識結果

3.4.1 で示した対象物に関して、認識を行う。認識成功例を Fig.4 へ、失敗例を Fig.5 へ示す。

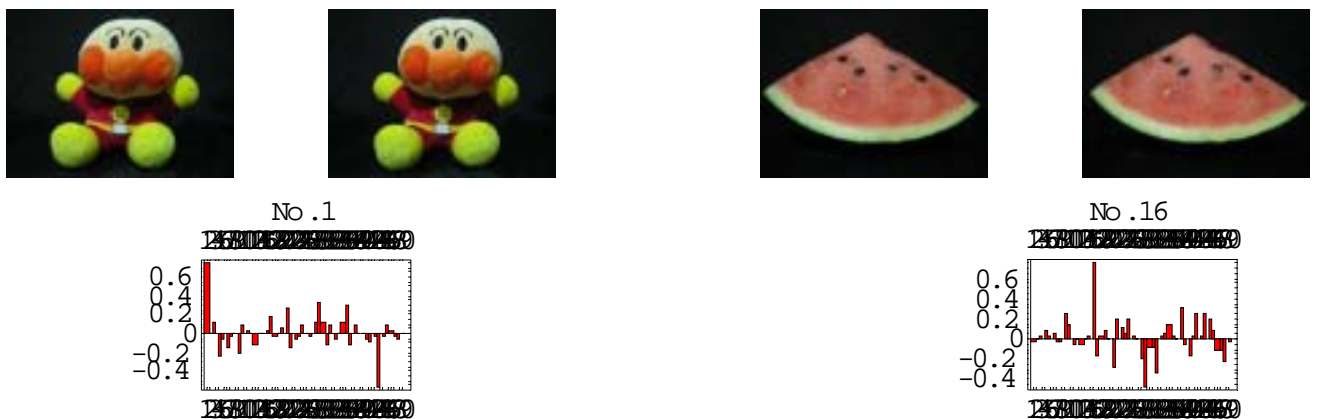


Fig.4 認識成功例

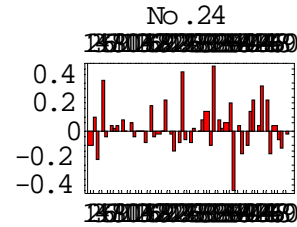
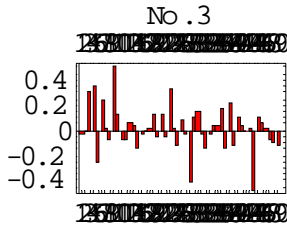


Fig.5 認識失敗例

Table.1

Sample No	Cognition	Result	Sample No	Cognition	Result
1	1	○	26	26	○
2	2	○	27	27	○
3	9	×	28	28	○
4	4	○	29	29	○
5	5	○	30	30	○
6	6	○	31	31	○
7	7	○	32	32	○
8	8	○	33	33	○
9	40	×	34	31	×
10	10	○	35	35	○
11	11	○	36	36	○
12	12	○	37	37	○
13	38	×	38	38	○
14	14	○	39	39	○
15	15	○	40	40	○
16	16	○	41	49	×
17	17	○	42	42	○
18	18	○	43	43	○
19	19	○	44	37	×
20	20	○	45	45	○
21	21	○	46	46	○

22	22	○	47	47	○
23	23	○	48	48	○
24	32	×	49	49	○
25	25	○	50	50	○

Table.1 より認識結果は 50 枚中 43 枚が認識成功し、認識率は 86.0%となった。この認識結果から良好な結果を得ることができた。

今回の認識では、一方向からの情報しか持っていないため、撮影する角度で色情報が異なってしまう。そこで対象物を数方向から撮影し、情報量を増やすことで、より良好な認識率を期待できると考えられる。

3.5 GVSPM 法の導入

3.4 において画像から固有パターンを抽出してデータベースを構築し、入力画像の固有パターンを抽出し、最小自乗法を用いてデータベースと入力画像をそれぞれの固有パターン間の一意性を評価し、静止画像の認識を行った。

ここでは固有パターンの一意性評価法に一般化ベクトルサンプルパターンマッチング法(GVSPM 法)を導入する。

3.5.1 GVSPM 法

ここでは、GVSPM 法の理論を述べる。

(10)式で与えられる一般的な線形システム方程式を考える。

$$\mathbf{Y} = \mathbf{C}\mathbf{X} \quad (10)$$

ここで、 \mathbf{Y} は n 次の入力ベクトル、 \mathbf{X} は m 次の解ベクトル、そして \mathbf{C} は n 行 m 列のシステム行列である。

(10)式は

$$\mathbf{X} = [x_1, x_2 \cdots x_m]^T, \quad \mathbf{C} = [\mathbf{C}_1 \quad \mathbf{C}_2 \quad \cdots \quad \mathbf{C}_m] \quad (11)$$

とすると、(12)式のように書き直すことが可能である。

$$\mathbf{Y} = \sum_{i=1}^m x_i \mathbf{C}_i \quad (12)$$

さらに(12)式の両辺を入力ベクトル \mathbf{Y} と列ベクトル \mathbf{C}_i のそれぞれのノルムで正規化して(13)式を得る。

$$\frac{\mathbf{Y}}{|\mathbf{Y}|} = \sum_{i=1}^m \left(x_i \frac{|\mathbf{C}_i|}{|\mathbf{Y}|} \right) \frac{\mathbf{C}_i}{|\mathbf{C}_i|} \quad (13)$$

ここで

$$\mathbf{X}' = \left[x_1 \frac{\mathbf{C}_1}{|\mathbf{Y}|} \quad x_2 \frac{\mathbf{C}_2}{|\mathbf{Y}|} \quad \cdots \quad x_m \frac{\mathbf{C}_m}{|\mathbf{Y}|} \right]^T, \quad \mathbf{C}' = \left[\frac{\mathbf{C}_1}{|\mathbf{C}_1|} \quad \frac{\mathbf{C}_2}{|\mathbf{C}_2|} \quad \cdots \quad \frac{\mathbf{C}_m}{|\mathbf{C}_m|} \right], \quad \mathbf{Y}' = \frac{\mathbf{Y}}{|\mathbf{Y}|} \quad (14)$$

とすると、(13)式は

$$\mathbf{Y}' = \mathbf{C}' \mathbf{X}' \quad (15)$$

となる。

(12)式は入力ベクトル \mathbf{Y} が必ずシステム行列の列ベクトル \mathbf{C}'_i の線形結合で与えられることを意味する。ここで反復解法である GVSPM 法の評価関数を考える。k 回目の反復解 $\mathbf{CX}^{(k)}$ と入力ベクトル \mathbf{Y} 間の角度の余弦成分

$$\begin{aligned} f(\mathbf{X}^{(k)}) &= \frac{\mathbf{Y} \cdot \mathbf{CX}^{(k)}}{|\mathbf{Y}| |\mathbf{CX}^{(k)}|} = \frac{\mathbf{Y} \cdot |\mathbf{Y}| \mathbf{C}' \mathbf{X}'^{(k)}}{|\mathbf{Y}| |\mathbf{Y}| |\mathbf{C}' \mathbf{X}'^{(k)}|} \\ &= \mathbf{Y}' \cdot \frac{\mathbf{C}' \mathbf{X}'^{(k)}}{|\mathbf{C}' \mathbf{X}'^{(k)}|} \end{aligned} \quad (16)$$

を解の評価関数とし、

$$f(\mathbf{X}^{(k)}) \rightarrow 1 \quad (17)$$

となる解ベクトル $\mathbf{X}^{(k)}$ が GVSPM 法の解である。

まず、第一近似解ベクトル $\mathbf{X}'^{(1)}$ を求めることを考える。反復計算の初期値を $\mathbf{X}'^{(0)}$ とすると(15)式は、

$$\mathbf{C}' \mathbf{X}'^{(0)} = \mathbf{Y}'^{(0)} \quad (18)$$

となる。両辺に \mathbf{C}'^T を掛けると、

$$\mathbf{C}'^T \mathbf{C}' \mathbf{X}'^{(0)} = \mathbf{C}'^T \mathbf{Y}'^{(0)} \quad (19)$$

となり、

$$\mathbf{C}'^T \mathbf{C}' \cong \mathbf{I} \quad (20)$$

が成り立つから、

$$\mathbf{X}'^{(0)} = \mathbf{C}'^T \mathbf{Y}'^{(0)} \quad (21)$$

となる。

反復計算の初期値 $\mathbf{X}'^{(0)}$ が(21)式で与えられるとすると、正規化された入力ベクトルの第一次偏差 $\Delta \mathbf{Y}'^{(1)}$ は

$$\Delta \mathbf{Y}'^{(1)} = \mathbf{Y}' - \frac{\mathbf{C}' \mathbf{X}'^{(0)}}{|\mathbf{C}' \mathbf{X}'^{(0)}|} \quad (22)$$

となる。(22)式が意味するところは、(16)式の、解の評価関数 f と等価な正規化ベクトルの差 $\Delta \mathbf{Y}'$ を用いて、

$$\Delta \mathbf{Y}' \rightarrow 0 \quad (23)$$

となる解ベクトル $\mathbf{X}'^{(1)}$ を求めることである。したがって、 \mathbf{I}_m を m 次の単位正方行列、 $\Delta \mathbf{X}'^{(1)}$ を解の誤差ベクトルとすると、第一近似解ベクトル $\mathbf{X}'^{(1)}$ は(24)式の形で書ける。

$$\begin{aligned} \mathbf{X}'^{(1)} &= \mathbf{X}'^{(0)} + \Delta \mathbf{X}'^{(1)} = \mathbf{X}'^{(0)} + \mathbf{C}'^T \Delta \mathbf{Y}'^{(1)} = \mathbf{X}'^{(0)} + \mathbf{C}'^T \left(\mathbf{Y}' - \frac{\mathbf{C}' \mathbf{X}'^{(0)}}{|\mathbf{C}' \mathbf{X}'^{(0)}|} \right) \\ &= \mathbf{C}'^T \mathbf{Y}' + \left(\mathbf{I}_m - \frac{\mathbf{C}'^T \mathbf{C}'}{|\mathbf{C}' \mathbf{X}'^{(0)}|} \right) \mathbf{X}'^{(0)} \end{aligned} \quad (24)$$

最終的に、k 回目の解ベクトル $\mathbf{X}'^{(k)}$ は(25)式で与えられる。

$$\begin{aligned}
 \mathbf{X}'^{(k)} &= \mathbf{X}'^{(k-1)} + \mathbf{C}'^T \Delta \mathbf{Y}'^{(k-1)} \\
 &= \mathbf{C}'^T \mathbf{Y}' + \left(\mathbf{I}_m - \frac{\mathbf{C}'^T \mathbf{C}'}{|\mathbf{C}' \mathbf{X}'^{(k-1)}|} \right) \mathbf{X}'^{(k-1)}
 \end{aligned}
 \tag{25}$$

式(25)で最大値を取る要素から静止画像は認識する。

3.6 GVSPM 法の静止画像認識への適用

3.5 に示した GVSPM 法を用いてデータベースと入力画像それぞれの固有パターン間の一意性を評価し静止画像認識を行い、精度の確認を行う。反復回数は 1000 回、2000 回、3000 回で行い、出力画像がそれぞれの結果で一致した場合を入力画像に対する出力画像とする。それぞれの出力画像が異なる場合は、認識不能とする。

3.6.1 データベース、入力画像

3.4 で使用した画像、Fig.2 をデータベース、Fig.3 を入力画像とする。

3.6.2 認識結果

認識成功例を Fig.6 へ、失敗例を Fig.7 へ、認識不能例を Fig.8 へ示す。

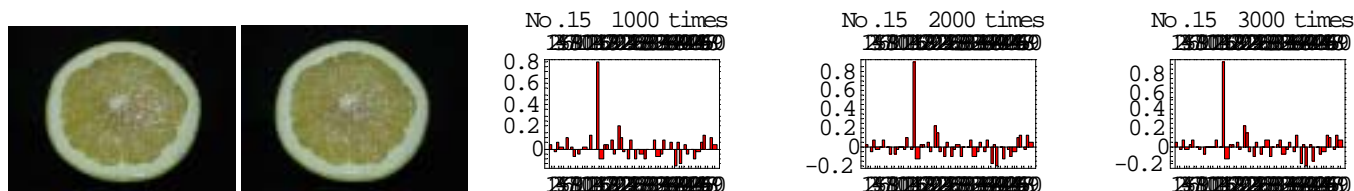


Fig.6 認識成功例

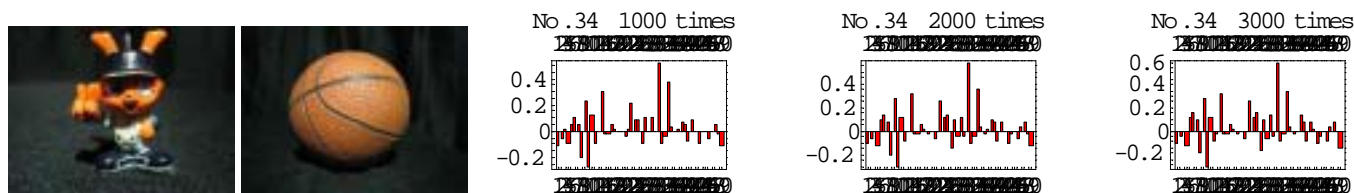


Fig.7 認識失敗例

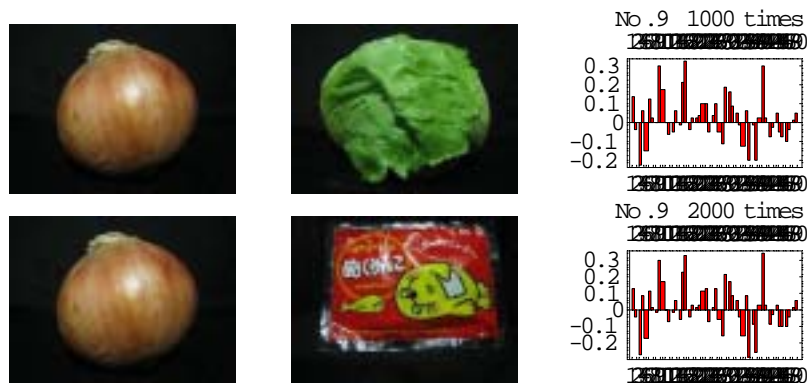




Fig.8 認識不能例

次に認識結果を Table.2 へと示す。○、×、□のそれぞれを認識成功、認識失敗、認識不能とする。

Table.2

Sample No	Cognition(1000)	Cognition(2000)	Cognition(3000)	Result
1	1	1	1	○
2	2	2	2	○
3	9	9	9	×
4	4	4	4	○
5	5	5	5	○
6	6	6	6	○
7	7	7	7	○
8	8	8	8	○
9	17	40	40	
10	10	10	10	○
11	11	11	11	○
12	12	12	12	○
13	38	38	38	×
14	14	14	14	○
15	15	15	15	○
16	16	16	16	○
17	19	17	17	○
18	18	18	18	○
19	19	19	19	○
20	20	20	20	○
21	21	21	21	○
22	22	22	22	○
23	23	23	23	○
24	24	24	32	○
25	25	25	25	○
26	26	26	26	○

27	27	27	27	○
28	28	28	28	○
29	29	29	29	○
30	30	30	30	○
31	31	31	31	○
32	32	32	32	○
33	33	33	33	○
34	31	31	31	×
35	35	35	35	○
36	36	36	36	○
37	37	37	37	○
38	38	38	38	○
39	39	39	39	○
40	40	40	40	○
41	40	49	49	×
42	49	42	42	○
43	43	43	43	○
44	29	29	37	×
45	45	45	45	○
46	46	46	46	○
47	47	47	47	○
48	48	48	48	○
49	49	49	49	○
50	50	50	50	○

Table.2 より認識結果は 50 枚中 44 枚が認識成功し、認識率は 88.0%となった。

今回の認識から、固有パターンの一意性評価方法を最小自乗法から GVSPM 法に変えたことによって、認識結果が向上し、固有パターンの評価方法によって認識結果が変わることが確認できた。

また、GVSPM 法の反復回数が 1000 回、2000 回、3000 回の時のそれぞれの結果は、50 枚中 44 枚、44 枚、44 枚が認識成功し、認識率はそれぞれ 88.0%、88.0%、88.0%となった。

GVSPM 法による固有パターンの一意性評価方法では、反復回数によって解が変わってしまう点が難点であるが、今回の認識結果では、反復回数によって解が変わった結果は 1 枚であり、良好な結果が得られた。また、反復回数によって解が変わる場合は、認識画像の候補として利用することによって、より精度の良い認識が行えることが期待できる。

4. 合成画像法による動画認識

R, G, B 成分固有パターンと線形システム方程式を用いた静止画像認識手法を動画認識へ一般化する。

4.1 合成画像法

動画は静止画像を時系列に並べた一連の画像であり、現在一般的に使われているビデオカメラは1秒間に 29.97 枚のフレーム画像を撮影可能である。これは、動画認識を静止画像認識法で行うには、すべてのフレーム画像による画素値分布から画像の普遍量を抽出する必要があることを意味する。

すべてのフレーム画像を1画像へ合成し、合成画像から R、G、B 成分固有パターンを抽出する。Fig.9 に動画の合成画像と固有パターンを示す。

R、G、B 成分固有パターンは配置、角度、変形などの影響を受けない性質を持つので、対象物が移動や変形をしても認識が可能である。

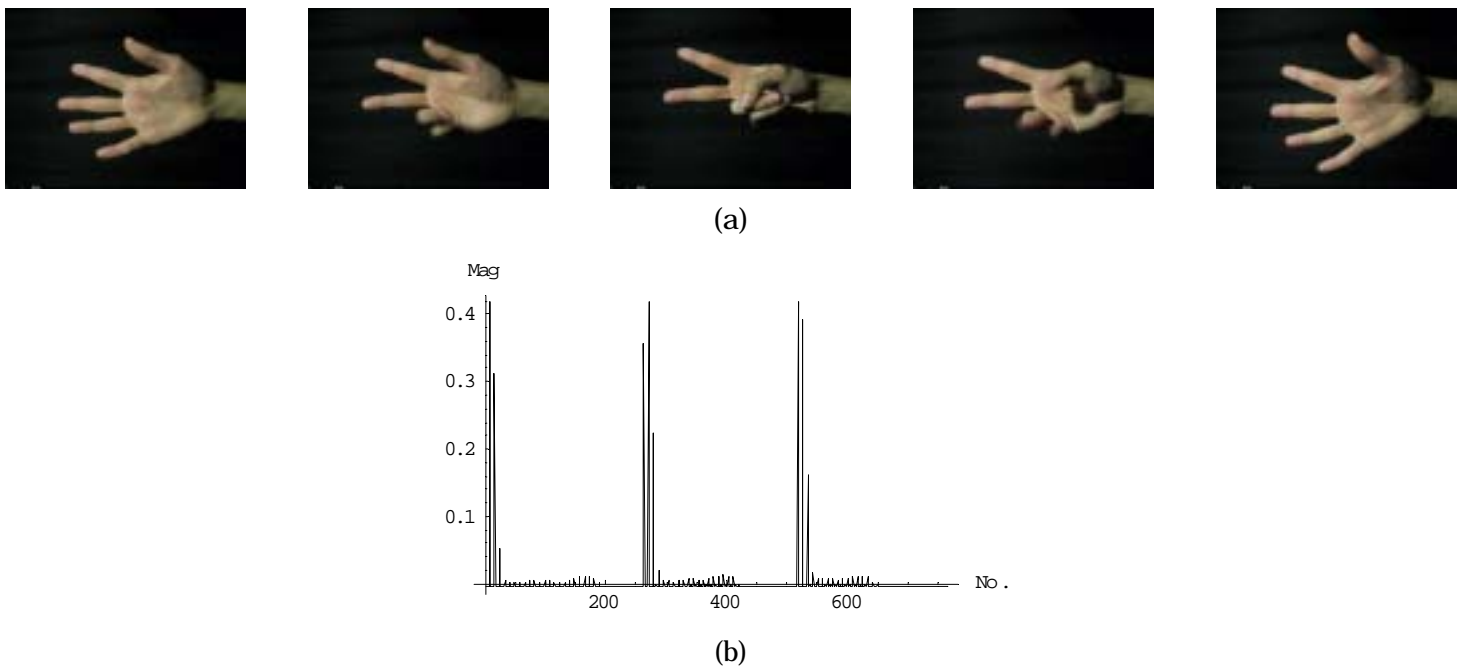


Fig.9 (a),(b) Sample Dynamic Image and Its RGB Eigen Pattern

5. 動画像認識の手話への応用

動画像認識の精度を見るために、簡単な動作による動画像認識を行う。ここでは手話の認識を行う。4.1 に示した合成画像法を用いて、3.1、3.2、3.3 に示した固有パターンと線形システム方程式によるデータベースと入力画像それぞれの固有パターン間の一意性を評価し動画像認識を行う。

5.1 同一人物による手話の識別

5.1.1 データベース、入力画像

背景の色情報が認識に影響しないように、白熱電球の下で、暗室にて撮影を行った画像を利用する。Fig.10 にデータベース、Fig.11 に入力画像を示す。Fig.10 と Fig.11 は、5人が20通りの手話を行っている動画像である。ここでは、5人、別々に認識を行う。同一人物による動作をわかりやすいように、5人の人物を Group A、Group B、Group C、Group D、Group E とする。



A1 A2 A3 A4 A5 A6 A7 A8 A9 A10



A11 A12 A13 A14 A15 A16 A17 A18 A19 A20

Group A



B1 B2 B3 B4 B5 B6 B7 B8 B9 B10



B11 B12 B13 B14 B15 B16 B17 B18 B19 B20

Group B



Fig.10 Database Images



A1 A2 A3 A4 A5 A6 A7 A8 A9 A10



A11 A12 A13 A14 A15 A16 A17 A18 A19 A20

Group A



B1 B2 B3 B4 B5 B6 B7 B8 B9 B10



B11 B12 B13 B14 B15 B16 B17 B18 B19 B20

Group B



C1 C2 C3 C4 C5 C6 C7 C8 C9 C10



C11 C12 C13 C14 C15 C16 C17 C18 C19 C20

Group C

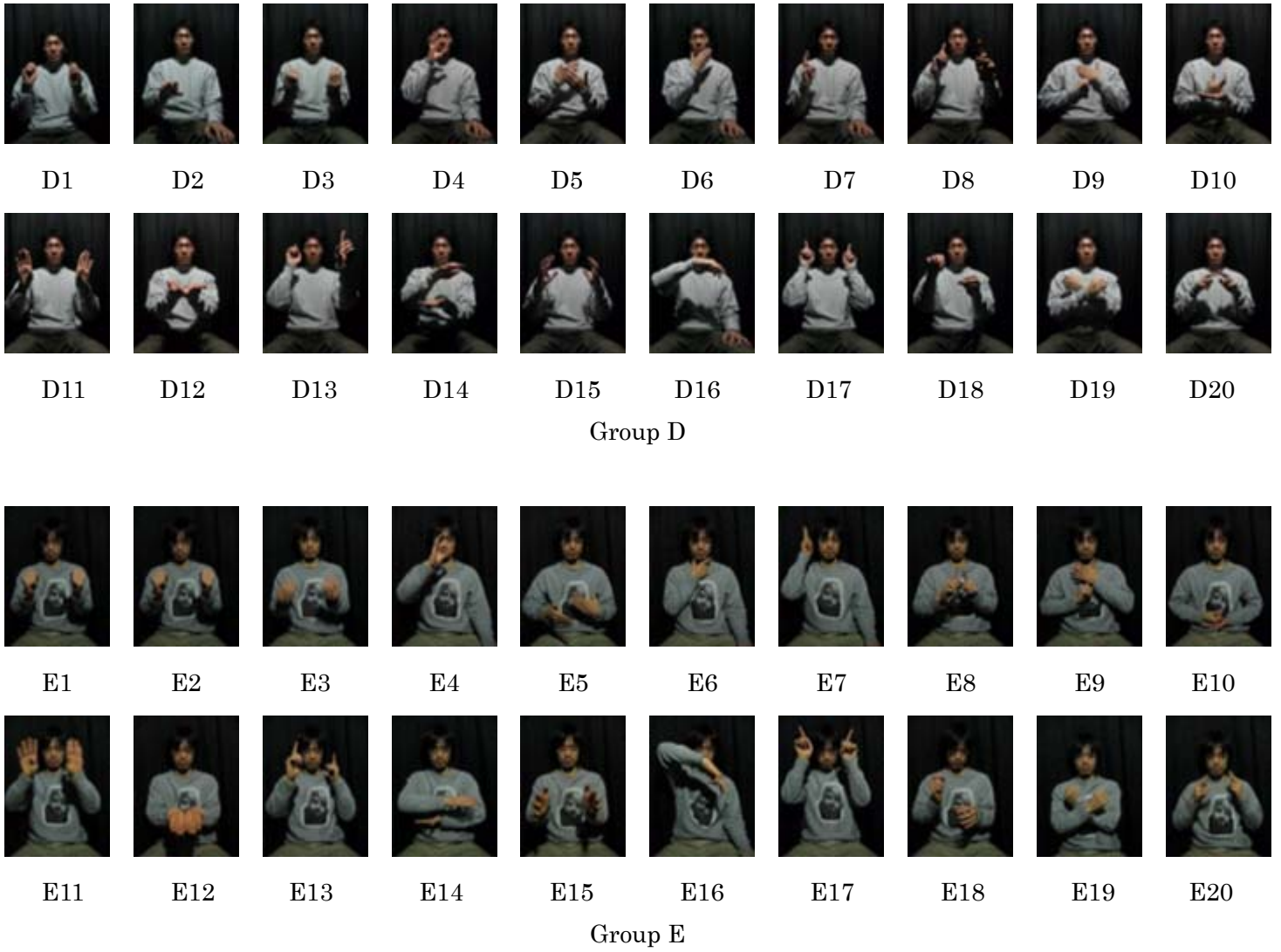


Fig.11 Test Images

5.1.2 認識結果

認識成功例を Fig.12 へ、失敗例を Fig.13 へ示す。





Fig.12 認識成功例



Fig.13 認識失敗例

次に認識結果を Table.3 へと示す。

Table.3
Group A

Sample No.	Cognition	Result	Sample No.	Cognition	Result
A1	A1	○	A11	A8	×
A2	A2	○	A12	A12	○
A3	A3	○	A13	A13	○
A4	A4	○	A14	A14	○
A5	A5	○	A15	A15	○
A6	A6	○	A16	A16	○
A7	A5	×	A17	A17	○
A8	A8	○	A18	A4	×
A9	A9	○	A19	A19	○
A10	A15	×	A20	A20	○

Group B

Sample No.	Cognition	Result	Sample No.	Cognition	Result
B1	B1	○	B11	B11	○
B2	B2	○	B12	B12	○

B3	B3	○	B13	B20	×
B4	B4	○	B14	B20	×
B5	B5	○	B15	B15	○
B6	B6	○	B16	B16	○
B7	B6	×	B17	B11	×
B8	B12	×	B18	B18	○
B9	B9	○	B19	B19	○
B10	B10	○	B20	B20	○

Group C

Sample No.	Cognition	Result	Sample No.	Cognition	Result
C1	C1	○	C11	C11	○
C2	C2	○	C12	C19	×
C3	C11	×	C13	C13	○
C4	C4	○	C14	C14	○
C5	C5	○	C15	C15	○
C6	C3	×	C16	C16	○
C7	C15	×	C17	C17	○
C8	C8	○	C18	C18	○
C9	C9	○	C19	C19	○
C10	C4	×	C20	C20	○

Group D

Sample No.	Cognition	Result	Sample No.	Cognition	Result
D1	D1	○	D11	D11	○
D2	D2	○	D12	D12	○
D3	D3	○	D13	D16	×
D4	D4	○	D14	D15	×
D5	D5	○	D15	D8	×
D6	D19	×	D16	D16	○
D7	D7	○	D17	D17	○
D8	D12	×	D18	D18	○
D9	D9	○	D19	D19	○
D10	D10	○	D20	D20	○

Group E

Sample No.	Cognition	Result	Sample No	Cognition	Result
E1	E1	○	E11	E11	○
E2	E2	○	E12	E12	○
E3	E3	○	E13	E13	○
E4	E4	○	E14	E14	○
E5	E3	×	E15	E15	○
E6	E6	○	E16	E16	○
E7	E7	○	E17	E17	○
E8	E12	×	E18	E18	○
E9	E13	×	E19	E20	×
E10	E10	○	E20	E20	○

Table.2 より認識結果は Group A、Group B、Group C、Group D、Group E のそれぞれは、20 枚中 16 枚、15 枚、15 枚、15 枚、16 枚が認識成功し、認識率は 80%、75%、75%、75%、80% となった。認識結果から、異なる動作を色情報によって認識した場合、7 割以上の認識が可能であることがわかる。動作の撮影方法によって、さらに認識率の向上が期待できることが考えられる。

5.2 数人による手話の識別

5.2.1 データベース、入力画像

5 人による 20 通りの手話を同時に認識する。5.1 で使用した画像、Fig.10 の Group A ~ Group E をデータベース、Fig.11 の Group A ~ Group E を入力画像とする。

5.2.2 認識結果

認識成功例を Fig.14 へ、失敗例を Fig.15 へ示す。

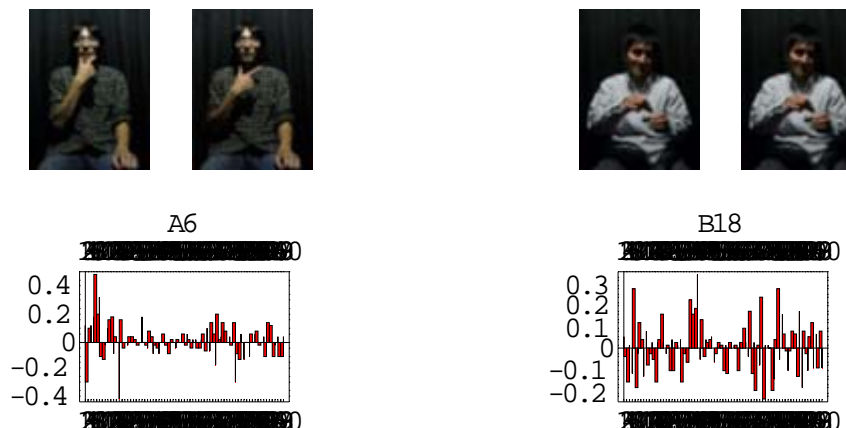


Fig.14 認識成功例

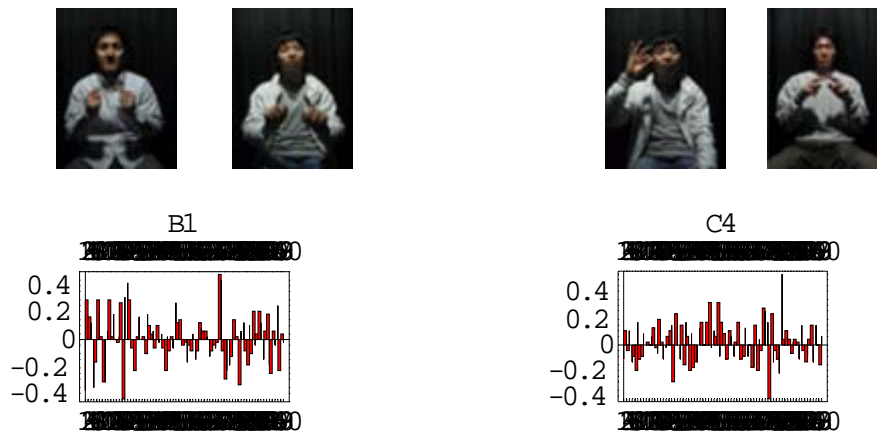


Fig.15 認識失敗例

次に認識結果を Table.4 へと示す。

Table.4

Sample No.	Cognition	Result	Sample No.	Cognition	Result
A1	A1	○	C11	A10	×
A2	A2	○	C12	C12	○
A3	A3	○	C13	C13	○
A4	A4	○	C14	C14	○
A5	A5	○	C15	C15	○
A6	A6	○	C16	C16	○
A7	A4	×	C17	C17	○
A8	A8	○	C18	C18	×
A9	A10	×	C19	C19	○
A10	A10	○	C20	C20	○
A11	A11	○	D1	D1	○
A12	A12	○	D2	D2	○
A13	A13	○	D3	D3	○
A14	A14	○	D4	D4	○
A15	A15	○	D5	D5	○
A16	A16	○	D6	D6	○
A17	A17	○	D7	D7	○
A18	A16	×	D8	D12	×
A19	A19	○	D9	D9	×
A20	A20	○	D10	D10	○

B1	C8	×	D11	D11	○
B2	B2	○	D12	D14	×
B3	B3	○	D13	A7	×
B4	B4	○	D14	D14	○
B5	B5	○	D15	D15	○
B6	A3	×	D16	D16	○
B7	B6	×	D17	D17	○
B8	B8	○	D18	D18	○
B9	B9	○	D19	D20	×
B10	D20	×	D20	D12	×
B11	A12	×	E1	E1	○
B12	B11	×	E2	E2	○
B13	D16	×	E3	E3	○
B14	B15	×	E4	E20	×
B15	D18	×	E5	E3	×
B16	B16	○	E6	E6	○
B17	B16	×	E7	E3	×
B18	B18	○	E8	E8	○
B19	A8	×	E9	E8	×
B20	D18	×	E10	E2	×
C1	C1	○	E11	E11	○
C2	C2	○	E12	E12	○
C3	A2	×	E13	E13	○
C4	D20	×	E14	E14	○
C5	A2	×	E15	E15	○
C6	E13	×	E16	E16	○
C7	D20	×	E17	E17	○
C8	A1	×	E18	E18	○
C9	C9	×	E19	E19	○
C10	D8	×	E20	E20	○

Table.3 より認識結果は 100 枚中 64 枚が認識成功し、認識率は 64%となった。認識結果から、数人の動作を色情報によって、同時に認識を行った場合、別々に認識した 5.1 の認識率より低下したことがわかる。別々に認識を行った場合、7 割以上の認識が可能であることから、今回の認識では対象画像が増えたことによって、適切な解が得られず、固有パターンの一意性評価方法が適切ではなかったことが考えられる。

5.3 GVSPM 法の動画認識への適用

ここで、4.1 に示した合成画像法を用いて、静止画像認識で用いた GVSPM 法を適用し動画認識を行う。5.2 で行った認識と同様に 5 人による 20 通りの手話を同時に認識する。反復回数は 1000 回、2000 回、3000 回で行い、出力画像がそれぞれの結果で一致した場合を入力画像に対する出力画像とする。それぞれの出力画像が異なる場合は、認識不能とする。

5.3.1 データベース、入力画像

5.2 と同様に、5.1 で使用した画像、Fig.10 の Group A ~ Group E をデータベース、Fig.11 の Group A ~ Group E を入力画像とする。

5.3.2 認識結果

認識成功例を Fig.16 へ、失敗例を Fig.17 へ、認識不能例を Fig.18 へ示す。

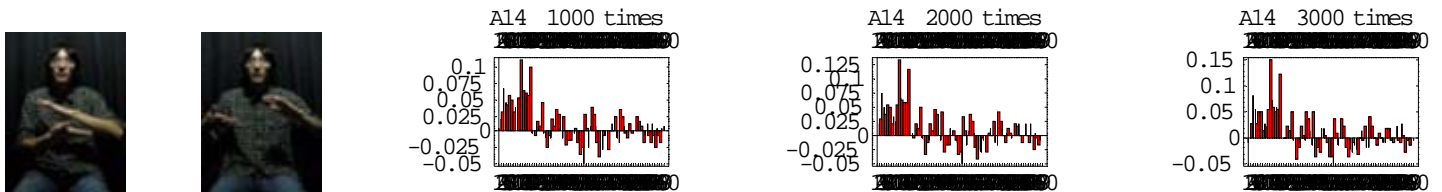


Fig.16

認識成功例

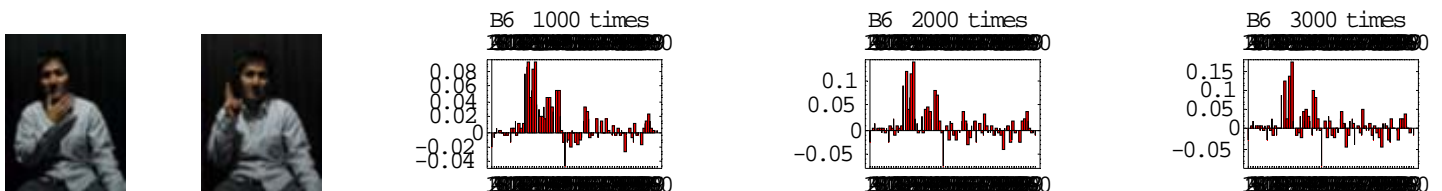


Fig.17

認識失敗例



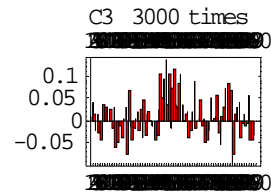


Fig.18 認識不能例

次に認識結果を Table.5 へと示す。○、×、 のそれぞれを認識成功、認識失敗、認識不能とする。

Table.5

Sample No.	Cognition(1000)	Cognition(2000)	Cognition(3000)	Result
A1	A1	A1	A1	○
A2	A1	A2	A2	
A3	A1	A1	A1	×
A4	A4	A4	A4	○
A5	A15	A15	A15	×
A6	A4	A4	A4	×
A7	A4	A4	A4	×
A8	A1	A1	A1	×
A9	A10	A10	A10	×
A10	A20	A11	A11	×
A11	A1	A11	A11	
A12	A1	A12	A12	
A13	A15	A15	A15	
A14	A14	A14	A14	○
A15	A15	A15	A15	○
A16	A16	A16	A16	○
A17	A17	A17	A17	○
A18	A17	A17	A18	
A19	A17	A19	A19	
A20	A20	A20	A20	○
B1	B2	B2	B2	×
B2	B2	B2	B2	○
B3	B3	B3	B3	○
B4	B17	B4	B4	
B5	C41	B8	B8	
B6	B7	B7	B7	×

B7	B2	B2	B7	×
B8	C1	C1	B2	
B9	B9	B9	B9	○
B10	B10	B10	B10	○
B11	B11	B11	B11	○
B12	B19	B12	B12	
B13	B7	B7	B7	×
B14	B20	B20	B20	×
B15	B7	B7	B7	×
B16	B16	B16	B16	○
B17	B17	B17	B17	○
B18	B18	B18	B18	○
B19	B20	B20	B20	×
B20	B20	B20	B20	○
C1	C1	C1	C1	○
C2	C1	C2	C2	
C3	C15	C6	C6	
C4	C15	C4	C4	
C5	C12	C12	C5	
C6	C9	C9	C9	×
C7	C15	C15	C15	×
C8	C8	C8	C8	○
C9	C9	C9	C9	○
C10	C12	C12	C11	
C11	C13	C13	C13	×
C12	C12	C12	C12	○
C13	C13	C13	C13	○
C14	C13	C14	C14	
C15	C15	C15	C15	○
C16	C16	C16	C16	○
C17	C17	C17	C17	○
C18	C18	C18	C18	○
C19	C19	C19	C19	○
C20	C20	C20	C20	○
D1	D1	D1	D1	○

D2	D2	D2	D2	○
D3	D3	D3	D3	○
D4	D4	D4	D4	○
D5	D5	D5	D5	○
D6	D6	D6	D6	○
D7	D7	D7	D7	○
D8	D11	D11	D13	×
D9	D14	D9	D9	
D10	D14	D10	D10	
D11	D11	D11	D11	○
D12	D14	D12	D12	×
D13	D4	D14	D16	×
D14	D14	D14	D14	○
D15	D14	D14	D14	×
D16	D16	D16	D16	○
D17	D17	D17	D17	○
D18	D18	D18	D18	○
D19	D14	D14	D9	
D20	D14	D14	D20	×
E1	E1	E1	E1	○
E2	E2	E2	E2	○
E3	E3	E3	E3	○
E4	E4	E4	E4	○
E5	E5	E5	E5	○
E6	E6	E6	E6	○
E7	E7	E7	E7	○
E8	E8	E8	E8	○
E9	E9	E10	E10	
E10	E10	E10	E10	○
E11	E11	E11	E11	○
E12	E12	E12	E12	○
E13	E16	E16	E16	×
E14	E14	E14	E14	○
E15	E15	E15	E15	○
E16	E16	E16	E16	○

E17	E17	E17	E17	○
E18	E18	E18	E18	○
E19	E19	E19	E19	○
E20	E20	E20	E20	○

Table.4 より認識結果は 100 枚中 57 枚が認識成功し、認識率は 57.0%となった。

今回の認識では、固有パターンの一意性評価方法を最小自乗法から GVSPM 法に変えたことによって、認識結果が低下した。認識率の低下の原因として、反復回数によって解が変わる場合が増え、認識不能が多かったためであると考えられる。

また、GVSPM 法の反復回数が 1000 回、2000 回、3000 回の時のそれぞれの結果は、100 枚中 58 枚、69 枚、71 枚が認識成功し、認識率はそれぞれ 58.0%、69.0%、71.0%となった。この結果より、この認識の固有パターンの一意性評価方法は最小自乗法より 3000 回の反復による GVSPM 法の方が適切であることがわかる。

5.4 クラスター分類を利用した GVSPM 法の動画像認識への適用

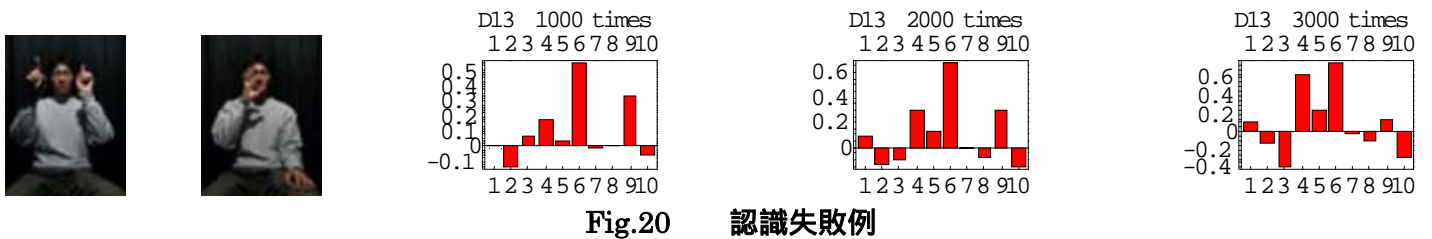
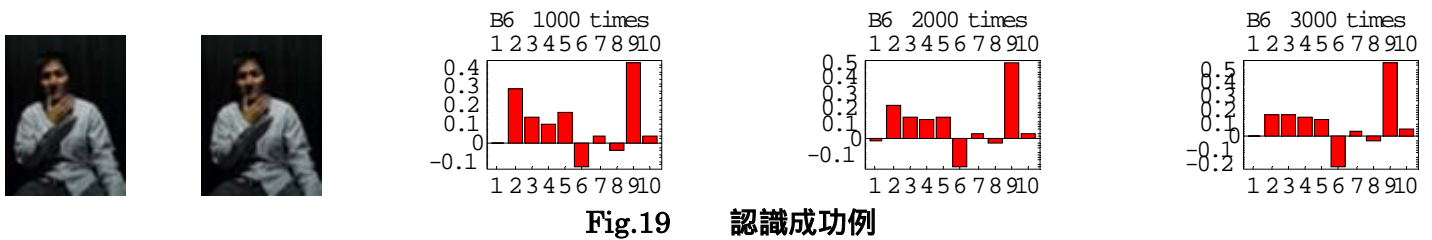
5.3 と同様に 4.1 に示した合成画像法を用いて、静止画像認識で用いた GVSPM 法を適用し動画像認識を行う。ここでは、最小自乗法によって得られた解ベクトルから最も大きい 10 個の要素をデータベースとして、GVSPM 法を適用する。反復回数は 1000 回、2000 回、3000 回で行い、出力画像がそれぞれの結果で一致した場合を入力画像に対する出力画像とする。それぞれの出力画像が異なる場合は、認識不能とする。

5.4.1 データベース、入力画像

5.3 と同様に、5.1 で使用した画像、Fig.10 の Group A ~ Group E をデータベース、Fig.11 の Group A ~ Group E を入力画像とする。

5.4.2 認識結果

認識成功例を Fig.19 へ、失敗例を Fig.20 へ、認識不能例を Fig.21 へ示す。



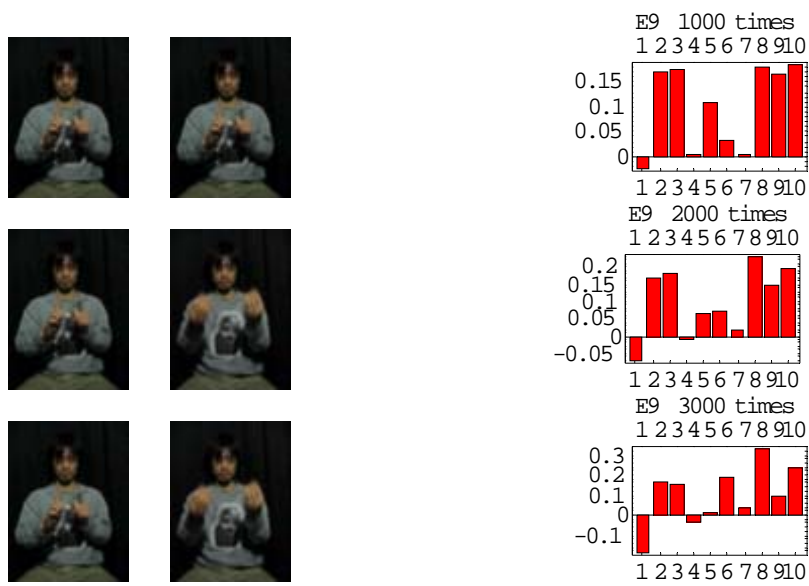


Fig.21 認識不能例

次に認識結果を Table.6 へと示す。○、×、□のそれぞれを認識成功、認識失敗、認識不能とする。

Table.6

Sample No.	Cognition(1000)	Cognition(2000)	Cognition(3000)	Result
A1	A1	A1	A1	○
A2	A2	A2	A2	○
A3	A1	A1	A3	
A4	A4	A4	A4	○
A5	A5	A5	A5	○
A6	A6	A6	A6	○
A7	A4	A4	A4	×
A8	A8	A8	A8	○
A9	A9	A9	A9	○
A10	A7	A7	A10	
A11	A11	A11	A11	○
A12	A12	A12	A12	○
A13	A15	A15	A15	×
A14	A14	A14	A14	○
A15	A15	A15	A15	○
A16	A16	A16	A16	○
A17	A17	A17	A17	○
A18	A18	A18	A18	○

A19	A19	A19	A19	○
A20	A20	A20	A20	○
B1	B2	B2	B2	×
B2	B2	B2	B2	○
B3	B3	B3	B3	○
B4	B4	B4	B4	○
B5	B8	B8	B8	×
B6	B6	B6	B6	○
B7	B6	B6	B6	×
B8	B9	B9	B9	×
B9	B9	B9	B9	○
B10	B10	B10	B10	○
B11	B11	B11	A3	○
B12	B5	B5	B5	×
B13	B20	B20	B20	×
B14	B15	B17	B17	
B15	B15	B15	B15	×
B16	B16	B16	B16	○
B17	B17	B17	B17	○
B18	B18	B18	B18	○
B19	B19	B19	B19	○
B20	B20	B20	B20	○
C1	C1	C1	C1	
C2	C2	C2	C2	○
C3	C3	C3	C3	○
C4	C4	C4	C4	○
C5	C3	C3	C3	○
C6	C6	C6	C6	○
C7	C10	C10	C10	×
C8	C8	C8	C8	○
C9	C9	C9	C9	○
C10	C10	C10	C10	○
C11	C11	C11	C11	○
C12	C12	C12	C12	○
C13	C13	C13	C13	○

C14	C14	C14	C14	○
C15	C15	C15	C15	○
C16	C16	C16	C16	○
C17	C17	C17	C17	○
C18	C18	C18	C18	○
C19	C19	C19	C19	○
C20	C20	C20	C20	○
D1	D1	D1	D1	○
D2	D2	D2	D2	○
D3	D3	D3	D3	○
D4	D6	D4	D4	
D5	D5	D5	D5	○
D6	D6	D11	D11	
D7	D7	D7	D7	○
D8	D11	D11	D11	×
D9	D15	D15	D9	
D10	D19	D19	D10	
D11	D11	D11	D11	○
D12	D12	D12	D12	○
D13	D4	D4	D4	×
D14	D14	D14	D14	○
D15	D15	D15	D15	○
D16	D17	D16	D16	
D17	D17	D17	D17	○
D18	D18	D18	D18	○
D19	D15	D15	D15	×
D20	D14	D14	D14	×
E1	E1	E1	E1	○
E2	E2	E2	E2	○
E3	E3	E3	E3	○
E4	E4	E4	E4	○
E5	E5	E5	E5	○
E6	E6	E6	E6	○
E7	E7	E7	E7	○
E8	E8	E12	E12	

E9	E8	E13	E13	
E10	E10	E10	E10	○
E11	E11	E11	E11	○
E12	E12	E12	E12	○
E13	E13	E13	E13	○
E14	E14	E14	E14	○
E15	E15	E15	E15	○
E16	E16	E17	E17	
E17	E17	E17	E18	
E18	E18	E18	E18	○
E19	E19	E19	E19	○
E20	E20	E20	E20	○

Table.5 より認識結果は 100 枚中 73 枚が認識成功し、認識率は 73.0%となった。

今回の認識では、5.3 に比べ、反復回数によって解が変わる場合が少なくなり、認識不能が減ったため、認識率が向上した。

また、GVSPM 法の反復回数が 1000 回、2000 回、3000 回の時のそれぞれの結果は、100 枚中 77 枚、77 枚、79 枚が認識成功し、認識率はそれぞれ 77.0%、77.0%、79.0%となった。この結果より、GVSPM 法の反復回数を変えることによって、80%以上の認識率も期待できると考えられる。

6. まとめ

本論文では、画像を構成する画素数、画像の2次元平面上の位置・角度などのハードウェアに依存する画像の性質を削除するため、画像の普遍量が R、G、B 成分より構成されていることを述べ、画像の固有パターン抽出手法を述べた。さらに、動画画像を合成静止画像として扱うことで、静止画像の画像認識手法を動画画像認識へ一般化し、これを用いて、動作の識別について述べた。

入力画像から抽出された固有パターンを線形システム方程式の入力ベクトルとして解が求まる。線形システム方程式は不適切であるため、最小自乗法を用いて近似解を導いた。また、固有パターン間の一意性を評価する方法として、一般化ベクトルサンプルパターンマッチング法(GVSPM 法)による逆問題的手法を提案した。さらに、最小自乗法と GVSPM 法を用いたクラスター分類による考え方を提案した。提案手法では、適切な反復回数が決まっていない点が難点であるが、最小自乗法を用いた認識結果と比較し、提案手法を用いることにより良好な認識結果を得ることを確認した。

よって、本論文で提案した方法は、極めて高い認識率が得られたことから、人間の視覚情報処理に近い識別能力を有することを明らかにした。

7. 参考文献

- [1] 佐藤隆紀、早野誠治、齋藤兆古「知的可視化情報処理による動画像認識」
可視化情報学会誌、Vol.22, No.1 (2002) pp.243-246.
- [2] 佐藤隆紀、早野誠治、齋藤兆古「知的可視化情報処理による動画像認識」
可視化情報学会誌、Vol.22, No.1 (2002) pp.243-246.
- [3] 丸山和夫、早野誠治、齋藤兆古「色情報を利用した知的動画像認識」
可視化情報学会誌、Vol.23, No.1 (2003) pp.95-98.

2005 年度
卒業論文

R G B 色情報を利用した画像認識法における
認識精度向上に関する考察

法政大学

工学部 電気電子工学科

学籍番号 01d3126

氏名 萩野 祐史

指導教授 斎藤 兆古

目次

1. 序論

2. 静止画像の固有パターン
 - 2.1 画像の色成分
 - 2.2 カラー画像の固有ベクトル
 - 2.3 固有パターン

3. 固有パターンの一貫性評価法
 - 3.1 データベース
 - 3.2 線形システム方程式
 - 3.3 解ベクトル
 - 3.4 静止画像認識
 - 3.5 最小自乗法による認識の向上
 - 3.6 二回の最小自乗法による認識の評価（その1）
 - 3.7 二回の最小自乗法による認識の評価（その2）

4. プログラム

5. まとめ

6. 参考文献

1. 序論

現在，人間の評価能力を前提として映像を計算機によって処理・加工する技術は比較的多く実用化されている。しかし、映像を認識する技術は完成度が低く、最終的には人間の視覚情報処理能力へ依存せざるを得ない。圧倒的な演算処理能力をもつ最新の計算機をもってしても、画像認識の分野においては人間の視覚情報処理能力には到底及ばない。映像を計算機によって認識させることは、計算機に人間と同等な視覚情報処理能力を与えることになり、現状の様々な手法では人間の視覚情報処理能力には匹敵し難い。

このような現状をふまえて、本論文では、電子的に計算機スクリーン上に可視化された映像から、画素数，配置，変形等に無関係となる映像の普遍量を抽出する概念を導入し、人間の視覚情報処理能力を計算機に与え、計算機自身によって映像を認識させることを試みる。

ここで我々が認識の対象とする映像は、人間の視覚によって認識し得る形態で表現された情報をいい、それらを総称して本論文では可視化情報と呼ぶ。可視化情報の例として、人間の視覚で捉えることが可能なカメラやビデオで撮影された画像等の映像がある。また、通常、人間の視覚情報処理機能で捉えることができない情報を映像化した赤外線映像や電子顕微鏡による映像等も可視化情報の一種である。さらに、人間の視覚で認識可能な形態へ変形された信号情報、例えば音声等を電子的にオシロスコープ等の機器で映像化した情報も可視化情報である。

人間の脳の視聴覚情報機能は、眼球を通して得られた映像を処理することによって外界の情報を認識するが、その対象となる視覚情報には、一定の構成ルールに従ってコード化された文字や言語等のコード化視聴覚情報と、コード化されていない可視化情報、すなわち非コード化視聴覚情報がある。前述の可視化情報は、非コード化視覚情報を処理する人間の視覚情報処理能力を前提としたデータであり、人間は人間の持つ視覚情報処理能力によって何らかの普遍量を感じ認識しているものと考えられる。従って、計算機によって人間の視覚情報処理能力を実現し認識するためには、可視化情報から計算機で処理可能な何らかの普遍量を抽出する必要がある。

計算機のスクリーン上の可視化情報は 2 次元平面上の画素(Pixel)で構成されている。各画素は可視光の波長によってそれぞれ赤、緑、及び青(以下それぞれ R、G、B と略記)成分の情報を持つ。また、可視化情報は反射光の強さ(Intensity)、色相(Tone)、及び色成分(Color component)の組み合わせからなり、それらの幾何学的配置で表現される。そのため、計算機のスクリーン上の可視化情報は、スクリーンの解像度や基準座標(視点)に依存する性質をもつ。このようなハードウェアに依存する性質を削減した可視化情報の普遍量を、本論文では映像の固有パターン(Eigen pattern)と呼び、この画像の固有パターンを用いて画像認識を試みる。

本論文で採用する画像認識手法は線形システム方程式の解を導くことに基づく方法である。しかしながら、得られた線形システム方程式は不適切なシステム方程式であるため、最小自乗法を用いて近似解を導く。換言すれば、本論文で採用する画像認識手法は、あらかじめ用意した画像データをデータベースとしその中から誤差ノルムが最小となる解要素を求める方法である。

本論文の構成は、最初に静止画像から固有パターンを抽出する方法について述べる。次に抽出された固有パターンの一致評価法について述べ、最小自乗法を用いてデータベースと入力画像それぞれの固有パターン間の一致性を評価し画像認識を行い、認識結果を報告する。次に認識率を向上させるために最小自乗法を二度行い、その評価を行い、実験結果を報告する。最後に、本論文をまとめる。

2. 静止画像の固有パターン

2.1 画像の色成分

計算機に取り込まれた画像は画素の集合であり、各画素は R、G、B 色成分で構成される。すなわち、 m 、 n をそれぞれ x 、 y 直行座標系の x 、 y 方向の画素数とすれば、画像データは $m \times n$ 次のベクトルとして扱うことができる。ベクトルとして表現されたカラー画像データにおける第 k 番目の画素の R、G、B 成分を R_k 、 G_k 、 B_k とすると、画素は式(1)で表される。

$$\begin{aligned} image &\in pixel_{i,j} \\ pixel_{i,j} &\in R_k, G_k, B_k \\ i &= 1, 2, 3, \dots, m \\ j &= 1, 2, 3, \dots, n \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \quad (1)$$

2.2 カラー画像の固有ベクトル

カラー画像の固有パターンは、画像中の R 成分の強度値、G 成分の強度値、B 成分の強度値で与えられ、1 画像に対して 3 組得られる。R 成分分布を I_R とすると、

$$\begin{aligned} I_R &= R_k \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \quad (2)$$

で与えられる。次に R 成分分布 I_R をダイナミックレンジ D の分布に正規化した正規化 R 成分分布 I_R^D を式(3)で定義する。 $Round[*]$ は括弧内の値を整数化する演算、 $Max[*]$ は括弧内の集合における最大値の要素を求める演算を示す。

$$\begin{aligned} I_R^D &\in Round \left[D \times \frac{R_k}{Max[I_R]} \right] \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \quad (3)$$

次に、式(3)で求められたダイナミックレンジ D の正規化 R 成分分布 I_R^D において 1 から D までの数値の出現頻度を求めヒストグラム化する。 $Count[* , p]$ は括弧内の集合における数値 p の個数を数える演算である。

$$\begin{aligned} E_R &\in Count[I_R^D, p] \\ p &= 1, 2, 3, \dots, D \end{aligned} \quad (4)$$

式(4)で求められた E_R を R 成分固有ベクトルという。G 成分、B 成分についても同様にして固有ベクトルを算出する。

2.3 固有パターン

固有パターンを E とすると固有パターン E は R 成分固有パターン、G 成分固有パターン、B 成分固有パターンの 3 個の固有パターンの集合からなり式(5)で与えられる。図 1 に静止画像の固有パターンの一例を示す。図 1 から、画像はそれぞれ独自の固有パターンを持つことが分かる。

$$E \in E_R, E_G, E_B \quad (5)$$

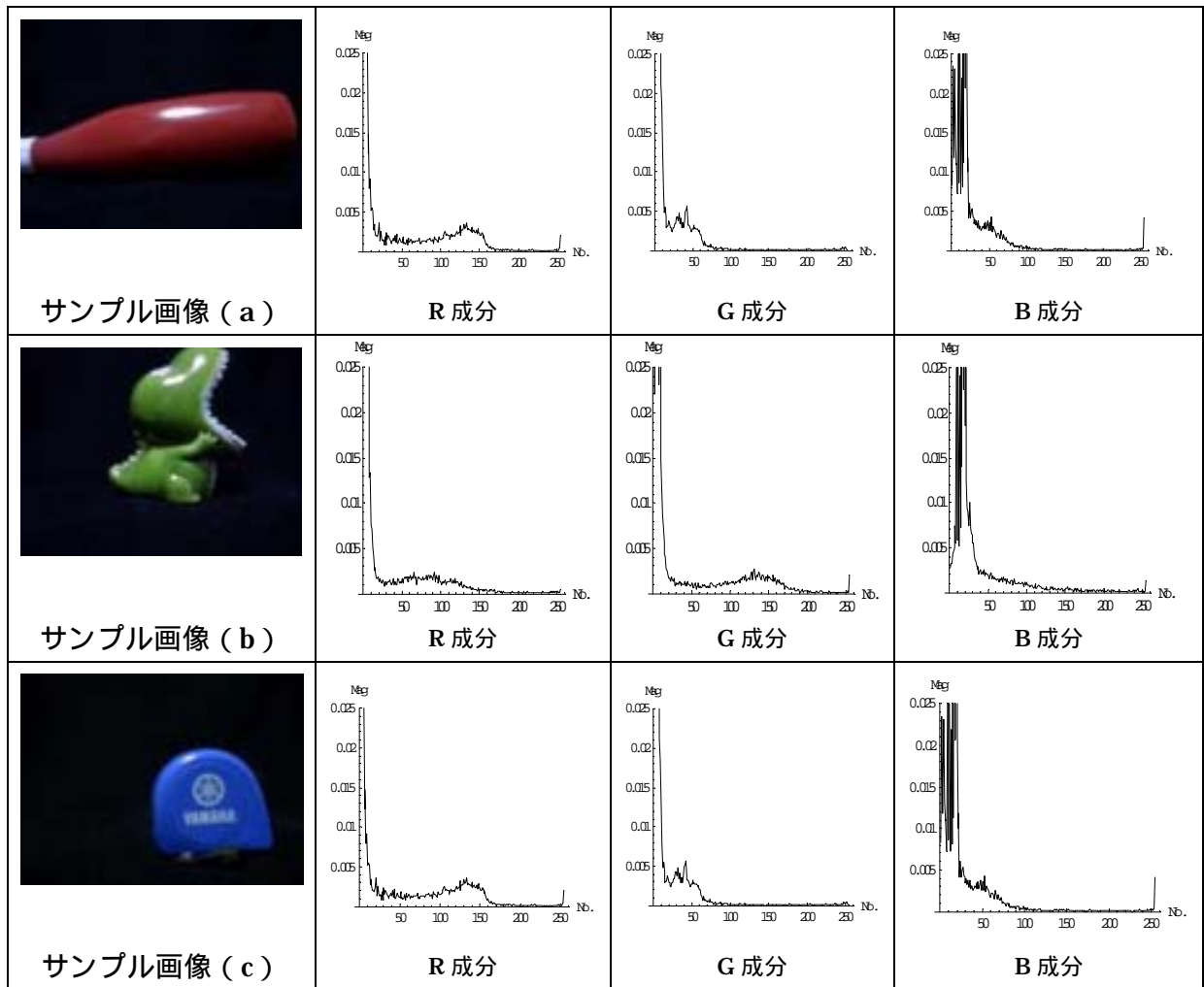


図 1 サンプル画像(a)(b)(c)の RGB 成分の固有パターン

3. 固有パターンの一貫性評価法

3.1 データベース

例として n 個の画像データから固有パターンを抽出しデータベースを構築する。

それぞれの画像データは R 成分固有パターン、G 成分固有パターン、B 成分固有パターンの 3 個の固有パターンは式(6)のシステム行列を構築可能とする。

$$\begin{aligned}
C_R &= [E_{R1}, E_{R2}, E_{R3}, \dots, E_{Rn}] \\
C_G &= [E_{G1}, E_{G2}, E_{G3}, \dots, E_{Gn}] \\
C_B &= [E_{B1}, E_{B2}, E_{B3}, \dots, E_{Bn}]
\end{aligned} \tag{6}$$

3.2 線形システム方程式

入力画像データの R、G、B 成分固有パターンをそれぞれ E_{YR}, E_{YG}, E_{YB} とすると R、G、B 成分の線形システム方程式は式(7)で表すことができる。 X_R, X_G, X_B は解ベクトルであり、式(6)を構成する固有パターンの重みに対応し、解ベクトルの要素で最大値をもつ要素が残りの要素に対して相対的に値が大きいほど、入力画像データと式(6)を構成する特定の固有パターンが等しいことを意味する。

$$\begin{aligned}
E_{YR} &= C_R \cdot X_R \\
E_{YG} &= C_G \cdot X_G \\
E_{YB} &= C_B \cdot X_B
\end{aligned} \tag{7}$$

3.3 解ベクトル

固有パターンの要素数はダイナミックレンジ D によって決まり、本論文では 0 から 255 の 256 段階の範囲で構成している。データベース数 n よりも固有パターンの要素数の方が大きいと仮定する。すなわち、式(7)のシステム方程式から解ベクトル X_R, X_G, X_B を求めることは不適切問題を解くことに帰する。この不適切問題の解法として最小自乗法を採用する。解ベクトル X_R, X_G, X_B の導出は式(8)で与えられる。

$$\begin{aligned}
X_R &= [C_R^T C_R]^{-1} C_R^T E_{YR} \\
X_G &= [C_G^T C_G]^{-1} C_G^T E_{YG} \\
X_B &= [C_B^T C_B]^{-1} C_B^T E_{YB}
\end{aligned} \tag{8}$$

式(8)より得られた R、G、B 解ベクトル X_R, X_G, X_B を畳み込み演算し、解ベクトル X を得る。*は畳み込み演算を示す。

$$X = X_R * X_G * X_B \tag{9}$$

式(9)で最大値を取る要素から静止画像は認識する。

3.4 静止画像認識

3.1、3.2、3.3 に示した固有パターンと線形システム方程式の静止画像認識手法を用いてデータベースと入力画像それぞれの固有パターン間の一致性を評価し画像認識を行い、精度の確認を行う。

3.4.1 データベース、入力画像

背景の色情報が認識に影響しないように、それぞれ同じ状況下で撮影を行った画像を利用する。図2にデータベース画像、図3に入力画像を示す。

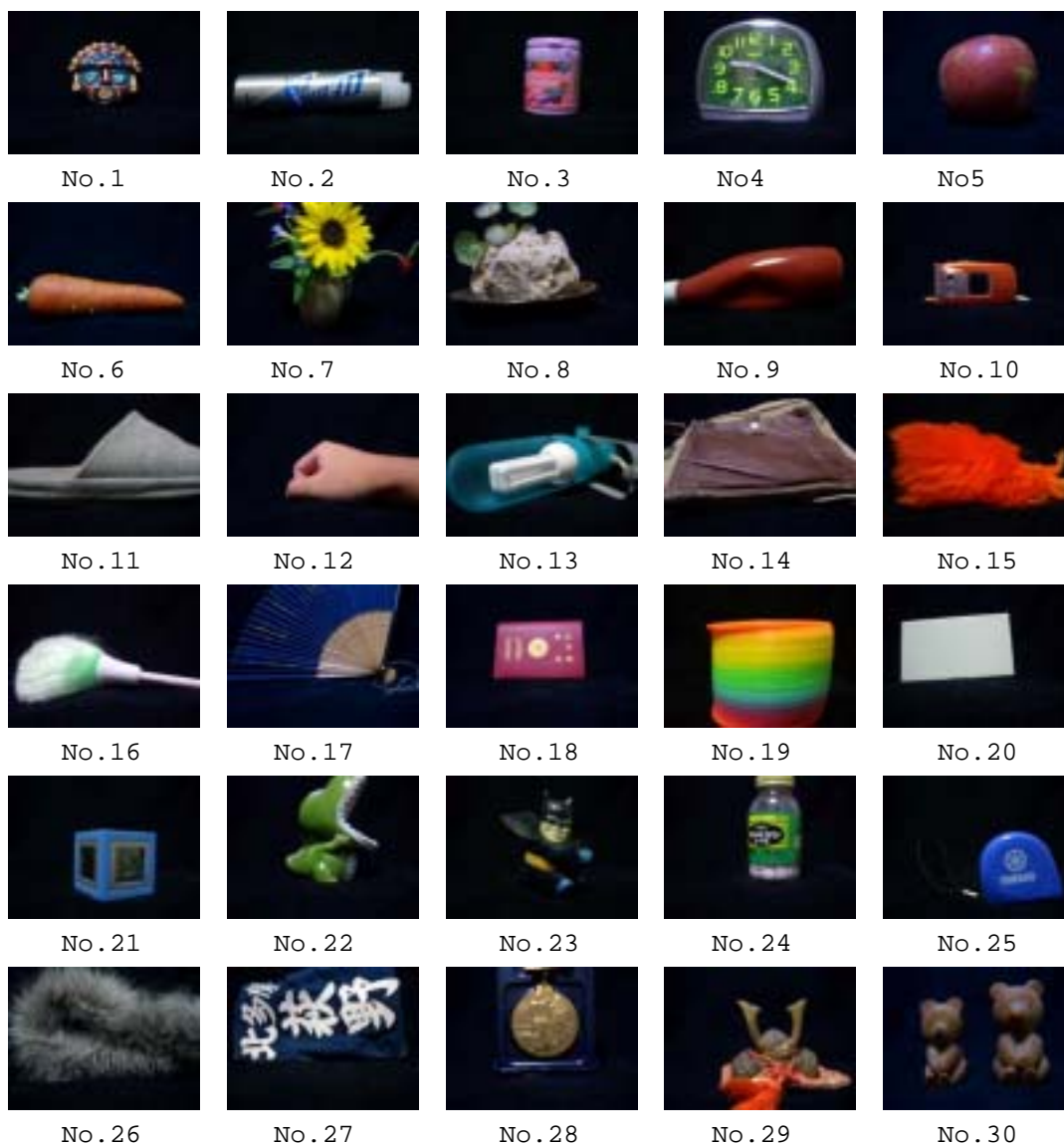


図2 データベース画像

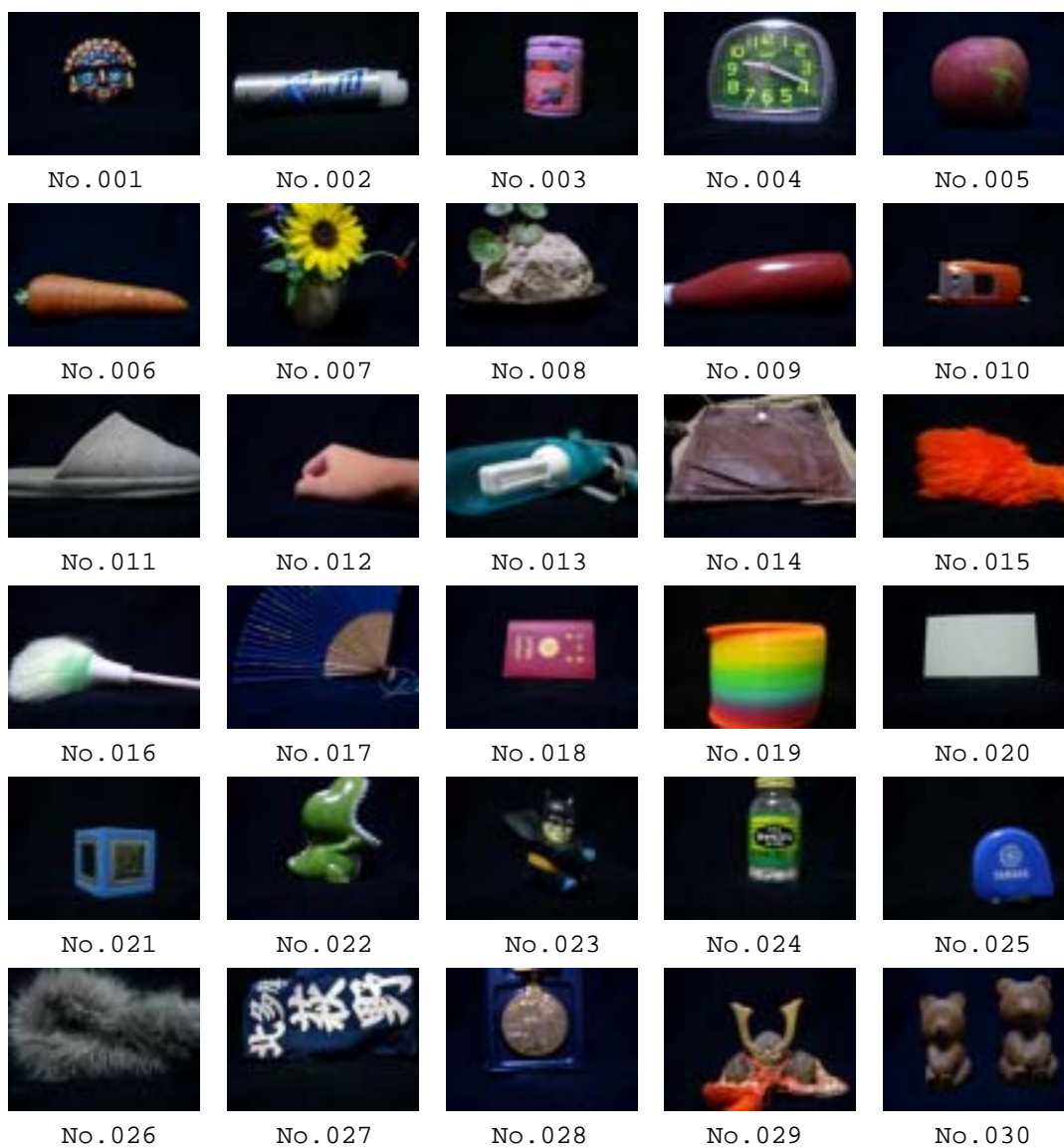


図3 入力画像

3.4.2 認識結果

次に認識結果を表1に示す。

表1

入力画像	出力画像	判定	入力画像	出力画像	判定
No.001	No.1		No.016	No.8	×
No.002	No.2		No.017	No.28	×
No.003	No.3		No.018	No.18	
No.004	No.4		No.019	No.19	
No.005	No.5		No.020	No.23	×
No.006	No.6		No.021	No.21	
No.007	No.7		No.022	No.22	
No.008	No.8		No.023	No.21	×
No.009	No.12	×	No.024	No.21	×
No.010	No.10		No.025	No.25	
No.011	No.11		No.026	No.26	
No.012	No.12		No.027	No.27	
No.013	No.13		No.028	No.18	×
No.014	No.14		No.029	No.29	
No.015	No.15		No.030	No.30	

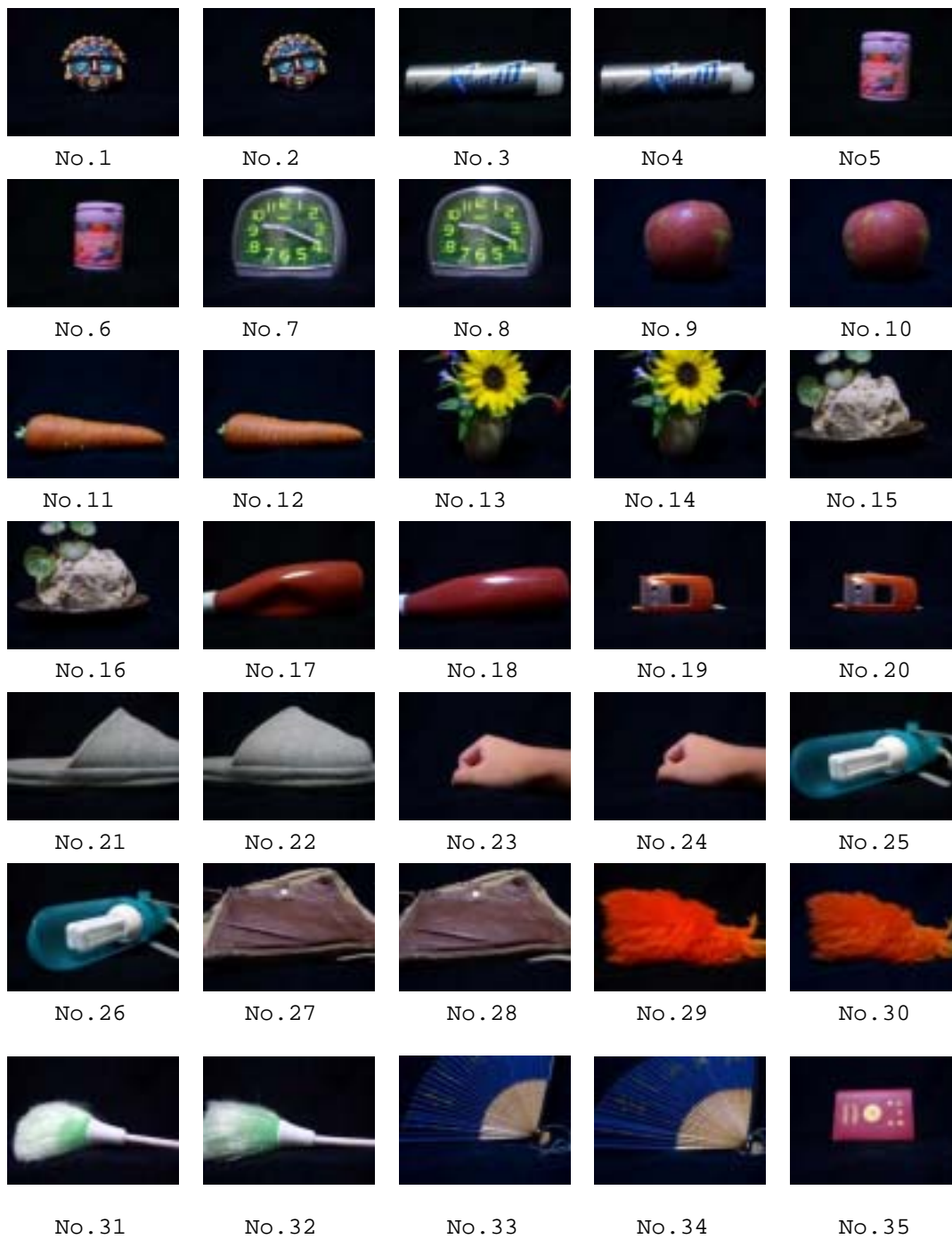
表1により30枚中23枚が認識成功となり、認識率76.7%であった。しかし画像認識と言うには80%以上の確率がなければシステムとしては不十分である。そこで次に認識率を上げるため、データ画像一種類につき二枚用意し、入力画像の特徴を現すデータ数を二倍に増やし認識させた。

3.5 最小自乗法による認識の向上

データ画像を30種類それぞれ2枚ずつ計60枚用意し、認識を行った。

3.5.1 データベース、入力画像

データベース画像は図4、入力画像は図3を使用する。



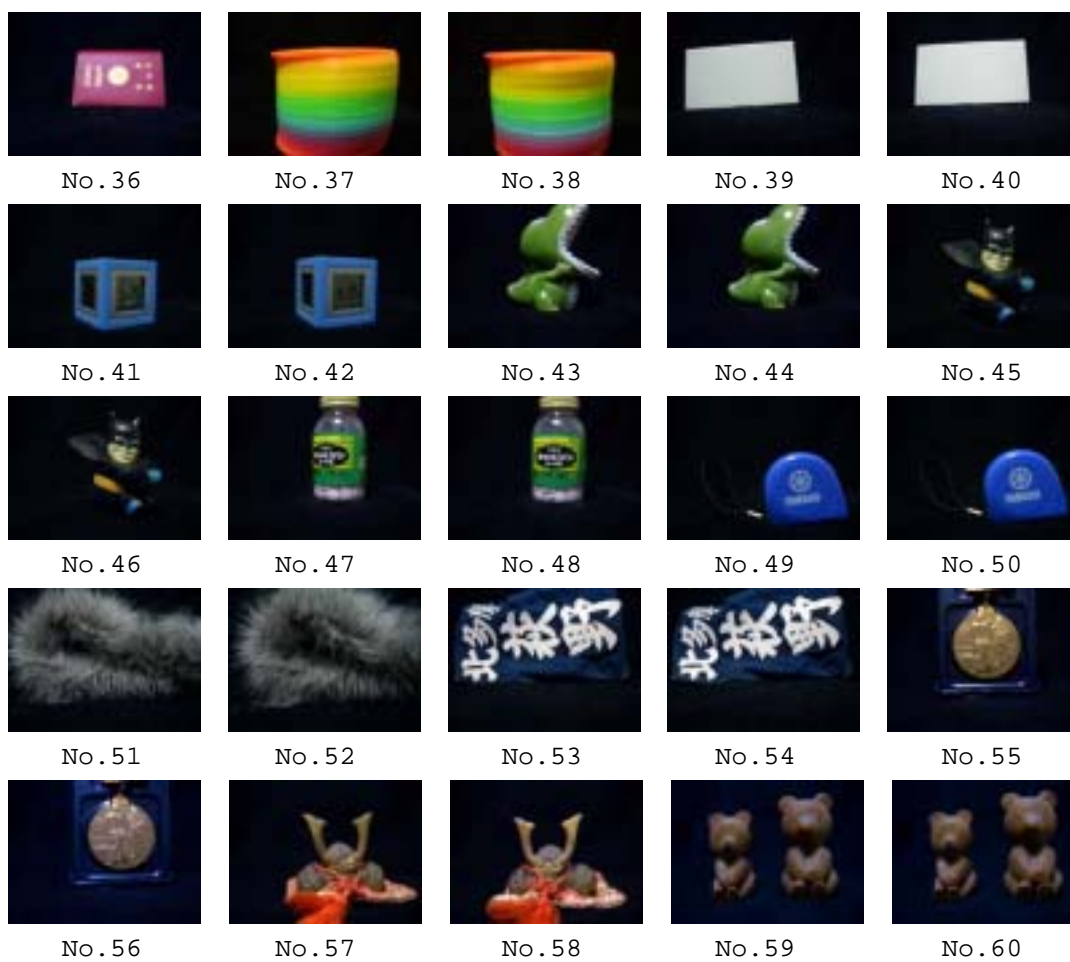


図4 データベース画像

3.5.2 認識結果

表2

入力画像	出力画像	判定	入力画像	出力画像	判定
No.001	No.2		No.016	No.32	
No.002	No.3		No.017	No.10	×
No.003	No.5		No.018	No.35	
No.004	No.7		No.019	No.37	
No.005	No.9		No.020	No.48	×
No.006	No.11		No.021	No.42	
No.007	No.13		No.022	No.44	
No.008	No.16		No.023	No.46	
No.009	No.18		No.024	No.46	×
No.010	No.57	×	No.025	No.50	
No.011	No.21		No.026	No.51	
No.012	No.23		No.027	No.14	×
No.013	No.25		No.028	No.56	
No.014	No.8	×	No.029	No.58	
No.015	No.48	×	No.030	No.60	

表2により30枚中23枚が認識成功となり、認識率76.7%であった。これは3.4.2の認識結果と同じ認識率であり、データ数を二倍に増やしても認識率が上がらないことがわかった。しかしこれはこのデータによる影響が大きいので、他の撮影画像であれば認識率は変わるかもしれない。

3.6 二回の最小自乗法による認識の評価（その1）

認識率を上げるため、一度の最小自乗法で入力画像の固有パターンに近いデータ画像を解ベクトル大きい順に何枚かに絞り、さらにそこから最小自乗法を利用し、認識結果を出力した。そして何枚に絞れば一番良い認識率がでるのか実験してみた。今回はデータ画像の偶数枚について行った。

3.6.1 データベース、入力画像

データベース画像は図2、入力画像は図3を使用する。

3.6.2 認識結果

表3

絞った枚数(枚)	認識率(%)
2	60.0
4	60.0
6	63.3
8	66.7
10	66.7
12	66.7
14	70.0
16	70.0
18	66.7
20	73.3
22	73.3
24	73.3
26	73.3
28	76.7

実験の結果、最高でも28個に絞っての76.7%であった。最小自乗法を二度使っても認識率が上がらなかった。これはデータ数が少ないため、入力画像の固有パターンに近い画像を絞ることが困難だったのではないかと考えられる。しかし表3を見ると一回目に絞る数が大きくなれば、認識率が上昇していることがわかる。

3.7 二回の最小自乗法による認識の評価（その2）

データ数が少なかったと考えられるので、データ画像30種類をそれぞれ2枚ずつ計60枚用意し、評価した。今回はデータ画像の4の倍数枚で行った。

3.7.1 データベース、入力画像

データベース画像は図4、入力画像は図3を使用する。

3.7.2 認識結果

表4

入力画像	一度目の最小自乗法で絞る解ベクトルの大きい画像の枚数(枚)													
	4	8	12	16	20	24	28	32	36	40	44	48	52	56
No.001														
No.002														
No.003	×													
No.004														
No.005														
No.006														
No.007														
No.008														
No.009														
No.010	×	×			×	×	×	×		×	×	×	×	
No.011														
No.012														
No.013	×			×										
No.014														×
No.015									×	×		×	×	×
No.016				×	×	×	×	×	×	×	×	×	×	
No.017					×	×	×		×	×	×		×	×
No.018														
No.019														
No.020	×	×	×						×	×	×	×	×	×
No.021														
No.022														
No.023	×													
No.024		×	×	×	×	×	×	×	×	×	×	×	×	×
No.025														
No.026														
No.027														×
No.028														
No.029														
No.030														
認識率(%)	83.3	90.0	93.3	90.0	86.7	86.7	86.7	90.0	83.3	80.0	83.3	83.3	80.0	80.0

実験の結果、一度目の最小自乗法でデータ画像を12枚に絞った時が認識率93.3%と最高の認識率になった。3.6の結果76.7%と比べても高い認識率が得られた。二回の最小自乗法を使う場合、データ画像を増やすと大幅に認識率が上がることがわかった。表4の認識率を見てもわかるが、12枚のときから離れるほど認識率が段々と落ちている。これはデータ画像60枚のうちで12枚に絞るという割合が最小自乗法を二回使う場合、一番良い認識結果が出るということである。つまりデータ画像に対して絞る枚数を20%にすると、高い認識率である。

4. プログラム

ここに今回使用した基本的な Mathematica のプログラムを載せておく。

<固有パターン>

```
Off[General::spell,General::spell1,MemoryConserve::start,MemoryConserve::end];
<<LinearAlgebra`MatrixManipulation`;
<<Graphics`;
<<Graphics`PlotField`

SetDirectory["C:¥Images"]
dog=FileNames["*.jpg"]

pic=Import[dog[[1]]];
Show[pic,ImageSize→200,AspectRatio→Automatic];

mat=InputForm[pic];
picRGB=mat[[1,1,1]];
dim=Dimensions[picRGB]
rgbData=Transpose[picRGB,{2,3,1}];

ListDensityPlot[rgbData[[1]],Mesh→False,Frame→False];
ListDensityPlot[rgbData[[2]],Mesh→False,Frame→False];
ListDensityPlot[rgbData[[3]],Mesh→False,Frame→False];
```

```

compRf=Flatten[rgbData[[1]]];
compGf=Flatten[rgbData[[2]]];
compBf=Flatten[rgbData[[3]]];

countR=Table[Count[compRf,i],{i,255}];
countG=Table[Count[compGf,i],{i,255}];
countB=Table[Count[compBf,i],{i,255}];

countR=countR/(dim[[1]]*dim[[2]]);
countG=countG/(dim[[1]]*dim[[2]]);
countB=countB/(dim[[1]]*dim[[2]]);

ListPlot[countR,PlotRange->{0,0.025},PlotJoined->True,AxesLabel->{"No. ", "
Mag"}];

ListPlot[countG,PlotRange->{0,0.025},PlotJoined->True,AxesLabel->{"No. ", "
Mag"}];

ListPlot[countB,PlotRange->{0,0.025},PlotJoined->True,AxesLabel->{"No. ", "
Mag"}];
eigenRGB=Table[Flatten[{countR,countG,countB}]];

ListPlot[eigenRGB,PlotRange->{0,0.025},PlotJoined->True,AxesLabel->{"No. "
,"Mag"}];

```


< 最小白乘法 >

```
Off[General::spell,General::spell1,MemoryConserve::start,MemoryConserve::end];
<<LinearAlgebra`MatrixManipulation`;
<<Graphics`;
<<Graphics`PlotField`;

SetDirectory["C:¥Temp¥images06¥data"]
image=FileNames["*.jpg"]
datedim=image//Dimensions

Do[pic[n]=Import[image[[n]],{n,(datedim[[1]])}];
Do[mat[n]=InputForm[pic[n]],{n,(datedim[[1]])}];
Do[picRGB[n]=mat[n][[1,1,1]],{n,(datedim[[1]])}];
Do[dim[n]=Dimensions[picRGB[n]],{n,(datedim[[1]])}];
Do[rgbData[n]=Transpose[picRGB[n],{2,3,1}],{n,(datedim[[1]])}];

Do[compRf[n]=Flatten[rgbData[n][[1]]],{n,(datedim[[1]])}];
Do[compGf[n]=Flatten[rgbData[n][[2]]],{n,(datedim[[1]])}];
Do[compBf[n]=Flatten[rgbData[n][[3]]],{n,(datedim[[1]])}];
Do[countR[n]=Table[Count[compRf[n],i],{i,255}],{n,(datedim[[1]])}];
Do[countG[n]=Table[Count[compGf[n],i],{i,255}],{n,(datedim[[1]])}];
Do[countB[n]=Table[Count[compBf[n],i],{i,255}],{n,(datedim[[1]])}];
Do[countR[n]=countR[n]/(dim[n][[1]]*dim[n][[2]]),{n,(datedim[[1]])}];
Do[countG[n]=countG[n]/(dim[n][[1]]*dim[n][[2]]),{n,(datedim[[1]])}];
Do[countB[n]=countB[n]/(dim[n][[1]]*dim[n][[2]]),{n,(datedim[[1]])}];

Do[eigenRGB[n]=Table[Flatten[{countR[n],countG[n],countB[n]}],{n,(datedim[[1]])}];

SetDirectory["C:¥Temp¥images06¥in"]
imagein=FileNames["*.jpg"]
picin=Import[imagein[[1]]];
Show[picin,ImageSize→200,AspectRatio→Automatic]
```

```

matin=InputForm[picin];
picinRGB=matin[[1,1,1]];
dimin=Dimensions[picinRGB]
rgbDatain=Transpose[picinRGB,{2,3,1}];
compRfin=Flatten[rgbDatain[[1]]];
compGfin=Flatten[rgbDatain[[2]]];
compBfin=Flatten[rgbDatain[[3]]];
countRin=Table[Count[compRfin,i],{i,255}];
countGin=Table[Count[compGfin,i],{i,255}];
countBin=Table[Count[compBfin,i],{i,255}];
countRin=countRin/(dimin[[1]]*dimin[[2]]);
countGin=countGin/(dimin[[1]]*dimin[[2]]);
countBin=countBin/(dimin[[1]]*dimin[[2]]);
eigenRGBin=Table[Flatten[{countRin,countGin,countBin}]];

ct=Table[eigenRGB[n],{n,(datedim[[1]])}];
cc=Transpose[ct];
inctcc=Inverse[ct.cc];
X=inctcc.ct.eigenRGBin;
xmax=Ordering[X,-1]

Show[pic[(xmax[[1]])],ImageSize->200,AspectRatio->Automatic];

```

5. まとめ

本論文では、まず画像の普遍量が R、G、B 成分より構成されていることを述べ、画像の固有パターン抽出手法を述べた。画像の普遍量を抽出する方法として、固有パターンを導入し、画像を構成する画素数、画像の x y 直交座標上の位置・角度といったハードウェアに依存する性質を削減することが可能となり、画像本来の情報を抽出する一方法を確立した。

入力画像から抽出された固有パターンを線形システム方程式の入力ベクトルとして解が求まる。線形システム方程式は不適切であるため、最小自乗法を用いて近似解を導いた。そして認識率を上げるため一度目の最小自乗法で解ベクトルの大きいものをデータ

画像から数を絞って選び、そしてその中からさらに最小自乗法を用い一枚を出力させた。そしてその絞る枚数について評価を行った。

その結果、データ画像 30 種類 30 枚の場合、絞る枚数を多くすれば認識率が高くなっていたが、データ画像 30 種類をそれぞれ 2 枚ずつの 60 枚の場合、絞る数を 12 枚に絞ると高い認識率を得た。前者の実験では入力画像に対してのデータが少なかったために、絞る数が多ければデータが多くなるので、3.6.2 の様な結果が出たのだと考えられる。一方後者は、データの数は入力画像を認識させうるのに十分なデータ量があったので、データを何枚に絞れば高認識率が出るのか、という問題についてひとつの傾向が実験結果として示されたと考えられる。

6. 参考文献

- [1] 佐藤 隆紀 「新画像認識法の提案とその磁界系非破壊検査への応用」
2002 年度 法政大学卒業論文
- [2] 寺西 正晃 「相関係数を用いた合成画像法による画像認識」
2004 年度 法政大学卒業論文

2005年度

卒業論文

色情報を利用した静止画像認識における信頼度
向上に関する考察と認識手法比較

法政大学

工学部 情報電気電子工学科

学籍番号 02d3091

氏名 星野賢治

指導教員 斎藤兆古

目次

1.	序論	3
2.	静止画像の固有パターン	5
2.1	画像の色成分	5
2.2	カラー画像の固有ベクトル	5
2.3	固有パターン	6
3.	最小自乗法による画像認識	8
3.1	線形システム方程式	8
3.2	解ベクトル	8
3.3	静止画像認識	9
4.	GVSPM 法による画像認識	19
4.1	GVSPM 法の導出	19
4.2	静止画像認識	21
5.	相関係数法による画像認識	24
5.1	相関係数の導出	24
5.2	静止画像認識	24
6.	RGB 成分の独立認識における画像合成認識	27
6.1	静止画像認識	27
7.	まとめ	31
8.	参考文献	32

1. 序論

現在、画像を計算機によって処理・加工するには人間が評価することを前提としている。この意味で、人間の視覚能力を前提として画像をハンドリングする画像処理は比較的多く実用化されている。しかし、映像を認識する技術は完成度が低く、最終的には人間の視覚情報処理能力に依存せざるを得ない。例を挙げると、同一被写体をデジタルカメラで撮影した場合、撮影条件が完全に同一でない限り同一の画像を表す数値データをとらない、すなわち、機械は同一の被写体と認識できない。換言すれば、まったく同一の画像のみ認識できる。最新の計算機をもってしても、画像認識の分野においては人間の視覚情報処理能力には到底及ばない。映像を計算機によって認識させることは、計算機に人間と同等な視覚情報処理能力を与えることになり、現状の様々な手法では人間の視覚情報処理能力には匹敵し難い。このような現状をふまえて、本論文では、電子的に計算機スクリーン上に可視化された映像から、画素数、配置、変形等に無関係となる映像の普遍量を抽出する概念を導入し、計算機に人間の視覚情報処理能力を与え、計算機自身によって映像を認識させることを試みる。

ここで我々が認識の対象とする画像は人間の視覚によって認識しえる形態で表現された情報をいい、それらを総称して本論文では可視化情報と称する。その典型的な例として、人間の視覚によって感知できるカメラやビデオで撮影された画像などがある。また、人間の視覚情報処理機能で捉えることができない情報を映像化した赤外線映像や電子顕微鏡による映像等も可視化情報の一種である。さらに人間の視覚で認識可能な形態へ変形された信号情報、例えば、音声等を電子的にオシロスコープ等の機器で映像化した情報もその一種である。

人間の脳の視聴覚情報機能は、眼球を通して得られた映像を処理することによって情報を認識するが、その対象となる視覚情報には、一定の構成ルールに従ってコード化された文字や言語等のコード化視聴覚情報と、コード化されていない可視化情報、すなわち非コード化視聴覚情報がある。先に述べた可視化情報は、非コード化視覚情報を処理する人間の視覚情報処理能力を前提としたデータであり、可視化情報から計算機で処理可能な何らかの普遍量を抽出する必要がある。

計算機のスクリーン上の可視化情報は2次元平面上の画素 (Pixel) で構成されている。各画素は可視光の波長によってそれぞれ赤、緑、及び青 (以下それぞれ R,G,B と略記) 成分の情報を持つ。また、可視化情報は反射光の強さ (Intensity) 色相 (Tone) ,及び色成分 (Color component) の組み合わせからなり、それらの幾何学的配置で表現される。そのため、計算機のスクリーン上の可視化情報は、スクリーンの解像度や基準座標 (視点) に依存する性質を持つ。このようなハードウェアに依存する性質を削減した可視化情報の普遍量を本論文では画像の固有パターン (Eigen pattern) と呼び、前述の普遍量はこれに相

当し、この固有パターンを採用し画像認識を試みる。

本論文で採用する画像認識手法は線形システム方程式の解を導くことに基づく方法である。しかしながら、得られた線形システム方程式は不適切なシステム方程式であるため、最小二乗法及びGVSPM法(Generalized Vector Sampled Pattern Matching)を用いて近似解を導く。言い換えれば、本論文で採用する画像認識手法は、あらかじめ用意した画像データをデータベースとしその中からなんらかの手法により解要素を求める方法である。

本論文の構成は最初に静止画像から固有パターンを抽出する方法について述べる。次に抽出された固有パターンの一致度の評価法について述べる。すなわち、最小二乗法、GVSPM法、相関係数を用いてデータベースと入力画像それぞれの固有パターンの一致性を評価し画像認識を行い、認識結果及び手法間の相違、比較を報告する。最後に、本論文をまとめる。

2. 静止画像の固有パターン

2.1 画像の色成分

計算機に取り込まれた画像はスクリーン上の画素で構成されている。各画素は赤 (Red,R) 緑 (Green,G), および青 (Blue,B) の色情報で構成されている。よって、いま、計算機ディスプレイのスクリーンを X-Y 直交座標系で表現するとすれば、 m 、 n をそれぞれ x 、 y 直交座標系の x 、 y 方向の画素数とすれば、画像データは $m \times n$ 次のベクトルとして扱うことができる。ベクトルとして表現されたカラー画像データにおける第 k 番目の各画素の R,G,B 成分を R_k, G_k, B_k とすると、画素は式(1)と表される。

$$\begin{aligned} image &\in pixel_{i,j} \\ pixel_{i,j} &\in R_k, G_k, B_k \\ i &= 1, 2, 3, \dots, m \\ j &= 1, 2, 3, \dots, n \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \tag{1}$$

2.2 カラー画像の固有ベクトル

カラー画像の固有パターンの算出はまず、画像中の R 成分の強度値、G 成分の強度値、B 成分の強度値で与えられ、1 画像に対して 3 組得られる。R 成分分布を I_R とすれば、

$$\begin{aligned} I_R &= R_k \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \tag{2}$$

と与えられる。次に R 成分分布 I_R をダイナミックレンジ D の分布に正規化した正規化 R 成分分布 I_R^D を式(3)で定義する。下式の $Round[*]$ 、 $Max[*]$ はそれぞれ括弧内の値を整数化する演算、括弧内の集合における最大値の要素を求める演算である。

$$\begin{aligned} I_R^D &\in Round \left[D \times \frac{R_k}{Max[I_R]} \right] \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \tag{3}$$

さらに式(3)で求めたダイナミックレンジ D の正規化 R 成分分布 I_R^D において 1 から D まで

の数値の出現頻度を算出しヒストグラムにする。 $Count[* , p]$ は括弧内の集合における数値 p の個数を考える演算である。

$$\begin{aligned} E_R \in Count[I_R^D, p] \\ p = 1, 2, 3, \dots, D \end{aligned} \quad (4)$$

式(4)で求めた E_R を R 成分固有ベクトルといい、同様に G 成分、B 成分も固有ベクトルを算出する。

2.3 固有パターン

固有パターンを E とすると固有パターン E は先に述べたように R 成分固有パターン、G 成分固有パターン、B 成分固有パターンの 3 種の固有パターンの集合からできている。式 (5) で与えられる。

$$E \in E_R, E_G, E_B \quad (5)$$

Fig.1 に静止画像の固有パターンの一例を示す。Fig.1 から、画像特有のパターンが存在することがわかる。

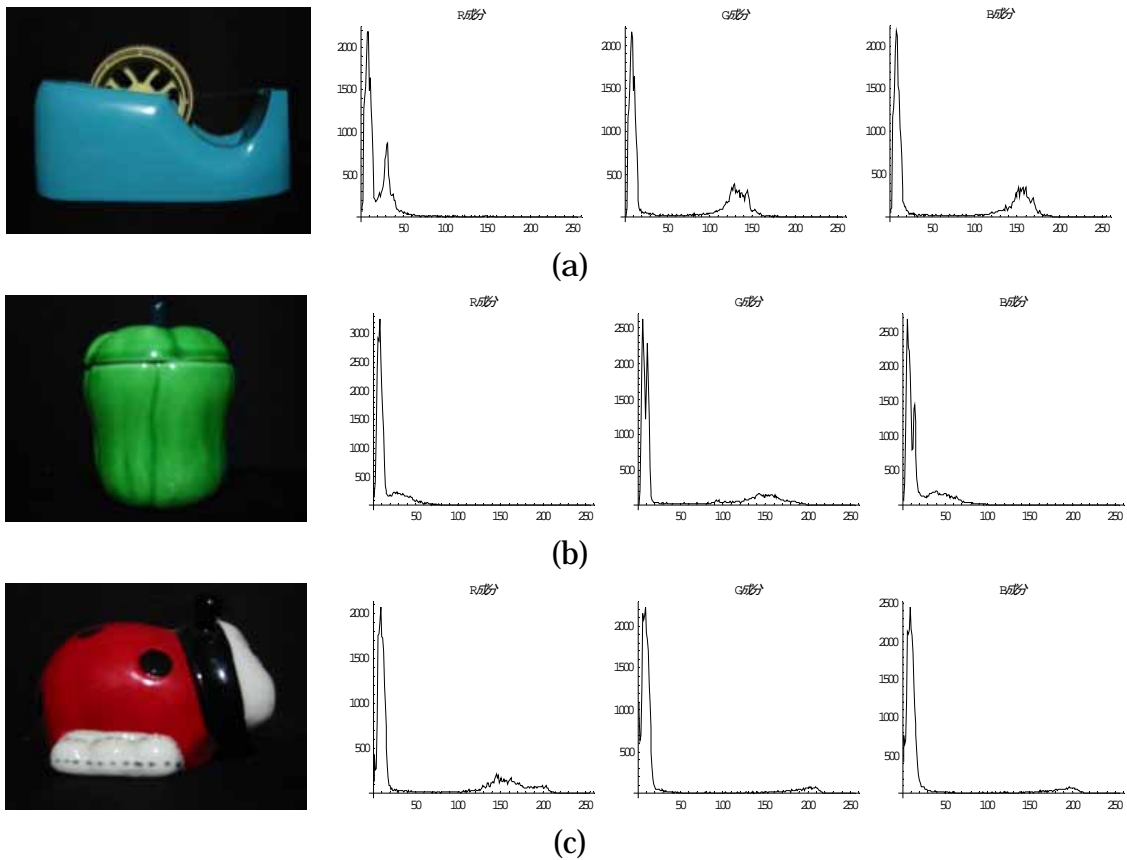


Fig.1 画像の固有パターン

Fig.1 からわかるようにそれぞれの画像が固有の赤・緑・青のヒストグラムパターン、すなわち、RGB 固有パターンをもつことがわかる。

3. 最小二乗法による画像認識

3.1 線形システム方程式

例として n 個の画像データから固有パターンを抽出しデータベースを構築する。それぞれの画像の固有パターンである R 成分固有パターン、G 成分固有パターン、B 成分固有パターンで式(6)のようなシステム行列を構築する。

$$\begin{aligned} C_R &= [E_{R1}, E_{R2}, E_{R3}, \dots, E_{Rn}] \\ C_G &= [E_{G1}, E_{G2}, E_{G3}, \dots, E_{Gn}] \\ C_B &= [E_{B1}, E_{B2}, E_{B3}, \dots, E_{Bn}] \end{aligned} \quad (6)$$

そして入力画像データ R 成分固有パターン、G 成分固有パターン、B 成分固有パターンをそれぞれ E_{YR}, E_{YG}, E_{YB} 、解ベクトルを X_R, X_G, X_B とすると R,G,B 成分の線形システム方程式は式 (7) で表すことができる。ここで解ベクトル X_R, X_G, X_B の要素で最大値をもつ要素が残りの要素に対して相対的に値が大きいほど、入力画像データと式 (6) を構成するデータベースの特定の固有パターンが等しい、すなわち、同じ画像の可能性が高いということの意味する。

$$\begin{aligned} E_{YR} &= C_R \cdot X_R \\ E_{YG} &= C_G \cdot X_G \\ E_{YB} &= C_B \cdot X_B \end{aligned} \quad (7)$$

3.2 解ベクトル

固有パターンの要素数はダイナミックレンジ D によって決まり、本論文では 0 から 255 の 256 段階の範囲で構成している。データベースの数がこの固有パターンの要素数より少ない場合、式(7)のシステム方程式から解ベクトル X_R, X_G, X_B を算出することは不適切問題を解くことに帰する。この不適切問題を本章では最小二乗法によって解く。最小二乗法はこういった不適切問題を解く時に誤差ノルムを最小とする解を適用し、最も適切な解を算出する手法である。この最小二乗法はこのような不適切なシステム方程式を解くことにあたっては最良な手法の一つである。解ベクトル X_R, X_G, X_B は以下のように導出する。

まず、式(7)のようなシステム方程式を与える。

$$\begin{aligned} E_1 &= C_1 \cdot X_1 \\ E_2 &= C_2 \cdot X_2 \\ &\vdots \\ E_n &= C_n \cdot X_n \end{aligned} \quad (8)$$

誤差を e とし式(8)を展開する。

$$e = \|CX - E\| = \left\{ (C_1 X_1 - E_1)^2 + \dots + (C_n X_n - E_n)^2 \right\}^{\frac{1}{2}}$$

$$e^2 = (CX - E)^T (CX - E) = C^T C X^2 - 2C^T E X + E^T E \quad (9)$$

最小となる e^2 を求めるため微分する。

$$\frac{de^2}{dX} = 2C^T C X - 2C^T E \quad (10)$$

式(10)の解が 0 となる X を求める。

$$X = \frac{C^T E}{C^T C} \quad (11)$$

よって、式(11)から式(7)の解を求めると

$$X_R = [C_R^T C_R]^{-1} C_R^T E_{YR}$$

$$X_G = [C_G^T C_G]^{-1} C_G^T E_{YG} \quad (12)$$

$$X_B = [C_B^T C_B]^{-1} C_B^T E_{YB}$$

式(12)より得られた R,G,B の解ベクトル X_R, X_G, X_B を畳み込み演算し、解ベクトル X を得る。*は畳み込み演算を示す。式(13)で最大値を取る要素から静止画像を認識する。

$$X = X_R * X_G * X_B \quad (13)$$

3.3 静止画像認識

3.1,3.2 で述べた最小二乗法による線形システム方程式の静止画像認識手法を用いてデータベースと入力画像それぞれの固有パターン間の一致性を評価し認識し、精度を報告する。

3.3.1 入力画像、データベース

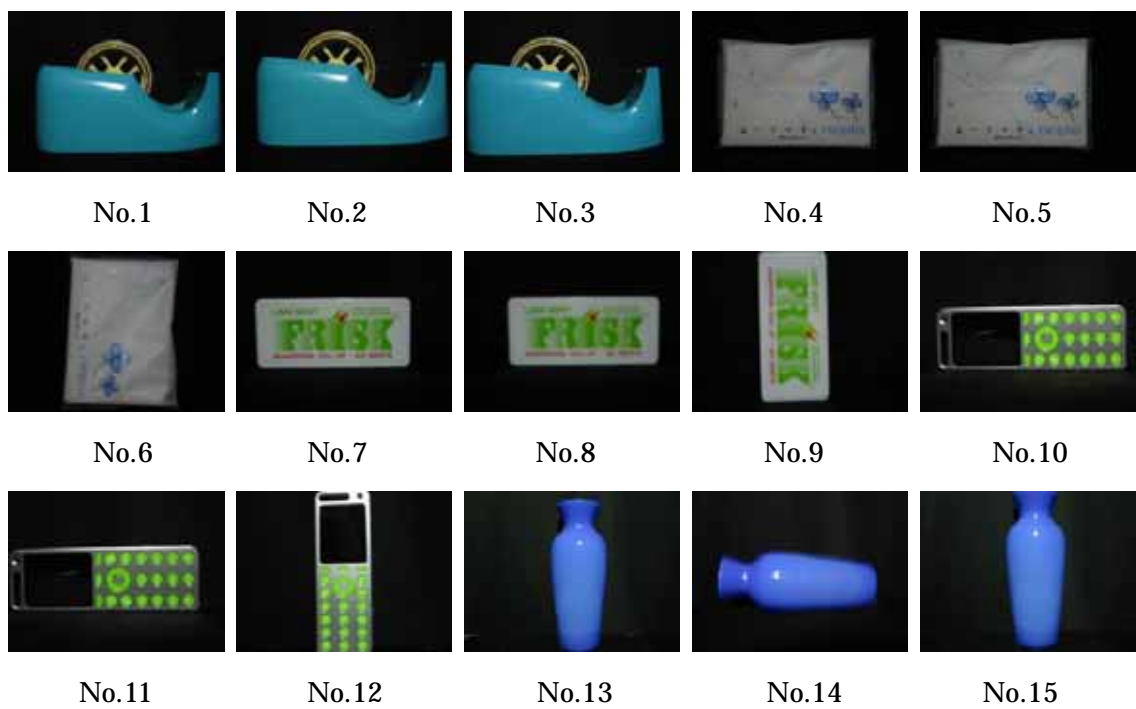
背景の情報が認識に作用しないように、白熱電球の下で暗室にて撮影をした画像によっ

て画像認識を行う。Fig.2 に入力画像を 50 枚、Fig.3 にデータベース画像を 150 枚を示す。





Fig.2 入力画像





No.16



No.17



No.18



No.19



No.20



No.21



No.22



No.23



No.24



No.25



No.26



No.27



No.28



No.29



No.30



No.31



No.32



No.33



No.34



No.35



No.36



No.37



No.38



No.39



No.40



No.41



No.42



No.43



No.44



No.45



No.46



No.47



No.48



No.49



No.50



No.51



No.52



No.53



No.54



No.55



No.56



No.57



No.58



No.59



No.60



No.61



No.62



No.63



No.64



No.65



No.66



No.67



No.68



No.69



No.70



No.71



No.72



No.73



No.74



No.75



No.76



No.77



No.78



No.79



No.80



No.81



No.82



No.83



No.84



No.85



No.86



No.87



No.88



No.89



No.90



No.91



No.92



No.93



No.94



No.95



No.96



No.97



No.98



No.89



No.100



No.101



No.102



No.103



No.104



No.105



No.106



No.107



No.108



No.109



No.110



No.111



No.112



No.113



No.114



No.115



No.116



No.117



No.118



No.119



No.120

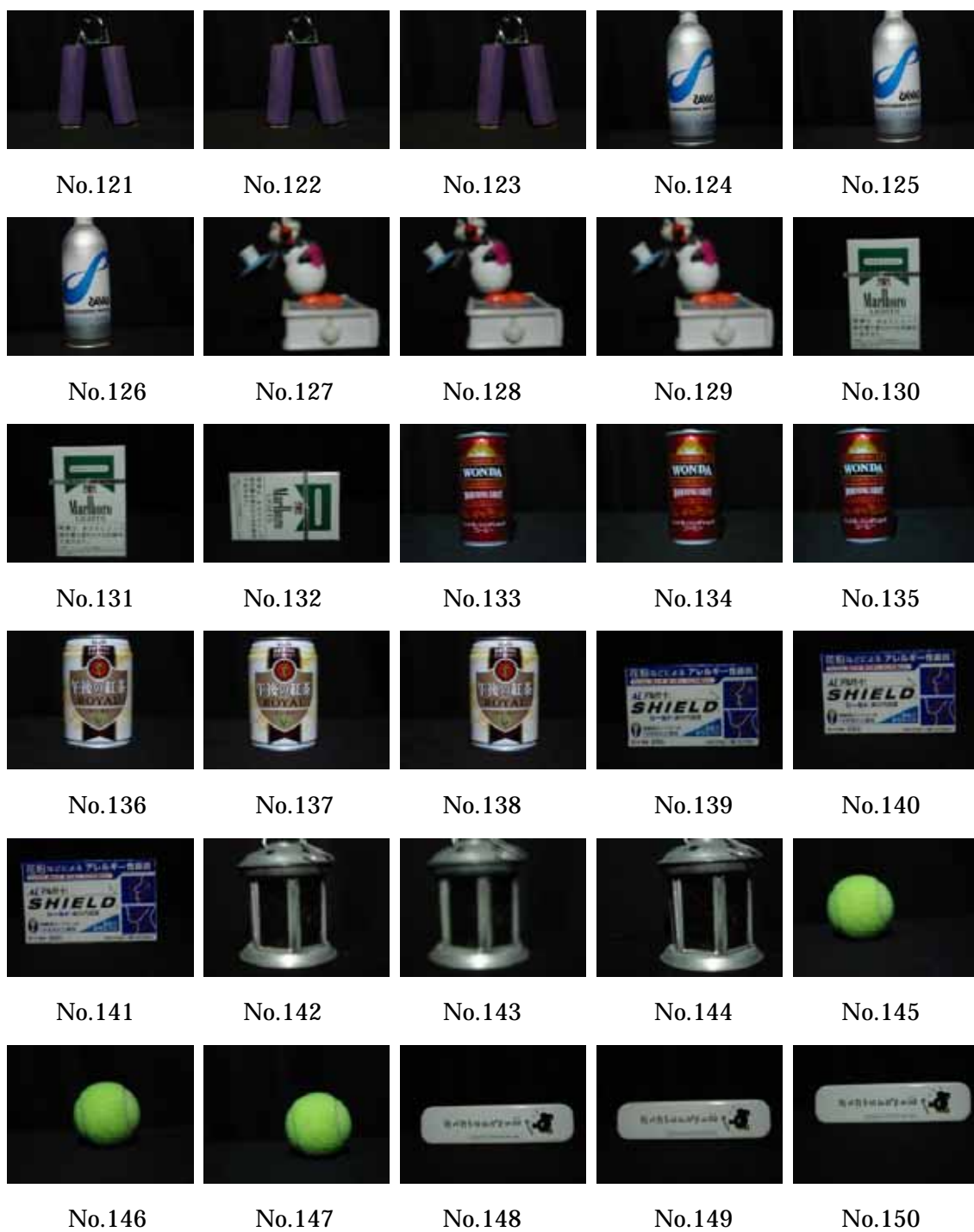


Fig.3 データベース画像

3.3.2 認識結果

認識結果を Table.1 に示す。出力 No.は Fig.3 のデータベース画像の No.に対応し、判定

○ は認識成功、× は認識失敗を示す。

Table.1

入力 No.	出力 No.	判定	入力 No.	出力 No.	判定
No.1	No.43	×	No.26	No.115	×
No.2	No.4		No.27	No.55	×
No.3	No.8		No.28	No.83	
No.4	No.12		No.29	No.85	
No.5	No.13		No.30	No.89	
No.6	No.16		No.31	No.91	
No.7	No.19		No.32	No.94	
No.8	No.22		No.33	No.98	
No.9	No.27		No.34	No.101	
No.10	No.122	×	No.35	No.104	
No.11	No.37	×	No.36	No.106	
No.12	No.34		No.37	No.109	
No.13	No.37		No.38	No.112	
No.14	No.133	×	No.39	No.117	
No.15	No.43		No.40	No.81	×
No.16	No.46		No.41	No.121	
No.17	No.50		No.42	No.124	
No.18	No.52		No.43	No.128	
No.19	No.133	×	No.44	No.130	
No.20	No.60		No.45	No.102	×
No.21	No.62		No.46	No.27	×
No.22	No.65		No.47	No.141	
No.23	No.67		No.48	No.144	
No.24	No.102	×	No.49	No.146	
No.25	No.74		No.50	No.149	

認識成功例と失敗例を Fig.4 に示す。

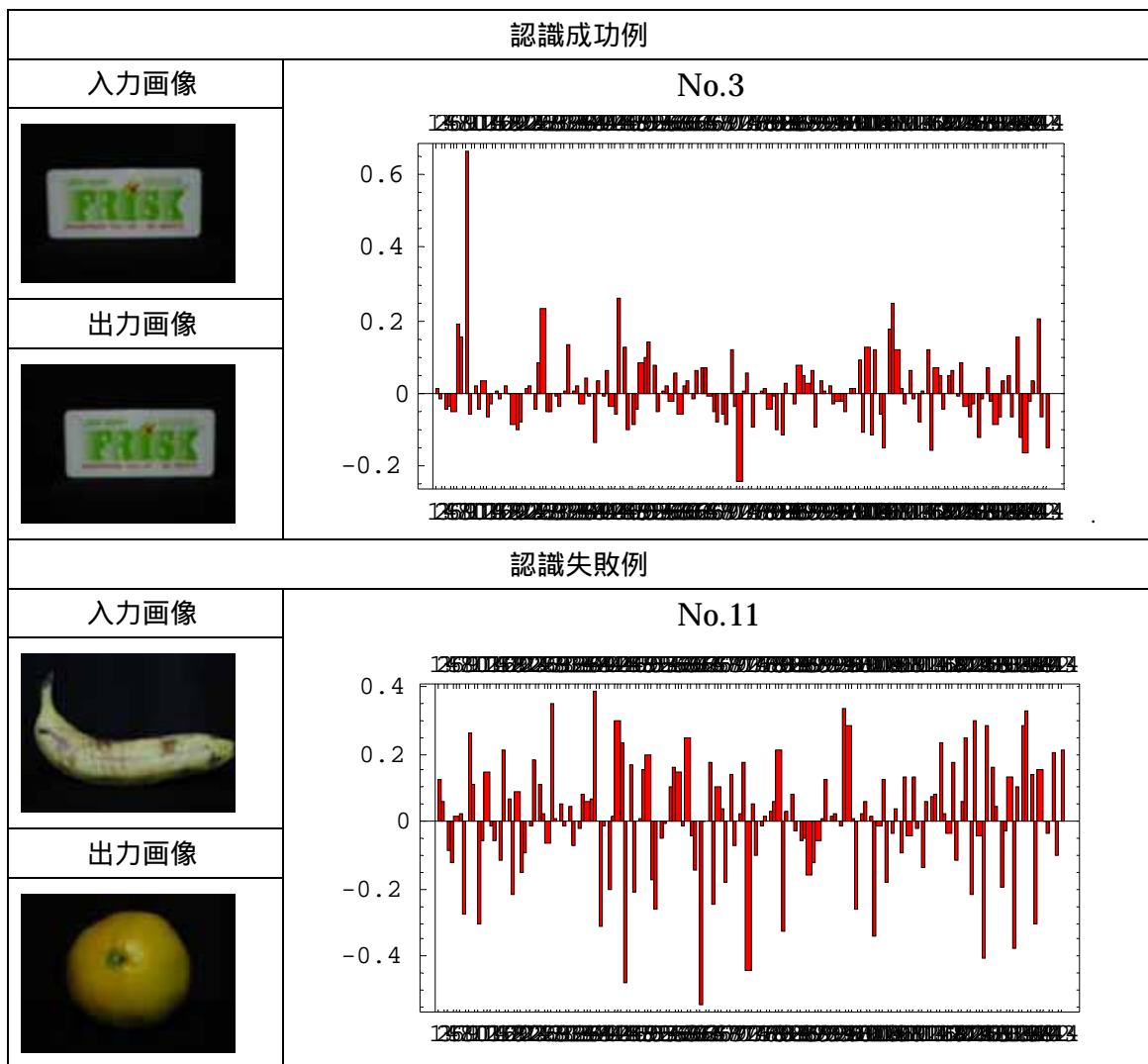


Fig.4 成功例と失敗例

Table.1 の結果から入力画像 50 枚中 39 枚の認識に成功し、認識率は 78.0%になった。認識失敗の考えられる原因としては、撮影時に微妙な光の反射などにより正確な固有パターンを得られなかったこと、画像は対象物に一方向から見た情報をもたないのでそれにより認識できなかった、などがあげられる。

3.3.3 認識成功率向上手法

3.3.2 では理想的な結果が得られなかった。そこで認識率向上手法として最小二乗法を二回適用する。最小二乗法は先に述べたように、式(9)の最大となったデータを出力するが、式(9)の数値が高い n 番目までのデータを取り出しその n 個で再度最小二乗法を試みる。本論文では n を 10 とする。

3.3.4 認識結果

認識結果を Table.2 に示す。出力 No.は Fig.3 のデータベース画像の No.に対応し、判定の ○ は認識成功、× は認識失敗を示す。

Table.2

入力 No.	出力 No.	判定	入力 No.	出力 No.	判定
No.1	No.2		No.26	No.134	×
No.2	No.4		No.27	No.81	
No.3	No.8		No.28	No.83	
No.4	No.12		No.29	No.85	
No.5	No.13		No.30	No.89	
No.6	No.16		No.31	No.91	
No.7	No.19		No.32	No.94	
No.8	No.22		No.33	No.98	
No.9	No.27		No.34	No.101	
No.10	No.28		No.35	No.104	
No.11	No.124	×	No.36	No.106	
No.12	No.34		No.37	No.109	
No.13	No.37		No.38	No.112	
No.14	No.40		No.39	No.117	
No.15	No.43		No.40	No.119	
No.16	No.46		No.41	No.121	
No.17	No.50		No.42	No.124	
No.18	No.52		No.43	No.128	
No.19	No.55		No.44	No.130	
No.20	No.60		No.45	No.134	
No.21	No.62		No.46	No.138	
No.22	No.65		No.47	No.141	
No.23	No.67		No.48	No.144	
No.24	No.70		No.49	No.146	
No.25	No.74		No.50	No.149	

Table.2 の結果より入力画像 50 枚中 48 枚の認識に成功し、認識率は 96%となった。3.3.2 の 78%の認識率と比較してみると、良好な結果といえる。本論文では n を 10 としたが、理想的な n の算出までは至らなく十分検討の余地はある。

4. GVSPM 法による画像認識

4.1 GVSPM 法の導出

3.1 で述べたように不適切な線形システム方程式を算出する場合、GVSPM 法の適用によっても導出することができる。式(14)のような線形システム方程式を与える。

$$Y = CX \quad (14)$$

ここで、 Y は n 次の入力ベクトル、 X は m 次の解ベクトル、 C は n 行 m 列のシステム行列である。そして、式(14)が

$$\begin{aligned} X &= [x_1, x_2, \dots, x_m]^T \\ C &= [C_1, C_2, \dots, C_m] \end{aligned} \quad (15)$$

と以下の式(16)のように変換できる。

$$Y = \sum_{i=1}^m x_i C_i \quad (16)$$

式(16)を入力ベクトル Y のノルムで正規化し、式(17)の関係が得られる。

$$\frac{Y}{|Y|} = \sum_{i=1}^m x_i \frac{|C_i|}{|Y|} \frac{C_i}{|C_i|} \quad (17)$$

ここで、

$$Y' = \frac{Y}{|Y|} \quad X' = \sum_{i=1}^m x_i \frac{|C_i|}{|Y|} \quad C' = \frac{C_i}{|C_i|} \quad (18)$$

とすれば、

$$Y' = C' X' \quad (19)$$

と表すことができる。

式(17)正規化した入力ベクトル Y' が重み付き解要素である $x_i|C_i|/|Y|$ と正規化された列ベクトル $C_i/|C_i|$ の線形結合として得られることを意味する。K 回目の反復解 $X^{(k)}$ で与えられる $CX^{(k)}$ と入力ベクトル Y 間の角度を目標関数 $f(X^{(k)})$ として次の式(20)で定義する。

$$f(X^{(k)}) = \frac{Y}{|Y|} \bullet \frac{CX^{(k)}}{|CX^{(k)}|} = Y' \bullet \frac{C'X^{(k)}}{|C'X^{(k)}|} \quad (20)$$

式(20)で目標関数 $f(X^{(k)})$ が 1 に収束する時、式(14)の解 X が求まる。つまりは、

$$f(X^{(k)}) \rightarrow 1 \quad (21)$$

この式(21)が GVSPM 法の目標関数である。

反復解の初期値を $X^{(0)}$ と設定し、式(22)を与える。

$$X^{(0)} = C'^T Y' \quad (22)$$

式(22)から、最初の偏差ベクトル $\Delta Y'^{(1)}$ が式(23)によって与えられる。

$$\Delta Y'^{(1)} = Y' - \frac{C'X^{(0)}}{|C'X^{(0)}|} \quad (23)$$

式(23)で、偏差 $\Delta Y'$ がゼロベクトルであれば、式(21)の目標関数は自動的に満足される。式(23)を一般化し、第 k 回目の反復解ベクトル $X^{(k)}$ は、第 k-1 回目の偏差ベクトル $\Delta Y'^{(k-1)}$ を用いて式(24)によって与えられる。ここで I_m は m 行 n 列の単位行列を意味する。

$$\begin{aligned} X^{(k)} &= X^{(k-1)} + C'^T \Delta Y'^{(k-1)} \\ &= C'^T Y' + \left(I_m - \frac{C'^T C'}{|C'X^{(k-1)}|} \right) X^{(k-1)} \end{aligned} \quad (24)$$

式(24)で最大値をとる数値より画像を認識する。

4.2 静止画像認識

4.1 で述べた GVSPM 法を適用し静止画像の固有パターンにおける一致性の評価をし精度の報告を行う。反復回数は 5000 回、7000 回の二回行う。

4.2.1 入力画像、データベース

3.3.1 で使用した Fig.2、Fig.3 を引き続き使用する

4.2.2 認識結果

認識結果を反復回数 5000 回を Table.2、7000 回を Table.3 に示す。出力 No.は Fig.3 のデータベース画像の No.に対応し、判定の ○ は認識成功、× は認識失敗を示す。

Table.2

入力 No.	出力 No.	判定	入力 No.	出力 No.	判定
No.1	No.3		No.26	No.76	
No.2	No.4		No.27	No.18	×
No.3	No.8		No.28	No.83	
No.4	No.11		No.29	No.85	
No.5	No.13		No.30	No.89	
No.6	No.16		No.31	No.91	
No.7	No.19		No.32	No.94	
No.8	No.22		No.33	No.97	
No.9	No.27		No.34	No.101	
No.10	No.28		No.35	No.104	
No.11	No.94	×	No.36	No.107	
No.12	No.34		No.37	No.109	
No.13	No.37		No.38	No.112	
No.14	No.40		No.39	No.117	
No.15	No.43		No.40	No.81	×
No.16	No.46		No.41	No.121	
No.17	No.49		No.42	No.124	
No.18	No.53		No.43	No.128	
No.19	No.133	×	No.44	No.130	
No.20	No.60		No.45	No.133	
No.21	No.62		No.46	No.138	
No.22	No.66		No.47	No.141	

No.23	No.67		No.48	No.143	
No.24	No.70		No.49	No.147	
No.25	No.73		No.50	No.150	

Table.3

入力 No.	出力 No.	判定	入力 No.	出力 No.	判定
No.1	No.3		No.26	No.115	×
No.2	No.4		No.27	No.18	×
No.3	No.8		No.28	No.83	
No.4	No.11		No.29	No.85	
No.5	No.13		No.30	No.89	
No.6	No.16		No.31	No.91	
No.7	No.19		No.32	No.94	
No.8	No.22		No.33	No.97	
No.9	No.27		No.34	No.101	
No.10	No.28		No.35	No.104	
No.11	No.94	×	No.36	No.107	
No.12	No.34		No.37	No.109	
No.13	No.37		No.38	No.112	
No.14	No.40		No.39	No.117	
No.15	No.43		No.40	No.81	×
No.16	No.46		No.41	No.121	
No.17	No.49		No.42	No.124	
No.18	No.53		No.43	No.128	
No.19	No.133	×	No.44	No.130	
No.20	No.60		No.45	No.133	
No.21	No.62		No.46	No.138	
No.22	No.66		No.47	No.141	
No.23	No.67		No.48	No.143	
No.24	No.70		No.49	No.147	
No.25	No.73		No.50	No.150	

反復回数 5000 回に設定時の成功例と失敗例を Fig.5 に示す。

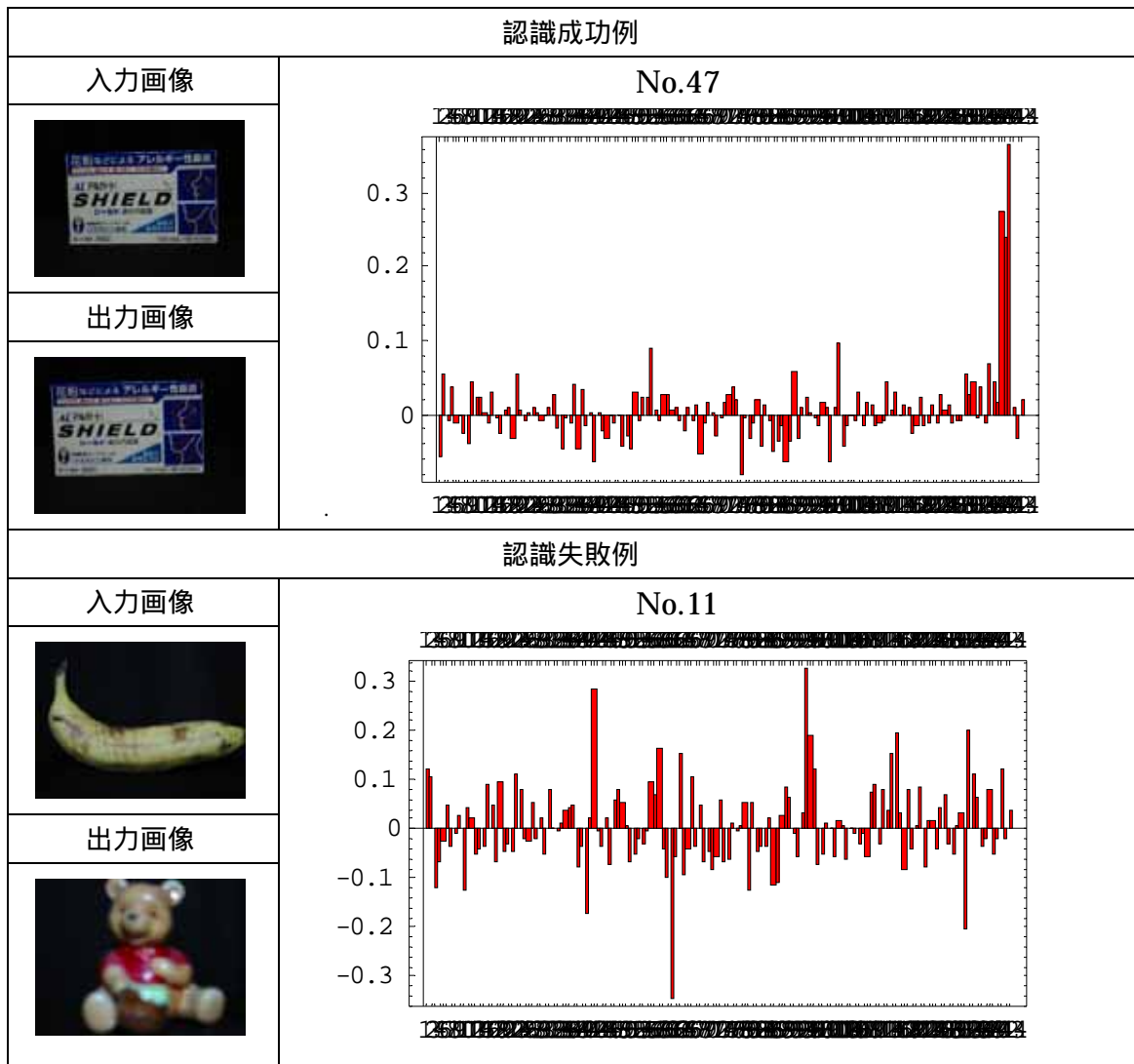


Fig.5

認識結果は反復回数を 5000 回に設定時で静止画像 50 枚中 46 枚成功し、認識率 92%、反復回数を 7000 回に設定時で 50 枚中 45 枚成功し、認識率 90%という結果になった。固有パターン的一致性評価法が最小二乗法より良好な結果を得られた。Table.2、Table.3 をみてわかるように反復回数が 7000 回になり認識率が下がったことから反復回数と認識率が必ずしも比例関係にあるわけではないことがわかる。

5. 相関係数による画像認識

先に述べた最小二乗法や GVSPM 法は画像認識の計算処理速度や認識率の低さなどの問題がある。その問題を解決するために、本章では相関係数の理論を述べ、導入する。

5.1 相関係数の導出

相関係数はある 2 変量 x, y 間の相関関係の強さを知る数値である。次のように定義する。 N 個の 2 変量のデータ配列、つまりは本論文での画像データを次のように与える。

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n) \quad (25)$$

次に、各々の変量の平均値を \bar{x}, \bar{y} とし、 x, y の分散を式(22)とする。

$$\begin{aligned} s_x^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \\ s_y^2 &= \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \end{aligned} \quad (26)$$

さらに、

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (27)$$

とする。式(27)は x, y の共分散で画像の関連性を示すものである。式(26),(27)より x, y の相関係数 r は

$$r = \frac{s_{xy}}{s_x s_y} \quad (28)$$

表現できる。式(28)によって最大値を算出しその要素により静止画像認識を行う。

5.2 静止画像認識

5.1 で示した相関係数を適用し静止画像の固有パターンの一貫性の評価をし精度の報告を行う。

5.2.1 入力画像、データベース

3.3.1 で使用した Fig.2、Fig.3 を引き続き使用する。

5.2.2 認識結果

相関係数を用いた静止画像認識の結果を Table.4 に示す。出力 No.は Fig.3 のデータベース画像の No.に対応し、判定の ○ は認識成功、× は認識失敗を示す。

Table.4

入力 No.	出力 No.	判定	入力 No.	出力 No.	判定
No.1	No.2		No.26	No.76	
No.2	No.4		No.27	No.81	
No.3	No.7		No.28	No.83	
No.4	No.11		No.29	No.85	
No.5	No.13		No.30	No.90	
No.6	No.16		No.31	No.91	
No.7	No.19		No.32	No.94	
No.8	No.22		No.33	No.97	
No.9	No.26		No.34	No.101	
No.10	No.28		No.35	No.104	
No.11	No.31		No.36	No.107	
No.12	No.34		No.37	No.109	
No.13	No.37		No.38	No.112	
No.14	No.40		No.39	No.101	×
No.15	No.43		No.40	No.119	
No.16	No.46		No.41	No.121	
No.17	No.49		No.42	No.126	
No.18	No.54		No.43	No.128	
No.19	No.46	×	No.44	No.130	
No.20	No.60		No.45	No.134	
No.21	No.62		No.46	No.138	
No.22	No.125	×	No.47	No.139	
No.23	No.67		No.48	No.143	
No.24	No.70		No.49	No.147	
No.25	No.73		No.50	No.148	

認識成功例と失敗例を Fig.6 に示す。

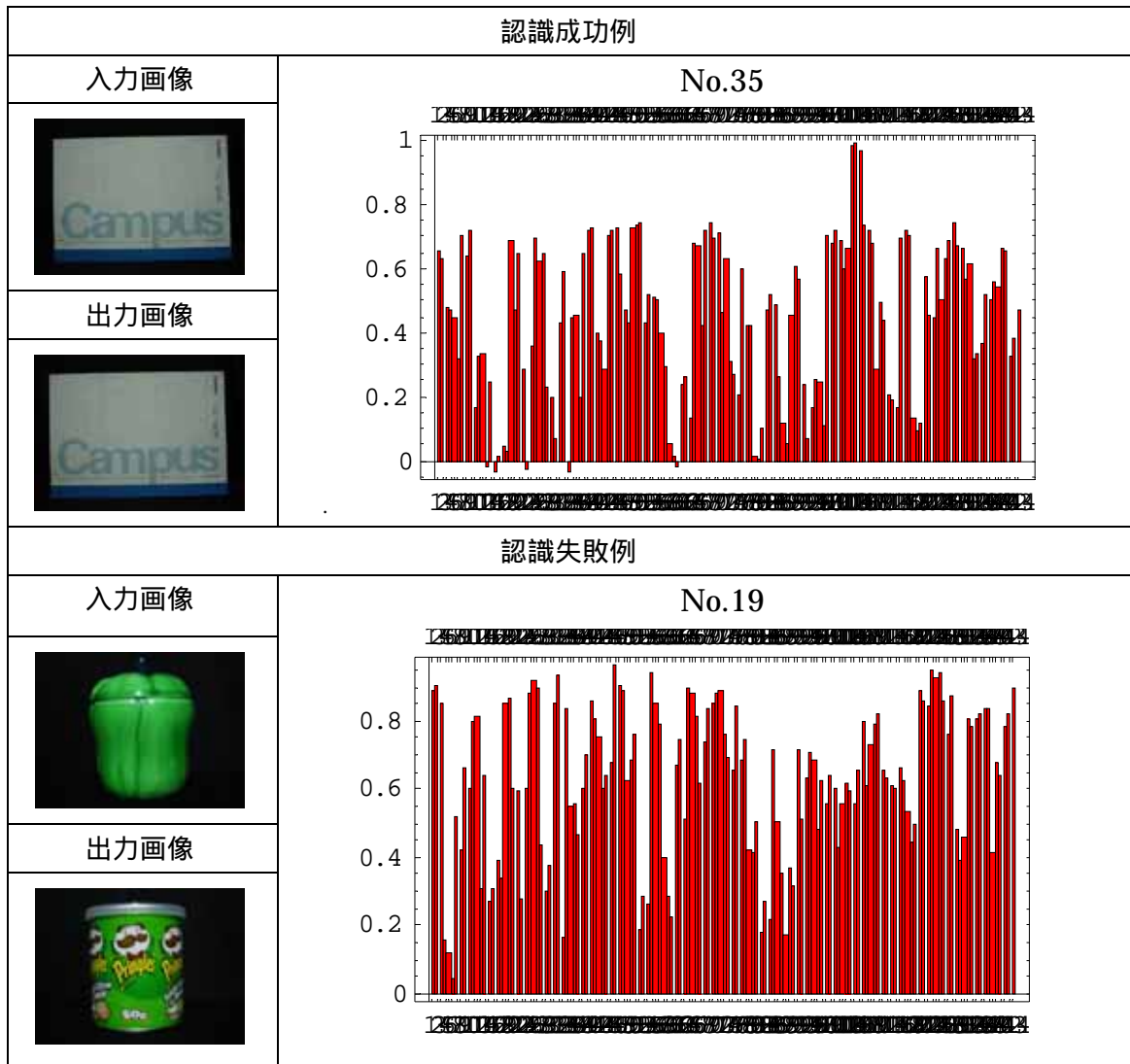


Fig.6

Table.3 の結果から入力画像 50 枚中 47 枚の認識に成功し、認識率は 96.0% になった。先に述べた最小二乗法、GVSPM 法と比べると（3.3.3 の手法を除く）良好な結果が得られ、2 手法の問題点の一つだった処理速度も速くなったことを確認した。

6. RGB成分の独立認識による画像合成認識

本論文では、これまで画像認識を行う際、一つの画像における RGB 成分の固有パターンを一つのデータ、つまりは一つの行列として計算機で処理・計算を行っていた。RGB 成分はそれぞれ 0 から 255 の要素がありそれらをひとつの要素としていた。それら RGB 成分をそれぞれ一つずつ認識して元の画像を認識し、精度を報告する。認識方法としてまず、それぞれの固有パターンを算出し、基準をもたすため 0 から 1 に正規化を行い R 成分固有パターン、G 成分固有パターン、B 成分固有パターンを掛け合わせ、最大をとる画像を出力する。仮に R 成分固有パターン、G 成分固有パターン、B 成分固有パターンがすべて最大とすれば出力数値は 1 となる。本章の目的は画像を R,G,B 成分の 3 種に分け一つずつ認識させ、1 つの認識結果にすることで、従来の認識方法より認識率が下がる可能性はあるが、精密に判定できるかにある。そして認識手法は第 4 章で述べた GVSPM 法を採用する。

6.1 静止画像認識

上記で述べた手法を用いて静止画像の固有パターンの一致性の評価を行い精度の報告を行う。反復回数は 5000 回、7000 回の二回行う。

6.1.1 入力画像、データベース

3.3.1 で使用した Fig.2、Fig.3 を引き続き使用する。

6.1.2 認識結果

認識結果を反復回数 5000 回及び、7000 回を Table.5 に示す。出力 No.は Fig.3 のデータベース画像の No.に対応し、判定の ○ は認識成功、× は出力数値が 1、× は認識失敗、総合判定は ○ が 5000 回及び、7000 回の出力数値が 1、○ は両データとも認識成功、× は両データとも認識失敗、○ はどちらか一方が認識失敗による認識不能を示す。

Table.5

入力 No.	反復回数	5000 回		反復回数	7000 回		
	出力 No.	出力数値	判定	出力 No.	出力数値	判定	総合判定
No.1	No.3	0.34		No.106	0.23	×	
No.2	No.4	0.68		No.4	0.68		
No.3	No.108	0.51	×	No.8	0.5		
No.4	No.11	0.79		No.11	0.75		
No.5	No.13	0.83		No.13	0.83		
No.6	No.16	1		No.16	1		
No.7	No.19	0.76		No.19	0.72		
No.8	No.22	0.41		No.22	0.39		

No.9	No.109	0.58	×	No.109	0.52	×	×
No.10	No.30	0.71		No.30	0.73		
No.11	No.31	0.38		No.31	0.42		
No.12	No.34	1		No.34	1		
No.13	No.37	0.92		No.37	0.94		
No.14	No.40	0.64		No.40	0.62		
No.15	No.43	0.89		No.43	0.89		
No.16	No.46	0.93		No.46	0.95		
No.17	No.49	0.96		No.49	0.94		
No.18	No.53	0.95		No.53	0.95		
No.19	No.57	0.43		No.57	0.41		
No.20	No.60	0.92		No.60	0.82		
No.21	No.62	0.6		No.62	0.57		
No.22	No.66	0.99		No.66	0.87		
No.23	No.39	0.92	×	No.39	0.94	×	×
No.24	No.70	0.58		No.70	0.61		
No.25	No.74	0.92		No.74	0.94		
No.26	No.115	0.25	×	No.115	0.25	×	×
No.27	No.18	0.44	×	No.81	0.44		
No.28	No.108	0.97	×	No.108	0.94	×	×
No.29	No.85	0.93		No.85	0.93		
No.30	No.89	0.95		No.89	0.95		
No.31	No.91	0.95		No.91	0.94		
No.32	No.95	0.74		No.95	0.74		
No.33	No.97	0.81		No.97	0.85		
No.34	No.101	1		No.101	1		
No.35	No.104	0.4		No.104	0.35		
No.36	No.107	1		No.107	1		
No.37	No.109	1		No.109	1		
No.38	No.112	0.55		No.112	0.51		
No.39	No.140	0.23	×	No.116	0.25		
No.40	No.120	0.54		No.120	0.56		
No.41	No.121	1		No.121	1		
No.42	No.32	0.66	×	No.32	0.69	×	×
No.43	No.128	0.99		No.128	1		

No.44	No.130	0.99		No.130	1		
No.45	No.133	0.41		No.133	0.39		
No.46	No.138	0.94		No.138	0.92		
No.47	No.141	1		No.141	1		
No.48	No.143	0.74		No.143	0.69		
No.49	No.147	0.78		No.147	0.76		
No.50	No.149	0.82		No.149	0.83		

反復回数 7000 回時の認識成功例と失敗例を Fig.6 に示す。

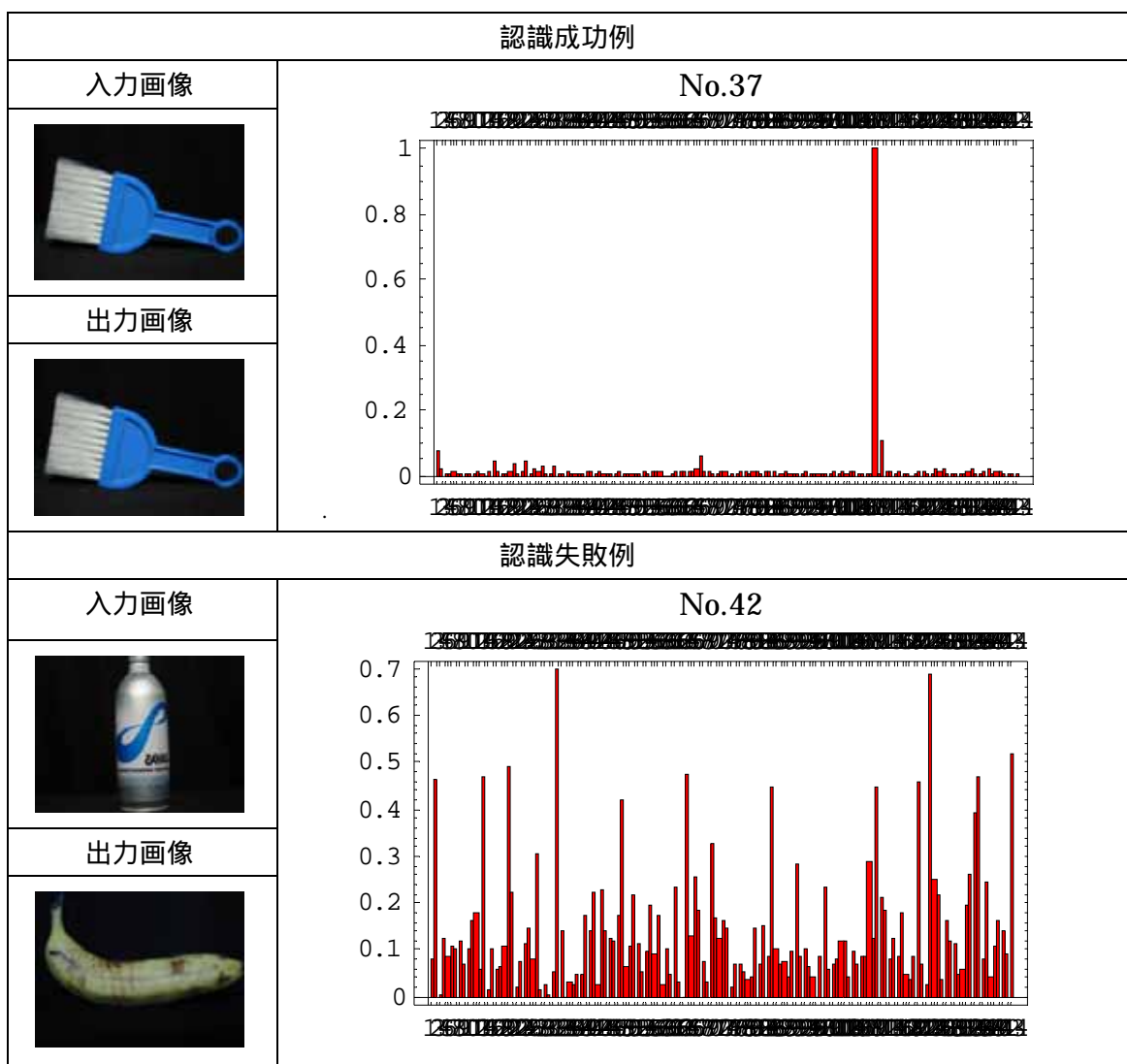


Fig.6

Table.5 の結果から反復回数 5000 回時、入力画像 50 枚中 42 枚の認識に成功し、認識率は

84.0%,7000 回時、入力画像 50 枚中 44 枚の認識に成功し、認識率は 88%、総合判定における認識率は 82%になった。この 82%という認識率は認識不能も認識失敗として算出されており、本論文での他の静止画像認識における認識率と数値で比べると良好な結果とはいえない。しかし、他の静止画像認識の認識成功における過程より条件が厳しく設定されており、信頼性が本章における静止画像認識は高いといえる。仮に、他の静止画像認識における認識率と同一の数値が算出された場合、本章における画像合成認識のほうが信頼性が高いと思われる。また認識不能と算出された画像も GVSPM 法の反復回数に依存するので予備出力として認識成功、認識失敗の両方でないカテゴリとして扱いたい。さらに出力数値が高く算出され認識成功に至るケースはかなり信頼性が高いといえ他の静止画像認識手法より細かく分析できる利点があると思われる。

最後に、本論文でいくつかの手法により静止画像認識を行ったが、手法問わず同じ入力画像を同じデータベースで認識失敗するケースがみられた。これは人間の視覚情報処理能力も個人により認識しづらいケースがあるのと同様に、計算機に認識しづらい固有パターンが存在すると思われる。

7. まとめ

本論文では、まず画像の普遍量とは RGB 成分で構成されていることを示しその固有パターンの抽出方法を述べた。そうすることにより、画像を構成する画素数、画像の x - y 直交座標上の位置・角度といったハードウェアに依存する性質を削減することができ、画像の情報を抽出し静止画像認識を行う手法を述べた。

画像の固有パターンからなる線形システム方程式は不適切であるため、最小二乗法及び、GVSPM 法により解を導出する手法を提案した。また認識率向上及び計算機の処理速度の問題を解決するため画像間の相関関係を求めることにより一致性を評価した相関係数法を提案した。さらに RGB 成分それぞれを固有パターンにし、正規化した数値データを掛け合わせ一致性を評価する合成認識法により、認識率の信頼性を向上させ認識結果からより多くの情報を読み取ることに成功した。またいくつかの静止画像認識手法を行うことにより計算機が手法問わず認識失敗するケースが存在することを報告した。

よって、本論文で提案した手法は、高い認識率が得られたことから、計算機の識別能力が人間の視覚情報処理に近づくという本論文の目的に達した。

8. 参考文献

- [1] 丸山和夫、早野誠治、斎藤兆古「色情報を利用した知的動画像認識」
可視化情報学会誌、Vol.23, No.1 (2003) pp.95-98.
- [2] 穴吹幸彦、早野誠治、斎藤兆古「GVSPM 法による 3 次元電流分布の可視化」
第 33 回可視化情報シンポジウム (2005)

2005年度

卒業論文

ウェーブレット変換を用いた色情報の圧縮による
静止画像認識

法政大学

工学部 情報電気電子工学科

学籍番号 02d3042

氏名 黒田洪平

指導教員 齋藤兆古

目次

1. 序論

2. 静止画像の固有パターン
 - 2.1 画像の色成分
 - 2.2 カラー画像の固有ベクトル
 - 2.3 固有パターン

3. 固有パターンの一貫性評価法
 - 3.1 線形システム方程式
 - 3.2 解ベクトル
 - 3.3 静止画像認識

4. ウェーブレット変換
 - 4.1 概要
 - 4.2 離散値系ウェーブレット変換
 - 4.3 ウェーブレット変換による圧縮
 - 4.4 静止画像認識

5. まとめ

6. 参考文献

1.序論

現在、画像を計算機によって処理・加工するには人間が見て評価することを前提としている。この意味で、人間の視覚評価能力を前提として画像をハンドリングする画像処理は比較的多く実用化されている。しかし、映像を認識する技術は完成度が低く、最終的には人間の視覚情報処理能力に依存せざるを得ない。例を挙げると、同一被写体をデジタルカメラで撮影した場合、撮影条件が完全に同一でない限り同一の数値データをとらない、つまり、機械は同一の被写体と認識できないのである。逆を言えば、まったく同一の画像と認識できる。しかしそのことを加味し、最新の計算機をもってしても、画像認識の分野においては人間の視覚情報処理能力には到底及ばない。映像を計算機によって認識させることは、計算機に人間と同等な視覚情報処理能力を与えることになり、現状の様々な手法では人間の視覚情報処理能力には匹敵し難い。このような現状をふまえて、本論文では、電子的に計算機スクリーン上に可視化された映像から、画素数、配置、変形等に無関係となる映像の普遍量を抽出する概念を導入し、計算機に計算機の正確さに加え人間の視覚情報処理能力を与え、計算機自身によって映像を認識させることを試みる。

ここで我々が認識の対象とする画像は人間の視覚によって認識しえる形態で表現された情報をいい、それらを総称して本論文では可視化情報と称する。その最たる例として、人間の視覚によって見ることができるカメラやビデオで撮影された画像などである。また、人間の視覚情報処理機能で捉えることができない情報を映像化した赤外線映像や電子顕微鏡による映像等も可視化情報の一種である。さらに人間の視覚で認識可能な形態へ変形された信号情報、例を挙げるなら音声等を電子的にオシロスコープ等の機器で映像化した情報もその一種である。

人間の脳の視聴覚情報機能は、眼球を通して得られた映像を処理することによって情報を認識するが、その対象となる視覚情報には、一定の構成ルールに従ってコード化された文字や言語等のコード化視聴覚情報と、コード化されてない可視化情報、すなわち非コード化視聴覚情報がある。先に述べた可視化情報は、非コード化視覚情報を処理する人間の視覚情報処理能力を前提としたデータであり、人間は人間の持つ視覚情報処理能力によって人間の視覚情報処理能力を実現し認識するためには、可視化情報から計算機で処理可能な何らかの普遍量を抽出する必要がある。

計算機のスクリーン上の可視化情報は2次元平面上の画素(Pixel)で構成されている。各画素は可視光の波長によってそれぞれ赤、緑、及び青(以下それぞれR,G,Bと略記)成分の情報を持つ。また、可視化情報は反射光の強さ(Intensity)、色相(Tone)、及び色成分(Color component)の組み合わせからなり、それらの幾何学的配置で表現される。そのため、計算機のスクリーン上の可視化情報は、スクリーンの解像度や基準座標(視点)に依存する性質を持つ。このようなハードウェアに依存する性質を削

減した可視化情報の普遍量を本論文では画像の固有パターン(Eigen pattern)と呼び、前述の計算機の普遍量はこれに相当し、この固有パターンを駆使し画像認識を試みる。

本論文で採用する画像認識手法は線形システム方程式の解を導くことに基づく方法である。しかしながら、得られた線形システム方程式は不適切なシステム方程式であるため、最小二乗法を用いて導く。また、固有パターンの情報量が膨大なため固有パターンを圧縮しそれを新たな固有パターンとして扱い画像認識を試みる。

本論文の構成は、最初に静止画像から固有パターンを抽出する方法について述べる。次にウェーブレット変換について述べ、固有パターンを圧縮した画像認識を行い、結果を報告する。最後に本文をまとめる。

2.静止画像の固有パターン

2.1 画像の色成分

計算機に取り込まれた画像は $x - y$ 直交座標上の画素で構成されている。その各画素は R, G, B 成分で構成されている。よって、 m, n をそれぞれ x, y 直行座標系の x, y 方向の画素数とすれば、画像データは $m \times n$ 次のベクトルとして扱うことができる。ベクトルとして表現されたカラー画像データにおける第 k 番目の各画素の R, G, B 成分を R, G, B とすると、画素は式(1)と表すことができる。

$$\begin{aligned} image &\in pixel_{i,j} \\ pixel_{i,j} &\in R_k, G_k, B_k \\ i &= 1, 2, 3, \dots, m \\ j &= 1, 2, 3, \dots, n \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \tag{1}$$

2.2 カラー画像の固有ベクトル

カラー画像の固有パターンの算出はまず、画像中の R 成分の強度値、 G 成分の強度値、 B 成分の強度値で与えられ、1 画像に対して 3 組得られる。 R 成分分布を I_R とすれば、

$$\begin{aligned} I_R &= R_k \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \tag{2}$$

と与えられる。次に R 成分分布 I_R をダイナミックレンジ D の分布に正規化した正規化 R 成分分布 I_R^D を式(3)で定義する。下式の $Round[*]$ 、 $Max[*]$ はそれぞれ括弧内の値を整数化する演算、括弧内の集合における最大値の要素を求める演算である。

$$\begin{aligned} I_R^D &\in Round \left[D \times \frac{R_k}{Max[I_R]} \right] \\ k &= 1, 2, 3, \dots, m \times n \end{aligned} \tag{3}$$

さらに式(3)で求めたダイナミックレンジ D の正規化 R 成分分布 I_R^D において 1 から D までの数値の出現頻度を算出しヒストグラムにする。 $Count[*], p]$ は括弧内の集合における数値 p の個数を考える演算である。

$$E_R \in \text{Count}[I_R^D, p] \quad (4)$$

$$p = 1, 2, 3, \dots, D$$

式(4)で求めた E_R を R 成分固有ベクトルといい、同様に G 成分、B 成分も固有ベクトルを算出する。

2.3 固有パターン

固有パターンを E とすると固有パターン E は先に述べたように R 成分固有パターン、G 成分固有パターン、B 成分固有パターンの 3 種の固有パターンの集合からできている。式(5)で与えられる。

$$E \in E_R, E_G, E_B \quad (5)$$

図 1 は左に静止画像とその右 R、G、B 各成分のヒストグラムを示す。図 1 から、画像特有のパターンが存在することがわかり、特有の普遍量が抽出されていることがわかる。

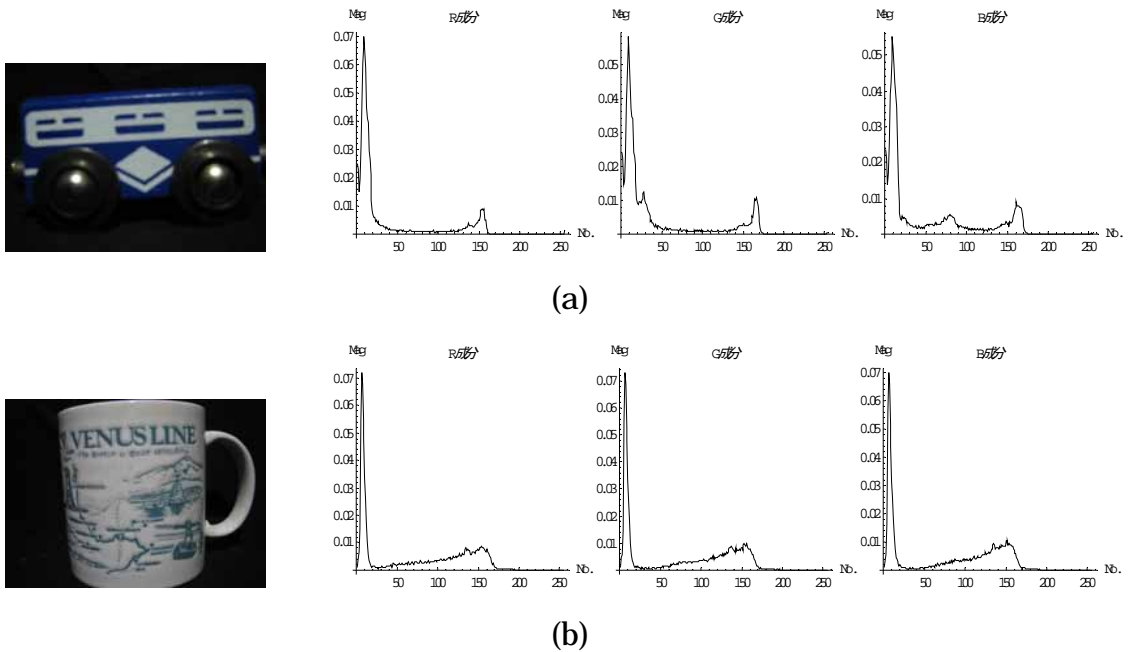


Fig.1 画像の固有パターン

3.固有パターンの一貫性評価法

3.1 線形システム方程式

例として n 個の画像データから固有パターンを抽出しデータベースを構築する。それぞれの画像の固有パターンである R 成分固有パターン、G 成分固有パターン、B 成分固有パターンで式(6)のようなシステム行列を構築する。

$$\begin{aligned} C_R &= [E_{R1}, E_{R2}, E_{R3}, \dots, E_{Rn}] \\ C_G &= [E_{G1}, E_{G2}, E_{G3}, \dots, E_{Gn}] \\ C_B &= [E_{B1}, E_{B2}, E_{B3}, \dots, E_{Bn}] \end{aligned} \quad (6)$$

そして入力画像データ R 成分固有パターン、G 成分固有パターン、B 成分固有パターンをそれぞれ E_{YR}, E_{YG}, E_{YB} 、解ベクトルを X_R, X_G, X_B とすると R,G,B 成分の線形システム方程式は式(7)で表すことができる。ここで解ベクトル X_R, X_G, X_B の要素で最大値をもつ要素が残りの要素に対して相対的に値が大きいほど、入力画像データと式(6)を構成するデータベースの特定の固有パターンが等しい、つまりは同じ画像の可能性が高いということの意味する。

$$\begin{aligned} E_{YR} &= C_R \cdot X_R \\ E_{YG} &= C_G \cdot X_G \\ E_{YB} &= C_B \cdot X_B \end{aligned} \quad (7)$$

3.2 解ベクトル

固有パターンの要素数はダイナミックレンジ D によって決まり、本論文では 0 から 255 の 256 段階の範囲で構成している。データベースの数がこの固有パターンの要素数より少ない場合、式(7)のシステム方程式から解ベクトル X_R, X_G, X_B を算出することは不適切問題を解くことに帰する。この不適切問題を本章では最小二乗法によって解く。この最小二乗法はこのような不適切なシステム方程式を解くことにあたっては最良な手法の一つである。解ベクトル X_R, X_G, X_B は以下のように導出する。

$$\begin{aligned} X_R &= [C_R^T C_R]^{-1} C_R^T E_{YR} \\ X_G &= [C_G^T C_G]^{-1} C_G^T E_{YG} \\ X_B &= [C_B^T C_B]^{-1} C_B^T E_{YB} \end{aligned} \quad (8)$$

式(8)より得られた R,G,B の解ベクトル X_R, X_G, X_B を畳み込み演算し、解ベクトル X を得

る。*は畳み込み演算を示す。式(9)で最大値を取る要素から静止画像を認識する。

$$X = X_R * X_G * X_B \quad (9)$$

3.3 静止画像認識

3.1,3.2 で述べた最小二乗法による線形システム方程式の静止画像認識手法を用いてデータベースと入力画像それぞれの固有パターン間の一致性を評価し認識し、精度を報告する。

3.3.1 入力画像

背景の情報が認識に作用しないように、白熱電球の下で暗室にて撮影をした画像を扱う。図2に入力画像を50枚を示す。





No.31



No.32

Fig.2 入力画像

3.3.2 データベース画像

入力画像と同じ環境で撮影したものを扱い、1つの入力画像に対して同じ対象物の画像を2つ用意した。図3にデータベース画像を示す。



No.1



No.2



No.3



No.4



No.5



No.6



No.7



No.8



No.9



No.10



No.11



No.12



No.13



No.14



No.15



No.16



No.17



No.18



No.19



No.20



No.21



No.22



No.23



No.24



No.25



No.26



No.27



No.28



No.29



No.30

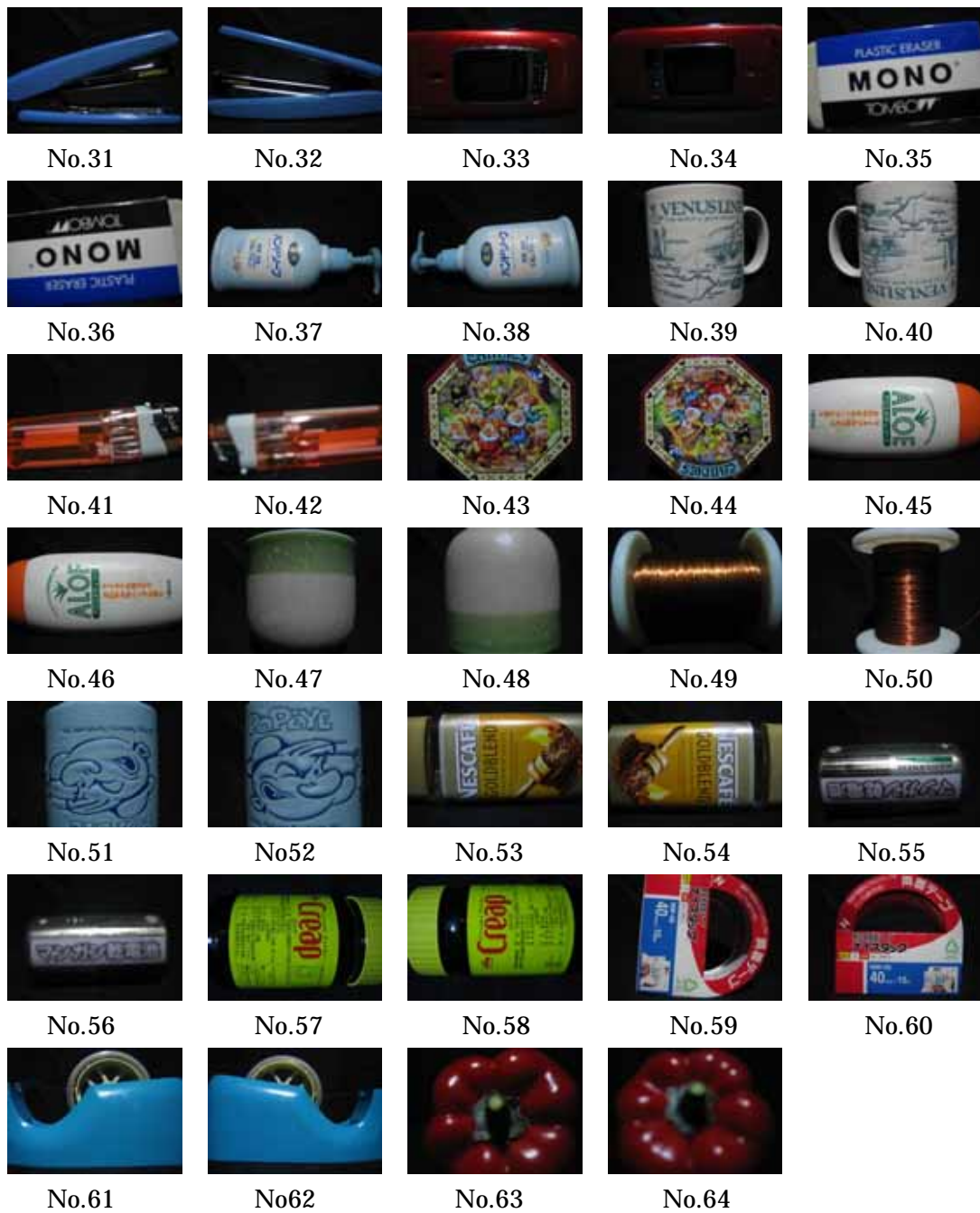


Fig.3 データベース画像

3.3.3 認識結果

認識結果を表.1 に示す。入力、出力 No.は図 2、図 3 の画像の No.に対応し、判定のは認識成功、×は認識失敗を示す。

Table.1 認識結果 1

入力 No.	出力 No.	判定	入力 No.	出力 No.	判定
No.1	No.1		No.17	No.33	
No.2	No.3		No.18	No.35	
No.3	No.6		No.19	No.38	
No.4	No.7		No.20	No.40	
No.5	No.9		No.21	No.41	
No.6	No.11		No.22	No.43	
No.7	No.11	×	No.23	No.45	
No.8	No.15		No.24	No.48	
No.9	No.44	×	No.25	No.49	
No.10	No.20		No.26	No.52	
No.11	No.22		No.27	No.54	
No.12	No.22	×	No.28	No.55	
No.13	No.26		No.29	No.58	
No.14	No.27		No.30	No.60	
No.15	No.27	×	No.31	No.61	
No.16	No.31		No.32	No.63	

認識成功例を図.4、認識失敗例を図5に示す。

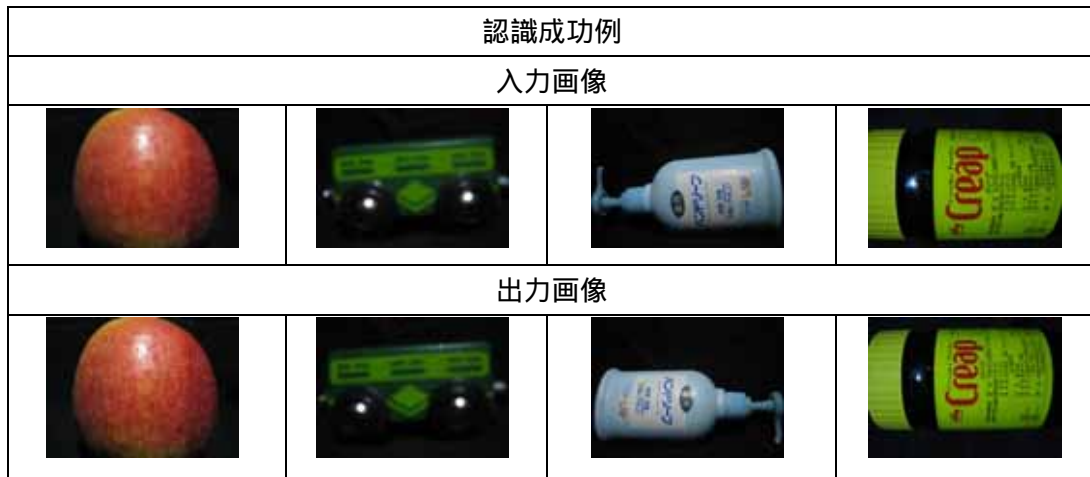


Fig.4 成功例

認識失敗例
入力画像 / 固有パターン

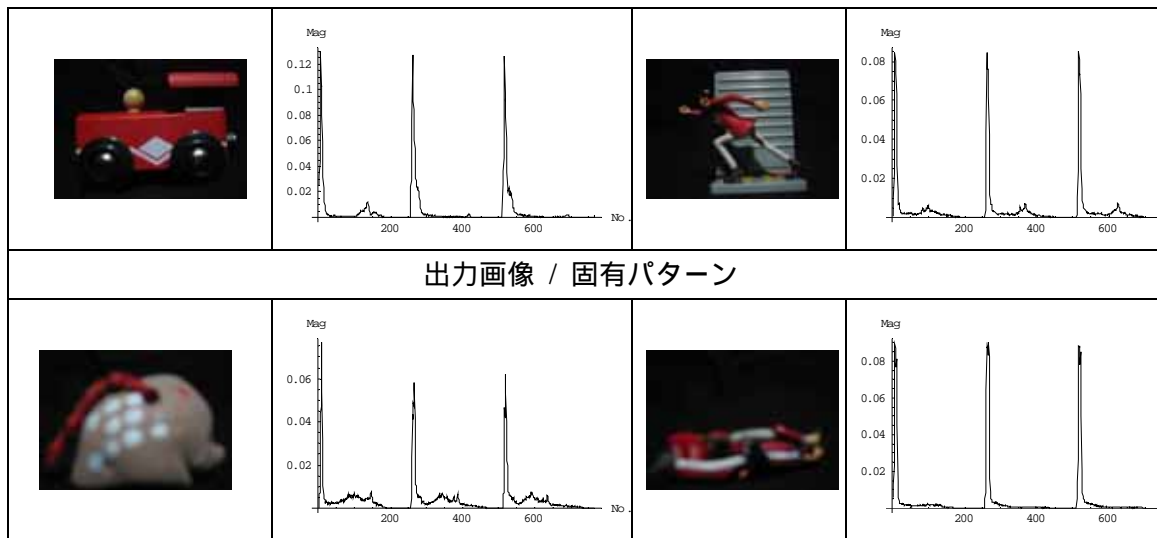


Fig.5.失敗例

表.1の結果から入力画像 32 枚中 28 枚の認識に成功し、認識率は 87.5%になった。認識失敗した原因としては、図 5 から入力画像と出力画像のそれぞれの固有パターンを比べると似ていることから失敗したと考えられる。人間の目では明かに違う画像でも、光源による反射光

4. ウェーブレット変換

4.1 概要

4.1.1 関数系の変換

古典的な関数変換の目的は、解析的に扱いにくい関数系を解析的に扱いやすい関数系へ変換することである。例えば、ラプラス(Laplace)変換は微積分演算を単純な掛け算や割り算へ化す変換である。また、関数系の変換とは、ある関数、例えば時間変化する関数を解析が容易な周期関数の和で表現することにも使われる。具体的な例としてフーリエ(Fourier)変換を考えよう。フーリエ変換は解析的に扱えない関数を解析的に扱える角周波数の異なる正弦波と余弦波の和で表現する変換である。換言すれば、フーリエ変換は解析的に扱いにくい関数系を解析的に扱いやすい関数系へ分解する変換と考えてもよい。

ウェーブレット変換は、ラプラス変換のように解析的な計算を容易な形へ変換する方法でなく、フーリエ変換のように与えられた関数を他の関数系の和で表現するために用いられる。いま、ある任意の時間 t をパラメータとする関数 $f(t)$ を一定値 a_0 、正弦波および余弦波の和で表現できるとしよう。すなわち、

を角周波数として、

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin(i\omega t) + \sum_{i=1}^{\infty} b_i \cos(i\omega t) \quad (10)$$

と仮定する。

問題は(10)式の係数、 a_0 、 a_i 、 b_i を決める方法である。いま、関数 $f(t)$ が時間 $t=0$ から $t=T$ の区間で(1)式の係数を決めることを考えれば、(10)式は、 $\omega = 2\pi/T$ であるから、

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i \frac{2\pi}{T} t\right) \quad (11)$$

とも書ける。

(11)式の両辺を時間 $t=0$ から $t=T$ の区間について積分すると、

$$\begin{aligned} & \int_0^T f(t) dt \\ &= \int_0^T \left\{ a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i \frac{2\pi}{T} t\right) \right\} dt \quad (12) \\ &= a_0 T \end{aligned}$$

が成り立つ。

したがって、定数項 a_0 は

$$a_0 = \frac{1}{T} \int_0^T f(t) dt \quad (13)$$

となる。

次に、(11)式の両辺に正弦波 $\sin[j(2\pi/T)t]$ 、 $j=1,2,3,\dots$ を掛け算し、時間 $t=0$ から $t=T$ の区間について積分する。

$$\begin{aligned}
& \int_0^T f(t) \sin\left(j \frac{2\pi}{T} t\right) dt \\
&= \int_0^T \left\{ a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i \frac{2\pi}{T} t\right) \right\} \times \sin\left(j \frac{2\pi}{T} t\right) dt \quad (14) \\
&= a_i \frac{2}{T}, i = j \text{ or } 0, i \neq j
\end{aligned}$$

(14)式から係数 a_i は

$$a_i = \frac{2}{T} \int_0^T f(t) \sin\left(i \frac{2\pi}{T} t\right) dt \quad (15)$$

として得られる。

同様に、(11)式の両辺に余弦波 $\cos\left[j \frac{2\pi}{T} t\right], j=1, 2, 3, \dots$ を掛け算し、時間 $t=0$ から $t=T$ の区間について積分することで

$$b_i = \frac{2}{T} \int_0^T f(t) \cos\left(i \frac{2\pi}{T} t\right) dt \quad (16)$$

として係数 b_i が得られる。

したがって、関数 $f(t)$ は、区間 $t=0$ から $t=T$ で、

$$\begin{aligned}
f(t) &= \frac{1}{T} \int_0^T f(t) dt + \\
& \sum_{i=1}^{\infty} \left\{ \frac{2}{T} \int_0^T f(t) \sin\left(i \frac{2\pi}{T} t\right) dt \right\} \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} \left\{ \frac{2}{T} \int_0^T f(t) \cos\left(i \frac{2\pi}{T} t\right) dt \right\} \cos\left(i \frac{2\pi}{T} t\right) \quad (17)
\end{aligned}$$

と書ける。これがいわゆるフーリエ変換の原型となるフーリエ級数であり、左辺の関数 $f(t)$ を右辺の計算が簡単な定数項と三角関数の和に変換している。

4.1.2 関数の直交性と線形性

フーリエ級数の考え方の中に重要な関数間で成り立つ性質、すなわち、関数の直交性(orthogonality)が使われている。

まず(4)式の係数 a が計算される過程を考えよう。(12)式は(11)式の両辺に定

数値1を掛け算し積分する演算である。このとき、

$$\int_0^T 1 \cdot \sin\left(i \frac{2\pi}{T} t\right) dt = 0, \int_0^T 1 \cdot \cos\left(i \frac{2\pi}{T} t\right) dt = 0, \quad (18)$$

$i = 1, 2, \dots$

の関係が成り立つために、(13)式の係数 a_0 がうまく計算できた。この関係を、定数値1と正弦波 $\sin[j(2\pi/T)t]$ 、および余弦波 $\cos[j(2\pi/T)t]$ 間の直交性と呼ぶ。同様に、(17),(18)式で計算される a_i 、 b_i は

$$\int_0^T \sin\left(i \frac{2\pi}{T} t\right) \sin\left(j \frac{2\pi}{T} t\right) dt = 0,$$
$$\int_0^T \cos\left(i \frac{2\pi}{T} t\right) \cos\left(j \frac{2\pi}{T} t\right) dt = 0, \quad (19)$$

$i \neq j$

なる直交性が成り立つことに基づいている。

(18),(19)式から、直交性とは与えられた関数を他の関数の和で表現しようとする場合、和となる関数の大きさ(係数)を一意的に決める条件であることがわかる。言い換えれば、与えられた関数を他の関数の和で表現できる条件である。

さて、ある任意の時間 t をパラメータとする関数 $f(t)$ は、区間 $t=0$ から $t=T$ で、一定値、正弦波および余弦波の和で表現できることがわかった。この変換は、一定値、正弦波および余弦波間で直交性が成り立つことが条件であった。この結果に至る過程を考えてみると、まず、展開される関数の和でもとの関数が表現されるとする大前提があることに気づくであろう。ある関数が他の関数の和で表現できる性質を線形性と呼ぶ。では、(15)式で計算される正弦波の係数 a_i が

定数 c_i の n 個の和で表現されるとしよう。すなわち、

$$a_i = c_i + c_i + \dots = nc_i \quad (20)$$

である。(20)式から、和を前提とする系では比例関係が成り立つことを意味することがわかる。すなわち、線形系とは比例関係が成立する系である。

4.1.3 連続系ウェーブレット変換

フーリエ級数で、和となる正弦波や余弦波を基底関数と呼ぶ。この基底関数に正弦波や余弦波のように直交性が成り立たない関数を採用することを考える。例えば、図1に示す、

$$\varphi(t) = (1 - t^2) \exp\left(-\frac{t^2}{2}\right) \quad (21)$$

で与えられる関数である。

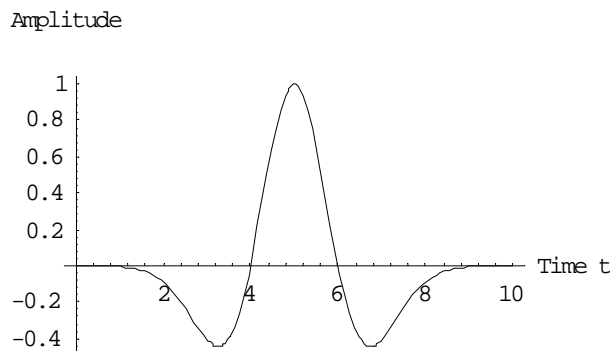


Fig.6. 非直交基底関数の例(メキシカンハット型)

この基底関数を用いて、

$$\Psi(t) = \frac{1}{\sqrt{|a|}} \varphi\left(\frac{t-b}{a}\right) \quad (22)$$

なる関数を考えよう。

(22)式で、パラメータ a は基底関数の広がり、すなわち、図6の時間 $t=1$ から $t=9$ までの周期 T を表し、 b は基底関数のピーク $t=5$ の時間位置を示す。フーリエ級数では、周期 T の異なる正弦波と余弦波を基底関数として考えた。(22)式では、周期 T に対応する a と時間位置 b をパラメータとする変換を行う。

(15), (16) 式と同様な形式の内積演算

$$T(a,b) = \frac{1}{\sqrt{c_\phi}} \int_{-\infty}^{\infty} \Psi^{(a,b)}(t) f(t) dt \quad (23)$$

を行い、 a 、 b をパラメータとする係数 $T(a,b)$ を計算する。

さて、この係数 $T(a,b)$ はどのような物理的な意味を持つであろうか。まず、パラメータ b で図1に示す波形のピーク時間位置が決まる。つぎに、パラメータ a で波形の周期 T が決まる。したがって、係数 $T(a,b)$ は、関数 $f(t)$ が周期 $T=a$ 、時間 $t=b$ でピークを持つ図1の波形と共通部分を取り出した値に相当することを意味する。言い換えれば、仮に関数 $f(t)$ が図1の波形と全く同じ波形であれば、係数 $T(a,b)$ は最大値をとる。これは、ある任意波形から、図1の波形と同じ波形成分を取り出す演算として(23)式がきわめて有効であることを意味する。この変換を連続ウェーブレット変換と呼ぶ。残念ながら、(21)式の基底関数はパラメータ a 、 b が異なっても直交性を持たない。このため、線形性も成り立たない。

4.1.4 離散値系ウェーブレット変換

離散値系(discrete)とはどのような系であろうか。離散値なる用語が生まれたのは計数型計算機を用いて数値計算を行う場合、連続関数をサンプリングして得られる数値で代表したことに起因する。例えば、関数 $f(t)$ を時間 $t=0$ から $t=T$ の区間で、 t ごとに n 個サンプリングして離散値系で表すと、

$$F = [f(\Delta t), f(2\Delta t), f(3\Delta t), \dots, f(n\Delta t)]^T \quad (24)$$

なるベクトルとなる。

同様に、(10)式を離散値系で書けば、

$$\begin{pmatrix} f(\Delta t) \\ f(2\Delta t) \\ f(3\Delta t) \\ \vdots \\ f(n\Delta t) \end{pmatrix} = a_0 \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \begin{pmatrix} \sin\left(\frac{2n}{T}\Delta t\right) & \sin\left(2\frac{2n}{T}\Delta t\right) & \sin\left(3\frac{2n}{T}\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}\Delta t\right) \\ \sin\left(\frac{2n}{T}2\Delta t\right) & \sin\left(2\frac{2n}{T}2\Delta t\right) & \sin\left(3\frac{2n}{T}2\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}2\Delta t\right) \\ \sin\left(\frac{2n}{T}3\Delta t\right) & \sin\left(2\frac{2n}{T}3\Delta t\right) & \sin\left(3\frac{2n}{T}3\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}3\Delta t\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sin\left(\frac{2n}{T}n\Delta t\right) & \sin\left(2\frac{2n}{T}n\Delta t\right) & \sin\left(3\frac{2n}{T}n\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}n\Delta t\right) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} \cos\left(\frac{2n}{T}\Delta t\right) & \cos\left(2\frac{2n}{T}\Delta t\right) & \cos\left(3\frac{2n}{T}\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}\Delta t\right) \\ \cos\left(\frac{2n}{T}2\Delta t\right) & \cos\left(2\frac{2n}{T}2\Delta t\right) & \cos\left(3\frac{2n}{T}2\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}2\Delta t\right) \\ \cos\left(\frac{2n}{T}3\Delta t\right) & \cos\left(2\frac{2n}{T}3\Delta t\right) & \cos\left(3\frac{2n}{T}3\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}3\Delta t\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \cos\left(\frac{2n}{T}n\Delta t\right) & \cos\left(2\frac{2n}{T}n\Delta t\right) & \cos\left(3\frac{2n}{T}n\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}n\Delta t\right) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \quad (25)$$

or

$$\mathbf{F} = a_0 \mathbf{I} + \mathbf{A}\mathbf{S} + \mathbf{B}\mathbf{C}$$

となる。ただし、 \mathbf{I} は n 次の単位列ベクトルである。

また、(25)式右辺の係数 a 、ベクトル \mathbf{S} と \mathbf{C} は、

$$a_0 = \frac{1}{n} \mathbf{I}^T \mathbf{F}, \mathbf{S} = \frac{2}{n} \mathbf{A}^T \mathbf{F}, \mathbf{C} = \frac{2}{n} \mathbf{B}^T \mathbf{F} \quad (26)$$

で与えられる。

さて、離散値系でフーリエ係数を計算する過程で、(25)式を

$$\mathbf{F} = a_0 \frac{1}{\sqrt{n}} \mathbf{I} + \sqrt{\frac{2}{n}} \mathbf{A} \mathbf{S} + \sqrt{\frac{2}{n}} \mathbf{B} \mathbf{C} \quad (27)$$

$$= a_0 \mathbf{I}^T + \mathbf{A}' \mathbf{S} + \mathbf{B}' \mathbf{C}$$

と書き直すと、

$$a_0 = \frac{1}{\sqrt{n}} \mathbf{I}' \mathbf{T} \mathbf{F}, \mathbf{S} = \sqrt{\frac{2}{n}} \mathbf{A}'^T \mathbf{F}, \mathbf{C} = \sqrt{\frac{2}{n}} \mathbf{B}'^T \mathbf{F} \quad (28)$$

の関係が成り立つ。 \mathbf{I} は n 次の単位行列である。この結果は、離散値系で正弦波や余弦波のフーリエ係数を求める場合、(28)式の係数行列 \mathbf{A}' 、 \mathbf{B}' の逆行列がそれぞれの転置行列で与えられることを意味する。言い換えれば、変換行列の逆行列が変換行列の転置行列で与えられることでフーリエ係数の直交性が満足される。

離散値系のウェーブレット変換は、連続系ウェーブレット変換の基底関数に対応する変換行列が常に直交性を満足する線形変換の一種である。ただし、ウェーブレット変換に用いる変換行列は、正弦波や余弦波のように連続な基底関数をサンプリングして構築されず、最初から離散値系で与えられる変換行列である。これは、従来の離散値系が、サンプリング個数を無限大にした場合、連続系に化することを前提にしているのに対し、サンプリング個数を無限大にしても連続系に化さない離散値系を前提とすることを意味する。換言すれば、離散値系のウェーブレット変換は、線形変換を行う演算が計数型計算機で行われることを大前提としている。この意味で離散値系ウェーブレット変換は、従来の離散値系と一線を画する離散値系である。

4.2 離散値系ウェーブレット変換の考え方

4.2.1 ウェーブレット変換行列

いま、 a と b なる数値を考えてみると、この数値の線形的な組合せは、

$$\begin{aligned} s &= a + b \\ d &= a - b \end{aligned} \quad (29)$$

が考えられる。さて、逆に s と d が与えられ a と b を求めようとするれば、

$$\begin{aligned} a &= \left(\frac{1}{2}\right)(s+d) \\ b &= \left(\frac{1}{2}\right)(s-d) \end{aligned} \quad (30)$$

となる。

(29)式の関係を行列を使って書くと、

$$\mathbf{X}' = C \cdot \mathbf{X} \quad (31)$$

と書ける。

ここで、 \mathbf{X}' , C , \mathbf{X} はそれぞれ、

$$\mathbf{X}' = \begin{bmatrix} s \\ d \end{bmatrix}, \mathbf{X} = \begin{bmatrix} a \\ b \end{bmatrix}, C = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (32)$$

である。他方、(30)式は、

$$\mathbf{X} = D \cdot \mathbf{X}' \quad (33)$$

と書ける。ただし、行列 D は、

$$D = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (34)$$

である。行列 C と D との関係は、

$$D \cdot C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (35)$$

となる。すなわち D は C の逆行列になっている。しかし、(34)式の行列の係数 $(1/2)$ を C と D に平等に分配して、

$$C = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, D = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (36)$$

とすれば、行列 D は C の転置行列であり、かつ逆行列となる。すなわち、

$$C^{-1} = C^T \quad (37)$$

である。次に a 、 b の 2 個の数値だけでなく、 a 、 b 、 c なる数値の組合せを考える。この場合、

$$\begin{aligned} s_1 &= a+b, s_2 = b+c, s_3 = c+a \\ d_1 &= a-b, d_2 = b-c, d_3 = c-a \end{aligned} \quad (38)$$

等の組合せが考えられる。この関係を行列で書くと、

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad (39)$$

となる。(39)式右辺の係数行列は 6 行 3 列の長方形行列であるため、(37)式のようにうまく逆行列を得られない。これは、組み合わせる数値の個数が奇数であると(33)-(37)式のような線形変換が適用できないことを意味する。では、 a 、 b 、 c 、 d の 4 個の数値の組みを考える。最初に考えられる組合せは、

$$\begin{bmatrix} s_1 \\ d_1 \\ s_2 \\ d_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad (40)$$

となる。これには、当然、

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} s_1 \\ d_1 \\ s_2 \\ d_2 \end{bmatrix} \quad (41)$$

が成り立つ。しかし、 a と b 、 c と d それぞれの組合せはできるが、 a と c 、 b と c 等の組合せはできない。このため、(38)式の左辺を並べ換えて、次の組合せを得る。

$$\begin{bmatrix} S_1 \\ D_1 \\ d_1 \\ d_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ d_1 \\ d_2 \end{bmatrix} \quad (42)$$

(40)式で、 a 、 b 、 c 、 d の4個の数値の組合せがすべて揃ったこととなり、結局、もとの $\{a, b, c, d\}^T$ のベクトルは(38)と(40)式の2回の線形変換で、和 S_1 と差 D_1 、 d_1 、 d_2 を要素とするベクトル $\{S_1, D_1, d_1, d_2\}^T$ に変形された。(40)式の逆変換は

$$\begin{bmatrix} s_1 \\ s_2 \\ d_1 \\ d_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} S_1 \\ D_1 \\ d_1 \\ d_2 \end{bmatrix} \quad (43)$$

によって行われ、(41)式の左辺を(39)式右辺のベクトルの形に並べ換えて、(39)式に代入することで完全な逆変換ができる。(38)式から(40)式までの変換をまとめて書くと、

$$\begin{bmatrix} S_1 \\ D_1 \\ d_1 \\ d_2 \end{bmatrix} = \left(\frac{1}{\sqrt{2}} \right)^2 \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \quad (44)$$

または、

$$\mathbf{S} = \mathbf{W} \cdot \mathbf{X} \quad (45)$$

となる。 \mathbf{W} はウェーブレット変換行列であり、この変換行列を構成する基底関数をハール(Haar)基底とよぶ。

(43)式の逆変換は \mathbf{W} の転置行列 \mathbf{W}^T が \mathbf{W} の逆行列になるから、

$$\mathbf{X} = \mathbf{W}^T \cdot \mathbf{S} \quad (46)$$

で与えられる。

次に、6個の数値の組合せを考えると、(38)式と同じ形で第1回の変換は、和が3組と差が3組となり、(27)式と同様に第2回の変換はできない。したがって、数値の組を和と差に分解する線形変換は、組を構成する数値の個数が、偶数でかつ、2、4、8、16、32等、2のべき乗でなければならない。

4.2.2 基底関数(base function)

いま、 n (2のべき乗)次のベクトルを \mathbf{X} とし、このベクトルに左から n 次の正方行列 \mathbf{C} を掛け算する。すなわち、

$$\mathbf{X}' = \mathbf{C}\mathbf{X} \quad (47)$$

なる線形変換を考える。行列 \mathbf{C} が次式の形で与えられるとすれば、

$$C = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 & 0 & 0 & \cdot & 0 & 0 & 0 & 0 \\ c_1 & -c_2 & c_1 & -c_0 & 0 & 0 & \cdot & 0 & 0 & 0 & 0 \\ 0 & 0 & c_0 & c_1 & c_2 & c_1 & \cdot & 0 & 0 & 0 & 0 \\ 0 & 0 & c_3 & -c_2 & c_1 & -c_0 & \cdot & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdot & c_0 & c_1 & c_2 & c_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdot & c_3 & -c_2 & c_1 & -c_0 \\ c_2 & c_3 & 0 & 0 & 0 & 0 & \cdot & 0 & 0 & c_0 & c_1 \\ c_1 & -c_0 & 0 & 0 & 0 & 0 & \cdot & 0 & 0 & c_3 & -c_2 \end{pmatrix} \quad (48)$$

となる。(48)式で第1行はベクトルの要素1から4までにそれぞれ係数 c_0, c_1, c_2, c_3 を重みとする平均値を取ることを意味する。第2行はベクトルの要素1から4までにそれぞれ係数 c_0, c_1, c_2, c_3 を重みとする差分を取ることを意味する。要するに(48)式の第1行は重みをつけた積分演算に対応するデジタルフィルターであり、第2行は重みをつけた微分演算に対応するデジタルフィルターである。第3,4行はそれぞれベクトルの要素3から6までに対する積分と微分演算を行うことを意味する。積分と微分演算はベクトルの2要素ずつシフトして循環する形で行われる。

つぎに、直交性を満足するために、

$$C^T C = I \quad (49)$$

となるように(41)式の要素を決めることを考える。 I は c と同じ次数の単位行列である。 C_7 は、

$$C^T = \begin{pmatrix} c_0 & c_3 & 0 & 0 & \cdot & 0 & 0 & 0 & 0 & c_2 & c_1 \\ c_1 & -c_2 & 0 & 0 & \cdot & 0 & 0 & 0 & 0 & c_3 & -c_0 \\ c_2 & c_1 & c_0 & c_3 & \cdot & 0 & 0 & 0 & 0 & 0 & 0 \\ c_3 & -c_0 & c_1 & -c_2 & \cdot & 0 & 0 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & c_2 & c_1 & c_0 & c_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdot & c_1 & -c_0 & c_1 & -c_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdot & 0 & 0 & c_2 & c_1 & c_0 & c_3 \\ 0 & 0 & 0 & 0 & \cdot & 0 & 0 & c_3 & -c_0 & c_1 & -c_2 \end{pmatrix} \quad (50)$$

であるから、

$$\begin{aligned} c_0^2 + c_1^2 + c_2^2 + c_3^2 &= 1 \\ c_2c_0 + c_3c_1 &= 0 \end{aligned} \quad (51)$$

が成り立つ。しかし、(51)式から係数を決めることはできない。このため、次の条件を追加する。

$$\begin{aligned} c_3 - c_2 + c_1 - c_0 &= 0 \\ 0c_3 - 1c_2 + c_1 - 3c_0 &= 0 \end{aligned} \quad (52)$$

(51), (52)式から、

$$c_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, c_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, c_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, c_3 = \frac{1-\sqrt{3}}{4\sqrt{2}} \quad (53)$$

として(41)式の要素が求められる。これをドビッシー (Daubechies) の4次アナライジングウェーブレット (基底) 関数と呼ぶ。

4.2.3 ウェーブレット変換の性質

アナライジングウェーブレットの係数を決めるために導入した拘束条件である(35)式で、最初の条件は入力データが一定値であれば常にゼロとなることを意味する。第2の条件は入力データが単調増加であっても常にゼロであることを意味する。したがって、入力データが2次関数以上の変化率を持つとき、第1段階の差分演算を受けた項はゼロでない値を持つ。第2段階は重み付き積分もしく

は平均値の項で2次関数以上の変化率を抽出する演算となる。このことから、ウェーブレット変換は変化率の大きな部分のみを抽出する演算であり、画像データなどでは画像の輪郭やコントラストの高い部分がゼロでないウェーブレットスペクトラムを与えることとなる。このため、画像データのすべてを記憶しなくても、画像の輪郭やコントラストの高い部分のみを記憶することが可能となり、結果として画像圧縮がなされる。画像データでなくより一般的な波形データで考えれば、波形データの持つ変化率が急峻な特徴のみを抽出するのにきわめて有効なデジタルフィルター演算とも解釈できる。

4.3 ウェーブレット変換による1次元圧縮

ウェーブレット変換によるデータ圧縮の原理は、もとのデータの特徴を表す絶対値の大きなスペクトラムのみを残し、他をゼロとすることにある。

4.3.1 例題

図7のようなサンプルデータを用意する。

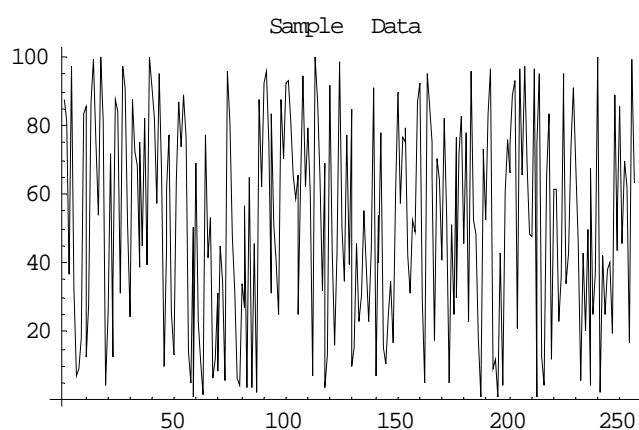


Fig.7 1次元サンプルデータ

このサンプルデータをウェーブレット変換したスペクトラムを図8に示す。尚、ウェーブレット変換にはハール基底を用いる。

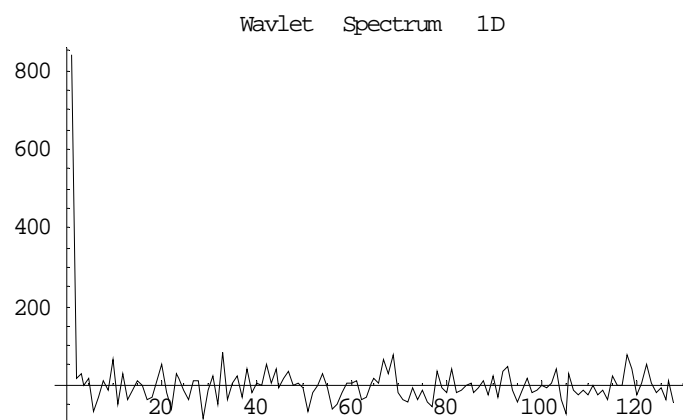


Fig8 . ウェーブレットスペクトラム

このスペクトラムをマザーウェーブレット近傍 128 までを残し、あとをゼロとして、逆ウェーブレット変換を行ったデータを図 9 に示す。

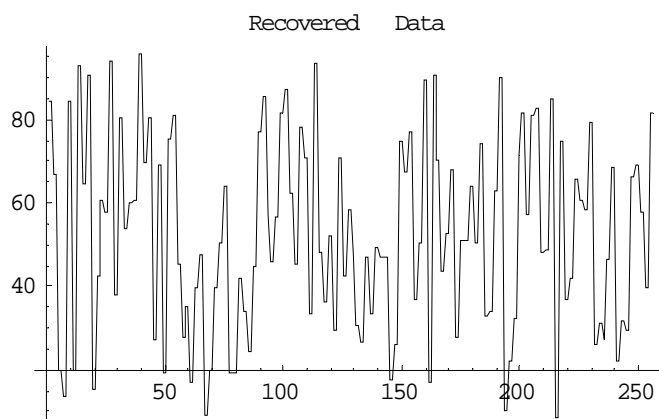


Fig 9 . 復元されたデータ

図 7 と図 9 の相関係数は 0.76 となり、マザーウェーブレット近傍に情報を集められることがわかる。

本論文では、R,G,B 各成分の色情報である 256 個のデータをウェーブレット変換により圧縮し、それを新しい固有パターンとして扱い、画像認識を試みる。

4.4 静止画像認識

4.3 で勘案した方法で画像認識を行った結果を述べる。入力画像とデータ画像は 3.3.1、3.3.2 とし、表の見方は 3.3.3 と同じとする

Table.2 認識結果 2

入力 No.	出力 No.	判定	入力 No.	出力 No.	判定
No.1	No.1		No.17	No.33	
No.2	No.3		No.18	No.35	
No.3	No.22	×	No.19	No.38	

No.4	No.7		No.20	No.40	
No.5	No.9		No.21	No.41	
No.6	No.11		No.22	No.43	
No.7	No.37	×	No.23	No.45	
No.8	No.22	×	No.24	No.48	
No.9	No.9	×	No.25	No.49	
No.10	No.20		No.26	No.52	
No.11	No.21		No.27	No.54	
No.12	No.22	×	No.28	No.55	
No.13	No.26		No.29	No.58	
No.14	No.27		No.30	No.60	
No.15	No.27	×	No.31	No.61	
No.16	No.31		No.32	No.63	

表 2 より 32 枚中 27 枚の認識し、認識率は 83.4% となった。認識率のていかとしては、圧縮により情報の欠如が大きなものと考えられるが、圧縮しない時と比べると、最小二乗法を解く際の計算は相当速くなった。

5. まとめ

本論文では、まず画像の普遍量とは RGB 成分で構成されていることを示しその固有パターンの抽出方法を述べた。そうすることにより、画像を構成する画素数、画像の x- y 直交座標上の位置・角度といったハードウェアに依存する性質を削減することができ、画像の情報を抽出し静止画像認識を行う手法を述べた。

画像の固有パターンからなる線形システム方程式は不適切であるため、最小二乗法により解を導出する手法を提案した。さらに RGB 成分それぞれを固有パターンをウェーブレット変換によって圧縮をし、解を求める際の計算速度向上を試みた。よって、本論文で提案した手法は、元来の固有パターンを圧縮して情報量を少なくした場合でも認識率はそれほど下がらず、計算速度の向上が望めることを報告する。

6. 参考文献

- [1] 斎藤兆古著、"Mathematicaによるウェーブレット変換"、朝倉書店
- [2] 斎藤兆古著、"ウェーブレット変換の基礎と応用 Mathematicaで学ぶ"、朝倉書店

2005 年度卒業論文

フーリエ変換による動物の声紋解析

法政大学

工学部

情報電気電子工学科

学籍番号 0 2 D3079

氏名 野崎 高広

指導教授 斎藤 兆古

目次

第 1 章 序論	2
第 2 章 フーリエ変換の基礎	3
2 . 1 概要	
2 . 2 関数の直交性と線形性	
第 3 章 データの間引き数による相関係数の違い	7
第 4 章 他動物との比較	1 0
第 5 章 まとめ	1 3
第 6 章 参考文献	1 4

第 1 章 序論

現在、人間の声に含まれる声紋データをもとに言語認識、または特定の人物認識の研究が進み実用化されています。しかし、動物の音声による認識はあまり耳にしません。その理由として、声帯を持たず体内の反響を利用して声を出す動物は体の大きさによって声が似てしまうからであり、例としてトラとライオンの声が区別不能だということです。もうひとつの理由として動物の音声の収録が困難だということです。動物は録音者の意思に応じて声を出すことが難しく、撮影場所も一定の環境での録音は非常に困難だと考えられるからです。

本論文では、今後の動物認識の研究のため、各動物の音声をフーリエコサイン変換により取り出した周波数情報の特徴の違いを見だし、間引きしたデータ数によるそれぞれのフーリエスペクトラムを相関係数法により相違を研究し、他の動物との周波数情報の比較を行いました。

第 2 章 フーリエ変換の基礎

2 . 1 . 1 概要

古典的な関数変換の目的は、解析的に扱いにくい関数系を解析的に扱いやすい関数系へ変換することです。例えば、ラプラス(Laplace)変換は微積分演算を単純な掛け算や割り算へ化す変換です。また、関数系の変換とは、ある関数、例えば時間変化する関数を解析が容易な周期関数の和で表現することにも使われます。具体的な例としてフーリエ(Fourier)変換を考えます。フーリエ変換は解析的に扱えない関数を解析的に扱える角周波数の異なる正弦波と余弦波の和で表現する変換です。換言すれば、フーリエ変換は解析的に扱いにくい関数系を解析的に扱いやすい関数系へ分解する変換と考えてもよいでしょう。

ウェーブレット変換は、ラプラス変換のように解析的な計算を容易な形へ変換する方法でなく、フーリエ変換のように与えられた関数を他の関数系の和で表現するために用いられます。いま、ある任意の時間 t をパラメータとする関数 $f(t)$ を一定値 a_0 、正弦波および余弦波の和で表現できるとします。すなわち、 ω を角周波数として、

$$f(t) = a_0 \sum_{i=1}^{\infty} a_i \sin(i\omega t) + \sum_{i=1}^{\infty} b_i \cos(i\omega t) \quad (1)$$

と仮定します。

問題は(1)式の係数、 a_0, a_i, b_i を決める方法です。いま、関数 $f(t)$ が時間 $t=0$ から $t=T$ の区間で(1)式の係数を決めることを考えれば、(1)式は、 $\omega=2\pi/T$ であるから、

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin(i \frac{2\pi}{T}t) + \sum_{i=1}^{\infty} b_i \cos(i \frac{2\pi}{T}t) \quad (2)$$

とも書けます。

(2)式の両辺を時間 $t=0$ から $t=T$ の区間について積分すると、

$$\begin{aligned} \int_0^T f(t)dt &= \int_0^T (a_0 + \sum_{i=1}^{\infty} a_i \sin(i \frac{2\pi}{T}t) + \sum_{i=1}^{\infty} b_i \cos(i \frac{2\pi}{T}t))dt \\ &= a_0 T \end{aligned} \quad (3)$$

が成り立ちます。

したがって、定数項 a_0 は

$$a_0 = \frac{1}{T} \int_0^T f(t)dt \quad (4)$$

となります。

次に、(2)式の両辺に正弦波 $\sin[j(2\pi/T)t]$ 、 $f=1,2,3,\dots$ と掛け算し、時間 $t=0$ から $t=T$ の区間について積分する。

$$\begin{aligned} \int_0^T f(t) \sin(j \frac{2\pi}{T}t) dt &= \int_0^T (a_0 + \sum_{i=1}^{\infty} a_i \sin(i \frac{2\pi}{T}t) + \sum_{i=1}^{\infty} b_i \cos(i \frac{2\pi}{T}t)) \times \sin(j \frac{2\pi}{T}t) dt = a_i \frac{2}{T}, \\ & \quad i = j \text{ or } 0, \\ & \quad i \neq j \end{aligned} \quad (5)$$

から係数 a_0 は

$$a_i = \frac{2}{T} \int_0^T f(t) \sin(i \frac{2\pi}{T}t) dt \quad (6)$$

として得られます。

同様に、(2)式の両辺に余弦波 $\cos[j(2\pi/T)t]$ 、 $f=1,2,3,\dots$ を掛け算し、時間 $t=0$ から $t=T$ の区間について積分することで、

$$b_i = \frac{2}{T} \int_0^T f(t) \cos(i \frac{2\pi}{T}t) dt \quad (7)$$

として係数 b_i が得られます。

したがって、関数 $f(t)$ は区間 $t=0$ から $t=T$ で、

$$f(t) = \frac{1}{T} \int_0^T f(t) dt + \sum_{i=1}^{\infty} \left(\frac{2}{T} \int_0^T f(t) \sin(i \frac{2\pi}{T} t) dt \right) \sin(i \frac{2\pi}{T} t) + \sum_{i=1}^{\infty} \left(\frac{2}{T} \int_0^T f(t) \cos(i \frac{2\pi}{T} t) dt \right) \cos(i \frac{2\pi}{T} t)$$

(8)

と書けます。これがいわゆるフーリエ変換の原型となるフーリエ級数であり、左辺の関数 $f(t)$ を右辺の計算が簡単な定数項と三角関数の和に変換しています。

2. 1. 2 関数の直交性と線形性

フーリエ級数の考え方の中に重要な関数間で成り立つ性質、すなわち、関数の直交性 (orthogonality) が使われています。

まず (4) 式の係数 a_0 が計算される過程を考えます。(3) 式は (2) 式の両辺に定数値 1 を掛け算し積分する演算です。このとき、

$$\begin{aligned} \int_0^T 1 \cdot \sin(i \frac{2\pi}{T} t) dt &= 0, \\ \int_0^T 1 \cdot \cos(i \frac{2\pi}{T} t) dt &= 0, \\ i &= 1, 2, \dots \end{aligned} \tag{9}$$

の関係が成り立つために、(4) 式の係数 a_0 がうまく計算できます。この関係を、定数値 1 と正弦波 $\sin[f(2\pi/T)t]$ 間の直交性と呼びます。同様に (6), (7) 式で計算される a_i, b_i は

$$\begin{aligned}
\int_0^T \sin(i \frac{2\pi}{T} t) \sin(j \frac{2\pi}{T} t) dt &= 0, \\
\int_0^T \cos(i \frac{2\pi}{T} t) \cos(j \frac{2\pi}{T} t) dt &= 0, \\
i &\neq j
\end{aligned}
\tag{10}$$

なる直交性が成り立つことに基づいています。

(9),(10)式から、直交性とは与えられた関数を他の関数の和で表現しようとする場合、和となる関数の大きさ(係数)を一意的に決める条件であることがわかります。言い換えれば、与えられた関数を他の関数の和で表現できる条件であります。

さて、ある任意の時間 t をパラメータとする関数 $f(t)$ は、区間 $t=0$ から $t=T$ で、一定値、正弦波および余弦波の輪で表現できることがわかりました。この変換は一定値、正弦波および余弦波間で直交性が成り立つことが条件でした。この結果に至る過程を考えると、まず、展開される関数の和でもとの関数が表現されるとする大前提があることに気づくでしょう。ある関数が他の関数の和で表現できる性質を線形性と呼びます。では、(6)式で計算される正弦波の係数 a_i が定数 c_i の n 個の和で表現できるとします。すなわち、

$$a_i = c_i + c_i + \dots = n c_i \tag{11}$$

です。(11)式から、和を前提とする系では比例関係が成り立つことを意味することがわかります。すなわち、線形系とは比例関係が成立する系です。

第 3 章 データの間引き数による相関係数の違い

相関係数により動物のなき声の違いを見出したいのですが、音声のデータ数を録音の際完璧にそろえることは不可能に近いと考えられます。そのため、この章では音声データのボリュームの最大値から 1/2 の部分を始点として、そこからのデータ数によって同種類の動物データ（象、ライオン、馬、猫、犬）の相関係数がどのような傾向をとるのか研究をしました。同種類の動物データは 2 つずつ用意し、互いのデータ数をそれぞれ 1 2 8、2 5 6、5 1 2、1 0 2 4、2 0 4 8、4 0 9 6 個にそろえ相関係数法に計算しました。表 1 は実験結果であり空白の部分はデータ数が不足したため計算を行えませんでした。

表 1

データ数	猫	犬	象	ライオン	馬
1 2 8	-0.0177	0.2056	0.2723	0.8547	-0.0147
2 5 6	0.0776	0.2000	0.2993	0.4922	-0.0138
5 1 2	0.2678	0.1730	0.2597	0.5535	0.0124
1 0 2 4	0.3900	0.2634	0.3662	0.3141	0.0765
2 0 4 8	0.2390	0.3236	0.4452	0.2375	0.0608
4 0 9 6		0.3382	0.5099	0.2409	0.0313
8 1 9 2		0.3915	0.5203	0.2560	0.0307
1 6 3 8 4			0.4600	0.1867	0.0139

実験の結果、犬、象はデータ数が多いほど、猫、馬は1024個から2048個付近、そしてライオンはデータ数が少ないほど相関係数が高くなりました。フーリエスペクトラムを見ると、単調なスペクトラムほどデータ数が少ないほど相関係数が高く、複雑なスペクトラムほどデータ数が多いほど相関係数が高くなると考えられます。しかし、相関係数の最大値が平均で0.5を下回り、馬にいたっては0.1未満と、とても低い結果でした。図1から5に元データのフーリエコサインスペクトラムを示します。

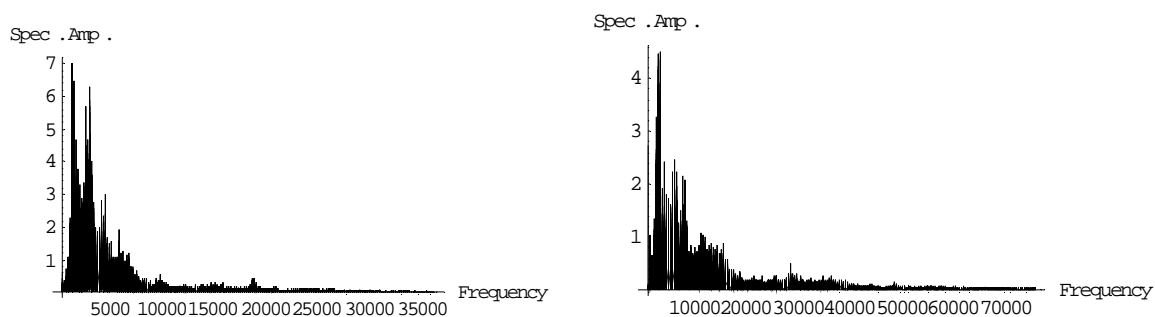


図1 フーリエコサインスペクトラム(象)

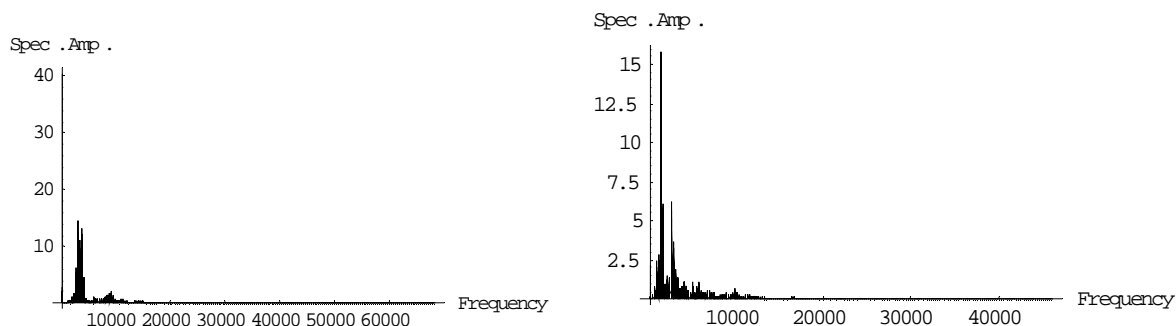


図2 フーリエコサインスペクトラム(ライオン)

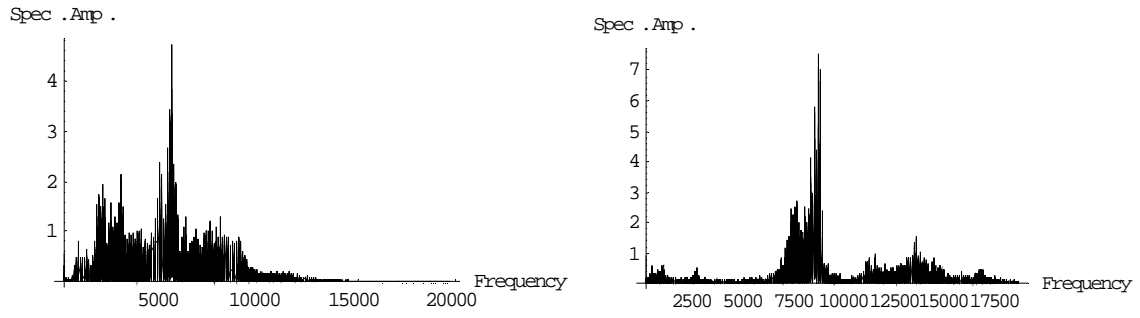


図 3 フーリエコサインスペクトラム(馬)

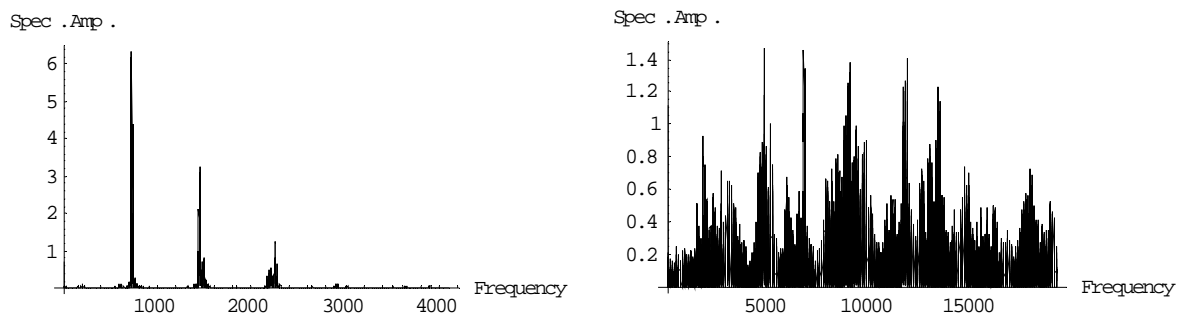


図 4 フーリエコサインスペクトラム(猫)

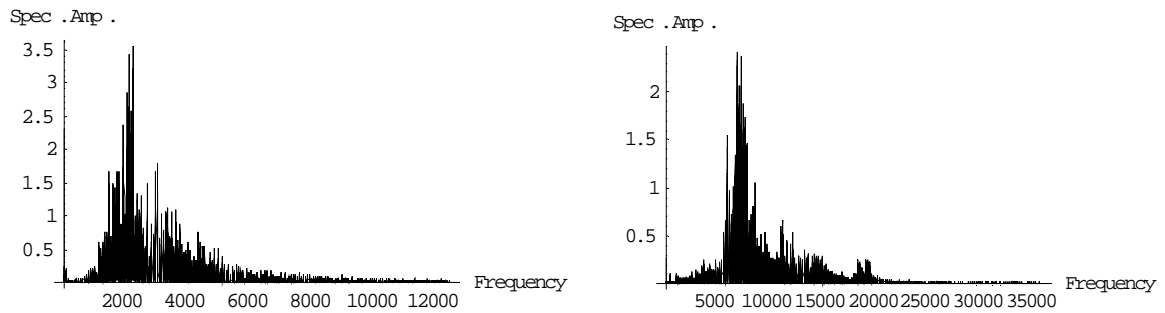


図 5 フーリエコサインスペクトラム(犬)

第 4 章 他動物との比較

本章では各動物がどの動物との相関係数が高いかを研究しました。音声データを動物ごとに入力データ、対象データに分け、データ数を 2048 個とし、フーリエコサイン変換のスペクトラムの相関係数を研究しました。図 6 から 10 はデータ数 2048 個にした各々のフーリエコサインスペクトラムです。ここで、図の左が入力データで右が対象データです。

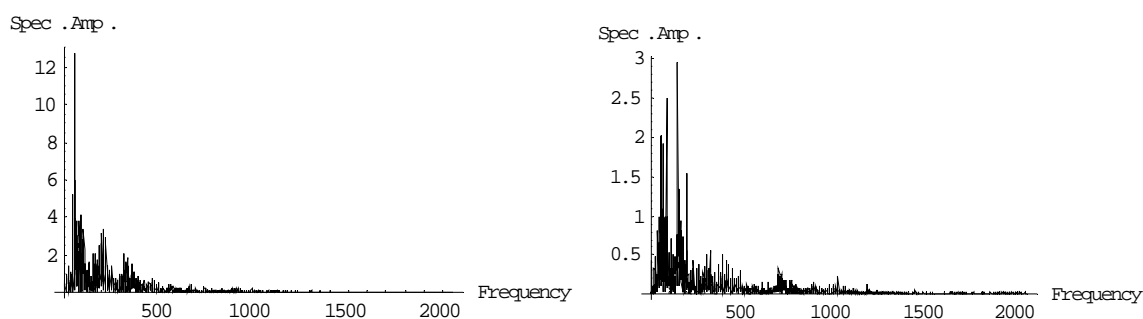


図 6 フーリエコサインスペクトラム(象)

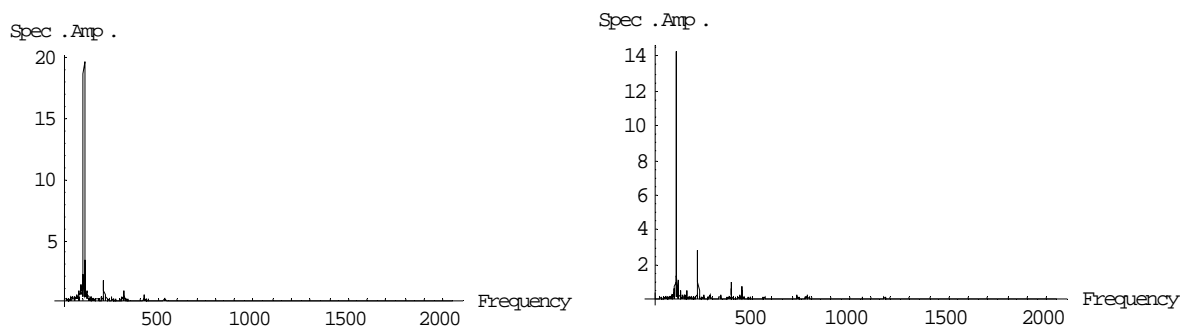


図 7 フーリエコサインスペクトラム(ライオン)

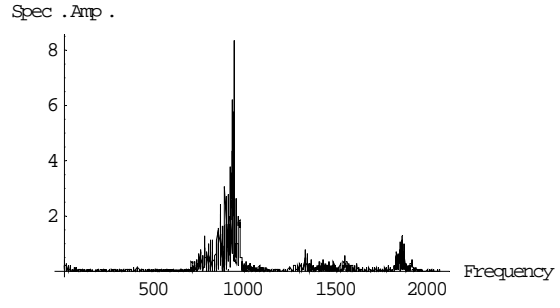
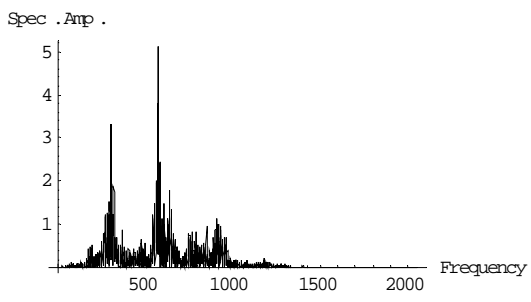


図 8 フーリエコサインスペクトラム(馬)

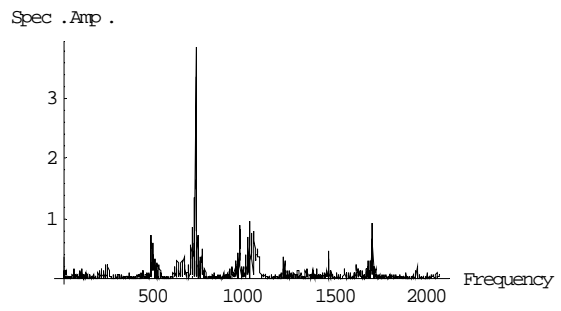
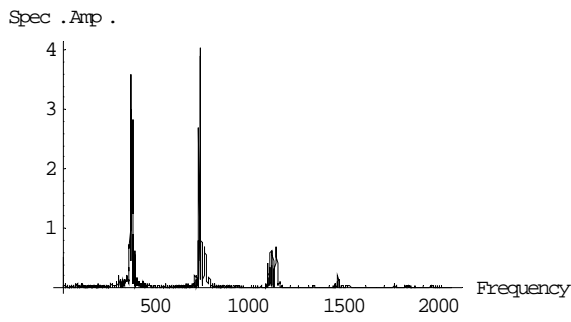


図 5 フーリエコサインスペクトラム(猫)

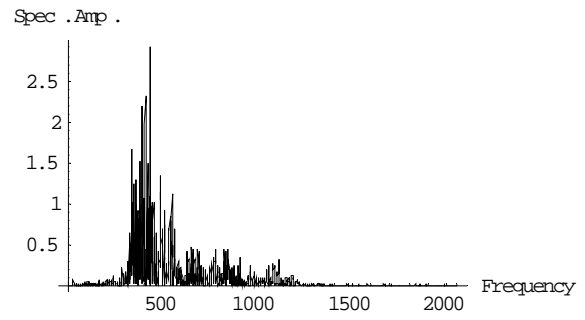
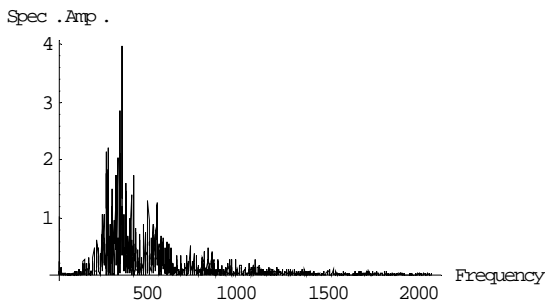


図 5 フーリエコサインスペクトラム(犬)

この実験の結果を表 2 に示す。

表 2

	犬 2	猫 2	象 2	ライオン 2	馬 2
犬 1	0.3236	0.1061	0.0632	-0.0158	0.1551
猫 1	-0.0199	0.2390	-0.0188	-0.0186	-0.0087
象 1	0.0858	0.0177	0.4452	0.2436	0.0383
ライオン 1	0.0789	0.0177	0.1029	0.2374	-0.0188
馬 1	-0.0634	-0.0174	-0.0648	-0.0294	0.0608

本章の実験の結果から同一の動物の相関係数がライオンを除いて最も大きな値を示しました。しかし、相関係数の数値が最大値でさえ 0.5 未満であるため認識のための数値としてはとても低いものです。

第 5 章 まとめ

本論文では動物の音声データの識別をする場合、データ数による相関係数の違いを調べ、また、データ数を一定とした場合のそれぞれの相関係数について研究しました。その結果、同種どうしの相関係数の数値が一番高い値を示したが相関係数の値は低いという結果となりました。そのため、周波数情報の特徴の大きい部分を取り出すことができれば認識が可能ではないかと考えられます。

声帯を持たない動物の鳴き声は体を利用して音声を模擬します。このため、体の条件、お腹が一杯か？姿勢などに依存性が大きいため、フーリエ変換を用いた声紋認識・識別は適用が困難であり、特にフーリエ変換やウェーブレット変換などでスペクトラム抽出は無意味に近いことが判明しました。

今回の実験で使用した各動物の音声データはインターネット上で公表されていたものを使用したため、録音状況や録音するための機器の性能にばらつきがみられました。そのため、雑音やマイクのリミッターを超える音声を録音してしまうデータもありました。このため、現時点で得られた本論文の結論は生の音声を録音機器や録音状況を統一して収録を行ったデータにより認識・識別を行うのが適切と考えられます。これは、動物が人間のように言語を持たない事が大きな原因と考えられます。

第 6 章 参考文献

- [1] 斎藤兆古著、“Mathematicaによるウェーブレット変換”、朝倉書店
- [2] 斎藤兆古著、“ウェーブレット変換の基礎と応用 Mathematicaで学ぶ”、朝倉書店

2005 年度
卒業論文

**中国に於ける
電磁界解析応用に関する文献調査
～ 清華大学博士論文を通して～**

法政大学

工学部 電気電子工学科

学籍番号 01d3030

氏名 岡村 聡

指導教授 斎藤 兆古

目次

1.	序論	-2-
2.	埋設導体系統の誘導電流分布と電磁遮蔽計算について	-3-
2. 1	計算モデルについて	
2. 2	埋設物の離散化モデル (遮蔽層電流を考慮しない場合)	
2. 2. 1	区間導体間での電界解析	
2. 2. 2	区間導体間での磁気結合解析	
2. 2. 3	区間導体間での横断面の電流補間関数	
2. 2. 4	埋設物の離散化方程式	
2. 3	金属管内の被覆物の離散化モデル	
2. 4	ケーブル系統の電流分布の解	
2. 5	通信ケーブルの電力線における電波妨害	
2. 6	計算例	
3.	地中の通信ケーブルを金属管で覆った場合と付近に 遮蔽線を敷く場合の効果の比較	-18-
3. 1	計算例と遮蔽効果の解析	
3. 2	両端の接地抵抗が遮蔽効果を及ぼす影響	
3. 3	金属管の外直径及び厚さが遮蔽効果を及ぼす影響	
3. 4	遮蔽線の材料や付設した位置が遮蔽効果を及ぼす影響	
4.	考察	-22-
5.	参考資料	-23-
6.	参考文献	-28-

1 序論

電力線が地中の通信ケーブル回線に危険な影響を与える解析、計算を行うことは電力線をどこに敷くか選ぶ際、とても重要な意義を持っている。危険な影響が一定の基準を超える場合、正常な運転状況では、電力線の電磁界は近くの通信回線に対して妨害を生む可能性があり、通信回線に雑音が生じ、通信の品質を下げる；また故障した状態で、電力線の強大な短絡電流は電界*¹と磁界*²によって通信回線付近で数百ボルト更には数千キロボルトの過電圧を発生する。それにより通信回線の設備と人員の安全が危惧される。ここ数年来、110kv以上の電圧をもつ電力線が続々と市街区域に入っていく、電力線と通信回線の公用地下道とその両方が複雑に絡み合っており、電力線と通信感染の干渉はすでに避けられない状況であり電力線が通信回線の電磁的な悪影響を与える問題が日に日に際立っている

通常通信ケーブル付近に遮蔽線（シールド線）*³を設備または通信ケーブルを金属管で覆うなどの方法で電磁的遮蔽保護をする。遮蔽線の設備または金属管等で覆うことを同時に考慮して埋設された金属管、遮蔽層や被覆などで構成されている複数層からなる導体系統の電磁遮蔽計算法はまだ報告されていない。

電磁遮蔽の基本計算理論は特に複雑ではないが、金属管または遮蔽線が通信ケーブルの遮蔽（シールド）効果に対して正確な計算を行うには多くの問題を解決しなければいけない、したがって適当な計算モデルを必要するだけでなく、通信ケーブル、電力線、遮蔽線、配置方法また大地のパラメータ等が関係する。近代の計算方法から述べると、解析的方法は*⁴は通信ケーブルと電力線が平行または交差など簡単な状況のみ用いられ、伝送線路法*⁵は通信ケーブルと電力線が重なって平行する場合にも用いられる。通信ケーブルと電力線の配置は複雑で、電界*¹と磁界*²を同時に考慮する場合で地中のケーブル電磁遮蔽問題は電磁界の数値計算技術*⁶によって計算を行う。

本論文は電界と磁界を同時に考慮する場合の地中導体の電磁遮蔽について計算モデルを提唱し、そのモデルを用いて金属管および遮蔽線が地下ケーブルに対する電磁遮蔽計算を行う。計算結果より、金属管と遮蔽線が通信ケーブルに及ぼす遮蔽効果を解析、比較をした。

筆者は清華大学邹军（zoujun）講師の「通信回線が電力線によって受ける電磁的危険な影響」“电力线路对通信线路电磁感应危险影响的研究”（Investigation of Telecommunication Lines Against Electromagnetic Danger Effects From Electric Power Lines）の文献を翻訳することで内容を紹介していく

2 埋設導体系統の誘導電流分布と 電磁遮蔽計算について

本章の内容は埋設ケーブル系統の誘導電流分布と電磁遮蔽計算についてである。電磁界の基本方程式から始まり、埋設導体の外表面の電界強度に関する積分方程式を導き、導体に対して縦方向の電流は適当な補間関数を選ぶことで、電界強度の積分方程式を広域電気ネットワーク問題に置き換えることを述べる。

2.1 計算モデルについて

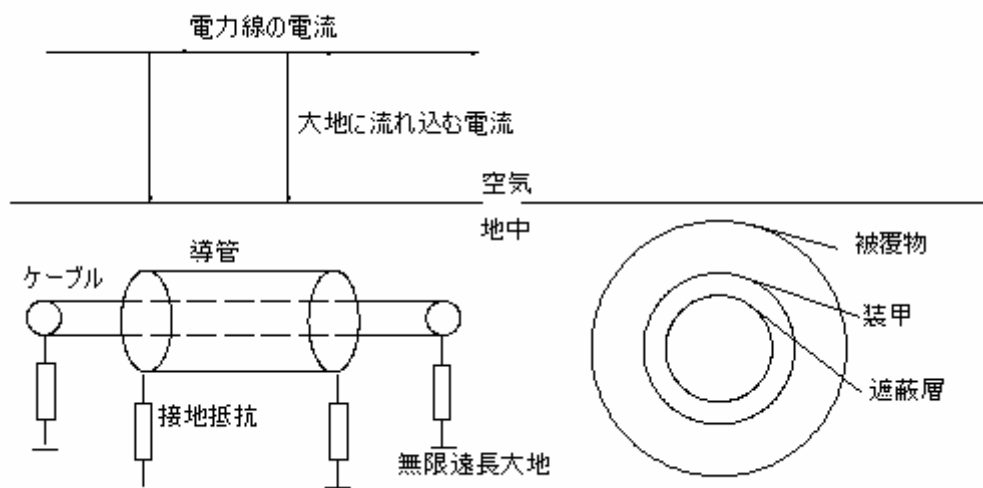


図1-1 図1-2 金属管—ケーブル系統の相關図（断面図）

図1-1と図1-2はそれぞれ架空電力線と金属管を覆った被覆ケーブル系統の側面図と断面図である。金属管、ケーブルの被覆物また遮蔽層より通信ケーブルの遮蔽導体を構成している。解析を明快にするため、以下の仮説をする

- 仮定（1）：遮蔽層、被覆物、金属管は同軸で一樣な管状の円柱体とする。計算、解析する場合、一樣な管状円柱体、つまり線状の導体とする
- 仮定（2）：遮蔽層と被覆物の間を絶縁体とし、その側面に流れる漏れ電流を省略することで、遮蔽層の電流は軸に沿って一樣に分布する

仮定（3）：被覆物と金属管の間の側面に流れる容量性電流を省略しその媒質中の漏れ電流のみ考え、よって被覆物－金属管間の側面の漏れインピーダンスは軸に一樣であると仮定する

遮断層の横断面の電流は距離によって変化はしないので離散化处理がされない。離散化される物体は三種類に分けられ

- (1)：直接大地と接触する金属管、被覆物は接地点を分散することを考慮しなければならない
- (2)：金属管内の被覆物と金属管の点接触、仮定3より媒質中の漏れ電流は物体の両端に集中すると近似できる
- (3)：遮蔽層の横断面の電流は一定に分布していることで、遮蔽層と金属管、被覆物の作用は電界のみである。したがって解析時に、まずは遮蔽層の電流を考慮しないで、金属管、被覆物を離散化した後にその影響を追加する。同時に、ケーブルの電流が電力線に対する反作用も省略し、電力線の電流は一樣とする

2. 2 埋設物の離散化モデル（遮蔽層電流を考慮しない場合）

埋設された金属管と埋設された被覆物（Armor）を同じ統一した記号で書き、物体の総数を N 、電力線を M 層、すべての層の電流を I_{Fn} ；電力線から大地に流れる電流を $M+1$ とし I_{En} とする。図2は埋設導体と電力線の座標図である、大地の表面で座標 (x, y, z) とし、ケーブル系統の中心を原点として成り立つ座標を $(0-\xi)$ とする。

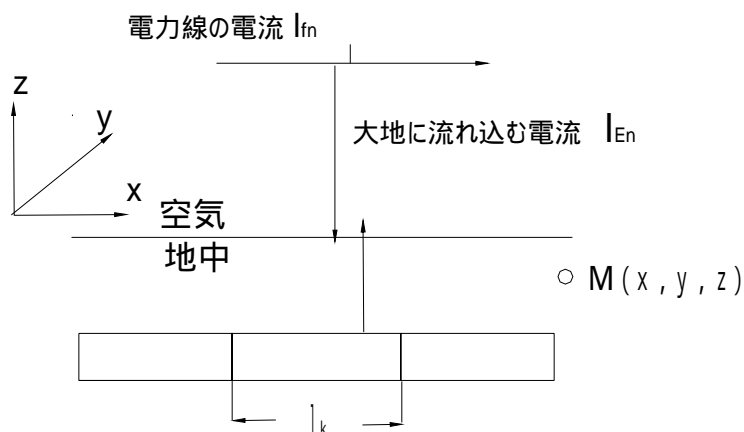


図2 電力線と埋設導体の概略図

$A_k(\xi)$ 、 $E_k(\xi)$ 、 $U_k(\xi)$ をそれぞれ導体外表面の第k区間での磁気ベクトルポテンシャル、電界強度と電位とする。正弦波定常条件下で、大地中のある1点での $A_e(x, y, z)$ 、 $E_e(x, y, z)$ 、 $U_e(x, y, z)$ を考えると、電磁界の理論によって、三者間が満足すべき式は

$$E_e(x, y, z) = -j\omega A_e(x, y, z) - \nabla U_e(x, y, z) \quad (1)$$

となる

M点が導体外表面に位置する場合、関数の連続性に基づき、 $A_e(x, y, z)$ 、 $U_e(x, y, z)$ が ξ 軸方向に沿った大きさは $A_{ek}(x, y, z) = A_k(\xi)$ 、 $U_{ek}(x, y, z) = U_k(\xi)$ となり、 $E_e(x, y, z)$ が ξ 軸方向に沿った E_{ek} は

$$E_{ek}(\xi) = -j\omega A_k(\xi) - \nabla U_k(\xi) \quad (2) \text{ となる}$$

導体外表面に沿った電界強度 $E_k(\xi)$ *7は

$$E_k(\xi) = Z_k I_k(\xi) \text{ と表される}$$

式中の Z_k は導体の外表面内での単位長あたりの抵抗となり、 $I_k(\xi)$ は埋設導体の縦向きの電流である。導体が埋設された大地は、 $E_{ek}(\xi) = E_k(\xi)$ 故に

$$Z_k I_k(\xi) + j\omega A_k(\xi) + \frac{\partial}{\partial \xi} U_k(\xi) = 0 \quad (3) \text{ となる}$$

第k区間での積分では電位のスカラ量の性質により、式(3)は

$$\int_{l_k} Z_k I_k(\xi_k) d\xi_k + \int_{l_k} j\omega A_k(\xi_k) d\xi_k + U_{k^r} - U_{k^l} = 0 \quad (4)$$

と書き直せる

式中 $\int_{l_k} j\omega A_k(\xi_k) d\xi_k$ は第k区間の縦向き誘導起電力であり、 k^l と k^r はkの左右の端点である。 U_{k^l} と U_{k^r} はk上の左右の両端が大地に対する電位の値を表している。

$A_k(\xi_k)$ と $U_k(\xi_k)$ はそれぞれ、1次側の電力線と2次側の誘導電流から生ずるポテンシャルの和からなり $A_k(\xi) = A_k^p(\xi) + A_k^s(\xi)$ $U_k(\xi) = U_k^p(\xi) + U_k^s(\xi)$ となる。よって式

(4) に代入することで

$$\int_{l_k} Z_k I_k(\xi_k) d\xi_k + \int_{l_k} j\omega A_k^S(\xi_k) d\xi_k + (U_{k^r}^P - U_{k^l}^P) = - \int_{l_k} j\omega A_k^P(\xi_k) d\xi_k - (U_{k^r}^P - U_{k^l}^P) \quad (5) \text{ となる。}$$

2.2.1 区間導体間での電界解析

電力線の電流 I_{Fn} が電柱区間毎に一樣である場合、第 k 区間での M_1 層の電力線が生じる縦向きの誘導起電力を合計すると、 $\int_{l_k} j\omega A_k^P(\xi_k) d\xi_k$ は

$$\int_{l_k} j\omega A_k^P(\xi_k) d\xi_k = \sum_{n=1}^M \tilde{Z}_{kn} I_{Fn} \quad (6) \text{ となる}$$

式中の \tilde{Z}_{kn} は第 M 層の電力線と第 k 区間の埋設ケーブルによる相互インピーダンス^{*4*5} である

$P(r_{ki})$ を第 i 区間での ξ_i と第 k 区間で ξ_k が位置する単位長の縦向きの相互インピーダンスとし、 r_{ki} を第 i 区間での ξ_i と第 k 区間で ξ_k 間の距離とする。 ϕ_{ki} を第 k 区間と第 i 区間の間でなす角とする。地中の金属管は一般にあまり深く埋設されないので、電界を計算する場合、金属管を大地の表面に位置するとみなし、大地を半無限大の様な電気媒質と仮定すると、変位電流を無視すれば、 $P(r_{ki})$ ^{*8} は

$$P(r_{ki}) = \frac{j\omega\mu_0}{2\pi r_{ki}} \left[\frac{1 - (1 + \gamma_1 r_{ki}) e^{-\gamma_1 r_{ki}}}{(\gamma_1 r_{ki})^2} \right] \quad (7)$$

式中、 μ_0 :真空誘磁率、 σ_1 :大地の導電率、 $\gamma_1 = \sqrt{j\omega\mu_0(\sigma_1 + j\omega\epsilon_1)} \approx \sqrt{j\omega\mu_0\sigma_1}$ を大地での変位電流を無視した場合の伝播定数とする

第 k 区分に位置する第 i 区分での電流 $I_i(\xi_i)$ より生じる磁気ベクトルポテンシャルを

$A_{ki}^p(\xi_k)$ とすると

$$\int_{l_k} j\omega A_k^s(\xi_k) d\xi_k = \int_{l_k} j\omega \left[\sum_{i=1}^N A_{ki}^s(\xi_k) \right] d\xi_k \quad (8)$$

式 (7) と式 (8) より

$$\begin{aligned} \int_{l_k} j\omega A_k^s(\xi_k) d\xi_k &= \int_{l_k} \left[\sum_{i=1}^N \int_{l_i} I_i(\xi_i) P(r_{ki}) \cos \varphi_{ki} d\xi_i \right] d\xi_k \quad (9) \\ &= \sum_{i=1}^N \int_{l_k} \int_{l_i} I_i(\xi_i) P(r_{ki}) \cos \varphi_{ki} d\xi_i d\xi_k \end{aligned}$$

よって電界部分は式 (6) と (9) *⁹を合わせることで

$$\int_{l_k} j\omega A_k(\xi_k) d\xi_k = \sum_{n=1}^M \tilde{Z}_{kn} I_{Fn} + \sum_{i=1}^N \int_{l_k} \int_{l_i} I_i(\xi_i) P(r_{ki}) \cos \varphi_{ki} d\xi_i d\xi_k \quad (10)$$

となる

2.2.2 区間導体間での磁気結合解析

大地中の変位電流を勘案した場合、大地表面の電流源の点と大地中の任意の点間での磁界は電流に関する電気映像法*¹⁰より

$$\tilde{Q}(r) = \frac{1}{2\pi(\sigma + j\omega\varepsilon)} \frac{1}{r} \quad (11) \quad \text{表すことができる}$$

式中、 ε ; 大地の誘電率、 r ; 電流源の点と大地中の任意の点間の距離。電力線の電流は電柱に沿って大地に注入した場合、式 (11) に基づいて、 $M+1$ 個の電力線より大地流れる電流 I_{En} が第 k 区間の左右両端で電位を生じ

$$\begin{cases} U_k^P = \sum_{n=1}^{M+1} \tilde{Q}(r_{k'n}) I_{En} \\ U_k^L = \sum_{n=1}^{M+1} \tilde{Q}(r_{k'l_n}) I_{En} \end{cases} \quad (12) \quad \text{となる}$$

式中、 $r_{k'l_n}$ と $r_{k'n}$; 第 N 個の電力線による大地に流れる電流と k 区分の左右両端の距離

式(11)と同様に地中での任意の二点の磁界系統は電気映像法*10より

$$Q(r) = \frac{1}{4\pi(\sigma + j\omega\varepsilon)} \left(\frac{1}{r} + \frac{1}{r'} \right) \quad (13) \text{ となる}$$

式中 r ; 源点と場の点間の距離、 r' ; 源点に相対する地面の鏡像の点と場点の距離。
有限長の埋設導体層の場点による電位は式(13)に対して線積分として得られ

$$\left\{ \begin{array}{l} U_{k'r}^S = \sum_{n=1}^N \int_{l_i} - \frac{d}{d\xi_i} I_i(\xi_i) Q(r_{k'r_n}) d\xi_i \\ U_{k'l}^S = \sum_{n=1}^N \int_{l_i} - \frac{d}{d\xi_i} I_i(\xi_i) Q(r_{k'l_n}) d\xi_i \end{array} \right. \quad (14) \text{ となる}$$

式中の $\tilde{Q}(r_{k'r_n})$ と $\tilde{Q}(r_{k'l_n})$ は第 i 区間での漏れ電流が第 k 区間の左右両端で作り上げる電位の電位係数である。

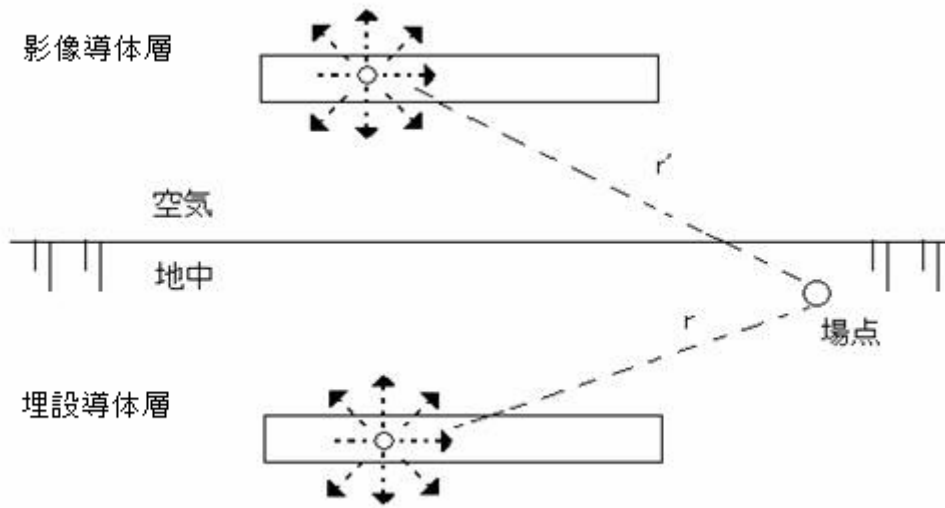


図3 埋設導体層とその映像導体層の図

よって式 (12) と式 (14) *⁹より各区間の磁界部分は、

$$U_{k^r} = \sum_{n=1}^N \int_{l_i} - \frac{d}{d\xi_i} I_i(\xi_i) Q(r_{k^r n}) d\xi_i + \sum_{n=1}^{M+1} \tilde{Q}(r_{k^r n}) I_{En} \quad (15-a)$$

$$U_{k^l} = \sum_{n=1}^N \int_{l_i} - \frac{d}{d\xi_i} I_i(\xi_i) Q(r_{k^l n}) d\xi_i + \sum_{n=1}^{M+1} \tilde{Q}(r_{k^l n}) I_{En} \quad (15-b)$$

となり

$r_{k^r n}$ と $r_{k^l n}$ は第 i 区間での ξ_1 から第 k 区間のそれぞれ左右両端までの距離である

2.2.3 区間導体間での横断面の電流補間関数

最後に導体ケーブル区間での線方向離散化法を用いて、第 k 区間に沿った縦向き電流を計算するには埋設物の金属管-被覆物の計算モデルを導き、このモデルに対する電流は適当な近似関数を用いて表現することが必要となる。埋設ケーブル系統に対して、漏れ電流の影響より、縦向き電流分布は比較的複雑になる。文献*¹¹からわかるように、埋設導体の縦向き電流分布は $e^{\Gamma x}$ 型分布に近似しており、その Γ とは導体電流の伝播定数である。文献*¹¹を用いて数値解析法を適用して、典型的なデータを設定し計算すると、低周波数条件下で $|\Gamma|$ の数値は 10^{-3} となるので $|\Gamma x| \ll 1$ が成り立つ場合、

$$e^{\Gamma x} \approx 1 + \Gamma x \quad (15) \text{ となり}$$

式 (15) が表しているように、導体の長さが十分小さいのなら、埋設物に対しても逐次直線近似関数を用いて電流分布を表したことで十分満足できる計算精度が期待できる。金属管内に覆われた被覆物に対しての電流分布は、逐次直線近似関数を用いて表示することができ、未知数の個数を削除するために、本論文は逐次直線近似を用いてその分布を表す。埋設物の横断面に対しての電流は一次直線補間を用いると

$$I_k(\xi_k) = a_k \xi_k + b_k \quad (16) \text{ となり}$$

式中、 a_k は単位長での側面による電流漏れ係数、 b_k は埋設物の電流の恒値の大きさであるが a_k b_k ともに未知数であるので、式 (16) より埋設物の漏れ電流は

$$-\frac{d}{d\xi_k} I_k(\xi_k) = -a_k \quad (17) \text{ となる}$$

金属管内の被覆物に対しても、電流は逐次定数補間法を用いることが出来るので

$$I_{(a)k}(\xi_k) = b_{(a)k} \quad (18) \text{ となる。}$$

2.2.4 埋設物の電流離散化方程式（遮蔽層電流を考慮しない場合）

式（10）、（15）、（18）を式（4）に代入して、積分の対称性に注意すると

$$\begin{aligned} Z_k I_k b_k + \sum_{i=1}^N \tilde{Z}_{ki} a_i + \sum_{i=1}^N Z_{ki} b_i - \sum_{i=1}^N [Z_{k^r i} - Z_{k^l i}] a_i = \\ - \sum_{n=1}^M \tilde{Z}_{kn} I_{Fn} - \sum_{n=1}^{M+1} [\tilde{Q}(r_{k^r i}) - \tilde{Q}(r_{k^l i})] I_{En} \end{aligned} \quad (19)$$

となり

$$Z_{k^l i} = \int_{l_i} Q(r_{k^l i}) d\xi_i \quad Z_{k^r i} = \int_{l_i} Q(r_{k^r i}) d\xi_i$$

$$Z_{ki} = \int_{l_k} \int_{l_i} P(r_{ki}) \cos \varphi_{ki} d\xi_i d\xi_k \quad \tilde{Z}_{ki} = \int_{l_k} \int_{l_i} \xi_i P(r_{ki}) \cos \varphi_{ki} d\xi_i d\xi_k$$

$Z_{ki}, \bar{Z}_{ki}, Z_{k^l i}, Z_{k^r i}$ は物理量であり

Z_{ki} ; 漏れ電流が流れない場合の第 k,i 区間での相互インピーダンス

\bar{Z}_{ki} ; 第 i 区間で漏れ電流が第 k,i 区間の相互インピーダンスに対しての変化分

$Z_{k^l i}$ と $Z_{k^r i}$ は第 i 区間の全漏れ電流が第 k 区間に対しての左,右端点での磁界係数

$k = 1, 2, \dots, N$ の時に式（19）を利用して N 個の方程式を形成する

2. 3 金属管内の被覆物体の離散化モデル

(遮蔽層電流を考慮しない場合)

金属管とその内側の被覆物は共に円柱形の導体である。有線伝送回路理論*⁵に基づいて方程式が導かれ

$$-\frac{d}{d\xi} [U_{(a)}(\xi) - U_{(p)}(\xi)] = [Z_{(ao)} + Z_{(pi)} - Z_{(ap)} + Z_{(pio)}] I_{(a)}(\xi) - Z_{(pio)} I_{(p)}(\xi) \quad (20)$$

式中の $Z_{(ao)}$ は被覆物の外表面における単位長のインピーダンス

$Z_{(pi)}$ は金属管内表面における単位長のインピーダンス

$Z_{(ap)}$ は被覆物と金属管の間で時間変化する磁束によって表されるインダクタンス

$Z_{(pio)}$ は金属管内の表面と外表面間における相互インピーダンス

$U_{(a)}(\xi)$ と $U_{(p)}(\xi)$ は被覆物体単体、金属管単体が ξ で位置する電位である。

金属管内の被覆物の電流は式 (18) より

$$I_{(a)k}(\xi_k) = b_{(a)k}$$

となり

式中の $b_{(a)k}$ は未知数である。

局部座標の原点は被覆物体の midpoint に定める。第 k 区間で積分すると式 (20) は

$$[U_{(a)k^l} - U_{(p)k^l}] - [U_{(a)k^r} - U_{(p)k^r}] = [Z_{(12)k} + Z_{(11)k}] b_{(a)k} + [Z_{(12)k} b_{(p)k}] \quad (21)$$

となり

仮説 (3) より、被覆物体の両端に抵抗 $R_{(ap)k^l}$ と $R_{(ap)k^r}$ を設置し、式 (25) より管道-

被覆物体の等価電気回路モデルが得ることが出来、これを図4に示す。接続点による連続条件より被覆管、被覆物体の間の横方向電圧は

$$U_{(a)k^l} - U_{(p)k^l} = R_{(ap)k^l} [b_{(a)k-1} - b_{(a)k}] \quad (22)$$

$$U_{(a)k^r} - U_{(p)k^r} = R_{(ap)k^r} [b_{(a)k} - b_{(a)k+1}] \quad (23)$$

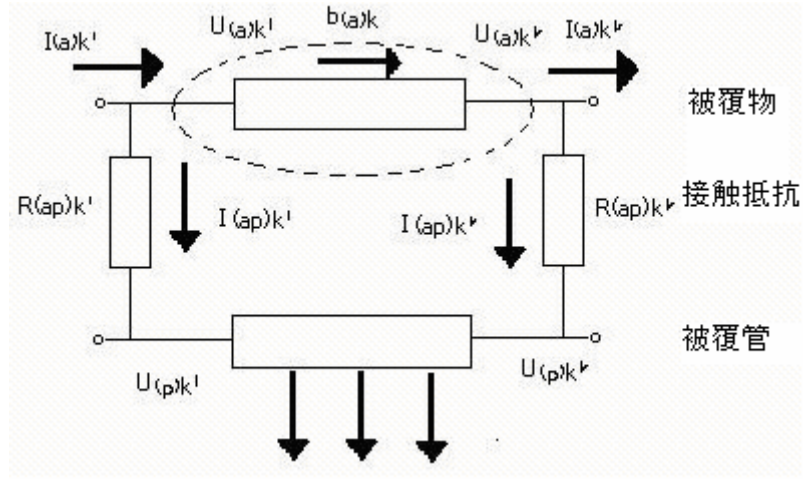


図4 被覆管-被覆物単体の回路図

式(22)(23)を式(21)に代入し整理すると

$$Z_{(12)k} b_{(p)k} - R_{(ap)k^l} b_{(a)k-1} + R_{(ap)k} b_{(a)k} - R_{(ap)k^r} b_{(a)k+1} = 0 \quad (24)$$

式の中で $R_{(ap)k} = R_{(ap)k^l} + R_{(ap)k^r} + Z_{(11)k} + Z_{(12)k}$ である

管内部に被覆物 $N_{(a)}$ 個設置した場合、式(26)を用いて $N_{(a)}$ 個の方程式を作れる

2.4 ケーブル系統の電流分布の解を求める

2.2節と2.3節では遮蔽層の電流作用は考慮していない、仮説(2)より遮蔽層内の電流 I_{pb} 、遮蔽層と被覆物、金属管の第 k 区間の相互インピーダンスを $Z_{k, pb}$ とし、遮蔽層での電流が金属管、被覆物より引き起こす起電力を考慮したあと、式(19)と(24)は修正され

$$Z_k I_k b_k + \sum_{i=1}^N \tilde{Z}_{ki} a_i + \sum_{i=1}^N Z_{ki} b_i - \sum_{i=1}^N [Z_{k^r i} - Z_{k^l i}] a_i + Z_{k, pb} I_{pb} = - \sum_{n=1}^M \tilde{Z}_{kn} I_{Fn} - \sum_{n=1}^{M+1} [\tilde{Q}(r_{k^r i}) - \tilde{Q}(r_{k^l i})] I_{En} \quad (25)$$

$$Z_{(12)k} b_{(p)k} - R_{(ap)k^l} b_{(a)k-1} + R_{(ap)k} b_{(a)k} - R_{(ap)k^r} b_{(a)k+1} + Z_{k, pb} I_{pb} = 0 \quad (26)$$

よって、未知数に対して総括的に述べると、金属管、被覆物を共にN個とし、未知数を a_k 、 b_k となり、従って $2N$ となる。金属管内の被覆物体を $N_{(a)}$ 個で未知数は $b_{(a)k}$ 個、従って $N_{(a)}$ 個であるので、遮蔽層中での電流 I_{pb} をもう1度考慮した場合、未知数が $2N + N_{(a)} + 1$ 個ある。

前節の埋設導体ケーブルと金属管内の被覆物体単元の離散化電気回路モデルが図5のように描かれる。1'、2'、6'、7'、9'、10'、11'と11'は物体の中心である。

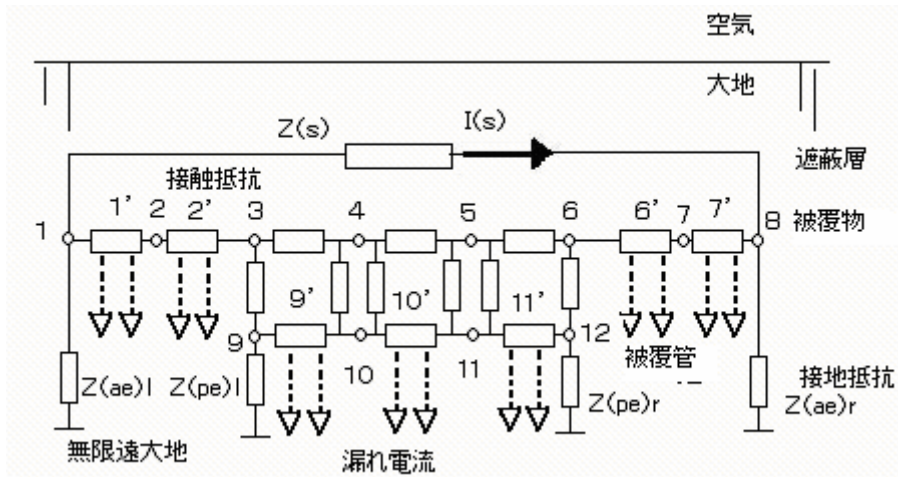


図5 地中導体系統の等価である電気通信ネットワーク図

電気通信ネットワークの解を求める際、難しいのは

- ① 電力線の電流によって、すべての区間に等しく印加される起電力の作用
- ② ある区間での相互誘導作用
- ③ 漏れ電流によって、埋設物を漏れ電流が分散することでその他の物体の両端の電位があがること

電気回路より述べると、埋設物と無限遠長である大地の間は抵抗性ネットワークが構成される。よって漏れ電流が物体に沿って一様に分布することによって、抵抗性ネットワークの相互抵抗は集中定数法が適用しがたい。図5の等価電気通信ネットワークモデルを用いて電流分布を求め、解析的結果より以下のように述べる事ができる

措置1：回路の電圧方程式の形成

- (1) 数値計算中の**主元素**の優勢を保障するため^{*12}、近接する埋設物の midpoint で構成する新たな区間を考慮し、それと無限遠点長の大地（接点0で表す）で構成された回路（図5中の回路1'-2'-0, 6'-7'-0, 9'-10'-0, 10'-11'-0）に関しては式(31)を用いて回路電圧方程式を描く。従ってN-3個の方程式である
- (2) 遮蔽層-大地間の回路（図5 0-1-8-0）、従って1つの方程式である
- (3) 被覆物-金属管によって構成される回路（図5 3-4-10-9, 4-5-11-10, 5-6-12-11）は式(31)を用いて回路電圧方程式を描く、従ってN_(a)個の方程式である

この結果よりN+N_(a)-2個の方程式が形成される

措置2：接点電流方程式の導出。

措置1よりN+N_(a)-2個の方程式が導かれたので、未知数は2N+N_(a)+1個であるため電流連続条件である接点数を勘案すると方程式の数が足りない

- (1) 埋設単位区間の第k、k+1区間（図5 2、3、6、7、10、11）から式(16)を用いてN-1個の方程式を導出する。

$$a_k \frac{1}{2} l_k + b_k = -a_{k+1} \frac{1}{2} l_{k+1} + b_{k+1} \quad (27)$$

従ってN-1個の方程式

- (2) 境界の接続点の電流処理（図5 1, 8, 9, 12）

金属管、被覆物の左右端の電圧分布は下記の式より

$$U_k^r = \sum_{i=1}^N Z_{ki}^{(r)} (-a_i) + \sum_{n=1}^{M+1} \tilde{Q}(r_{k'n}) I_{EN} \quad (28)$$

$$U_k^l = \sum_{i=1}^N Z_{ki}^{(l)} (-a_i) + \sum_{n=1}^{M+1} \tilde{Q}(r_{k'n}) I_{EN} \quad (29)$$

式(28)(29)と境界での接地抵抗によって境界接点で、地中に流れる電流（例として図5での接点1の地中に流れる電流はU₁^l/Z_(a)^l）接続点での電流連続性方程式は従って4個

である

措置1、措置2で形成した方程式の数は $N + N_{(a)} - 2 + N - 1 + 4 = 2N + N_{(a)} + 1$ であり、方程式の数は未知数の数に等しく、同時に措置1の中で積分方程式が近接する物体の中点を考慮したことによって、主元素優勢の連立方程式が保証され、方程式が解ける

金属管と被覆物の間には実際の計算で注意しなければならない問題である。計算より表されるように、分割個数の増加は計算精度を向上し、精度を高めるのを助ける。しかし単元を細分化することで無駄に計算に時間が膨大となる、またある状況下では誤差が蓄積することで計算精度が下落することを招くことになる。

2. 5 通信ケーブルの電力線における電波妨害

実際の工程ではケーブル遮蔽係数とケーブル軸の縦向き誘導起電力のため計算の目安が必要となる。遮蔽係数の定義は

$$k_0 = \frac{E_A}{E_{A0}} \quad (30)$$

式中、 E_A : 遮蔽系統がある場合、妨害を受ける導線Aでの
単位長の縦向き誘導起電力

E_{A0} : 遮蔽系統がない場合、妨害を受ける導線Aでの
単位長の縦向き誘導起電力

遮蔽線Pが妨害を受ける導線Aでの誘導起電力 E_{AP} の計算例はケーブル系統の遮蔽線Pの電流分布を求めたあと、各物体の誘導電流電磁界が妨害を受ける導線Aでの作用を累計することで、積分の対称性に注意し、

$$E_{AP} = \sum_{i=1}^N \int_{l_i} (a_i \xi + b_i) Z_{(AP)i} d\xi = \sum_{i=1}^N b_i Z_{(AP)i} \quad (31) \text{ とする}$$

$Z_{(AP)}$: 遮蔽線Pと妨害を受ける導線Aの間の相互インピーダンス

同様な計算で得られる E_{A0} 、 E_A 、よって遮蔽係数 k_0 が得られる。

2. 6 計算例

電力系統と電力ケーブルの配置を図6に示す。大地の表面に直角の座標系を打ち立て、その平面に俯瞰（鳥瞰）図6を示す。500kV電力線に高度19.5m仮設し、1段の縦向きの電流のみで、単相接地を事故を仮定し、線路に縦向きの電流を(1000、-1000)Aとし、大地に流れる電流を(1000、-1000)A、(-1000、1000)と分ける。ケーブル回線をHEQ₂型通信ケーブルとし、大地の表面に平行に埋設する。ケーブルの埋設1m、その他のパラメータは以下のとおりである。

被覆物の外側直径：65.8mm 厚さ：0.5×2mm

遮蔽層内側直径：41.6mm 厚さ：2.05mm

鋼管外側直径：80.0mm 厚さ：6mm

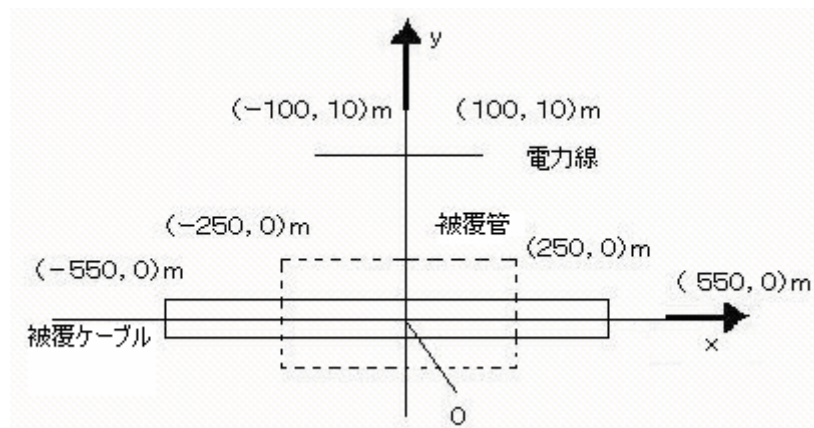


図6 線路の鳥瞰図

ケーブル、鋼管の両端の接地抵抗は0（理想設置）とする。

金属管とケーブル被覆物間の接地抵抗も0とし、金属管と被覆物は1系統点を通る良好な接触とし、ケーブルは11区間、金属管は5区間に分けることができる。

ケーブル系統の理想遮蔽係数とケーブル軸上の誘導起電力を計算する。

状況1 ケーブルが鋼管に保護されていない状態での計算表1にまとめた。

表1 ケーブルが鋼管に保護されていない状態

項目	複素数 / V	絶対値 / V
電力線で生じる縦向き起電力	-56.537 - j 44.469	71.930
被覆物で生じる縦向き起電力	17.896 + j 6.076	18.899
遮蔽層で生じる縦向き起電力	-2.784 + j 31.581	31.704
すべての縦向き起電力	-41.553 - j 6.971	42.134

表1に基づいてケーブルが鋼管に保護されていない状態で、遮蔽係数 $k_1 = 0.514 - j 0.281$ 、より $|k_1| = 0.586$ 。設計手帳よりHEQ₂型通信ケーブルの理想遮蔽係数は0.3~0.6となり、この方法を用いて合理的に計算をする。

状況2 ケーブルが鋼管で保護されている場合の計算表3にまとめた。

表2 ケーブルが鋼管に保護されている状態

項目	複素数 / V	絶対値 / V
電力線で生じる縦向き起電力	-56.537 - j 44.469	71.930
被覆物で生じる縦向き起電力	6.572 + j 1.480	6.737
遮蔽層で生じる縦向き起電力	7.109 + j 17.573	18.956
鋼管で生じる縦向き起電力	18.834 + j 35.968	40.601
すべての縦向き起電力	-24.021 - j 7.592	25.193

表2に基づいてケーブルが鋼管に保護されている状態で、遮蔽係数 $k_2 = 0.197 - j 0.289$ 、より $|k_2| = 0.350$ 。より、鋼管を設けたほうがケーブルを保護するのに有利である

本章は地中の遮蔽系統の電流分布に通用する計算モデルを提唱した。実際にこのモデルは通信回線、都市の金属導線の設計などへ実用化が可能である。ここでのモデルの使用制限条件は大地を半無限大の導電媒質に仮定した部分である

3 地中の通信ケーブルを金属管で覆う場合と付近に遮蔽線を施設した場合の効果の比較

本章は送電線と通信回線系統で、被覆していない通信ケーブルを金属管で覆った遮蔽保護とケーブルに沿って平行に設けた遮蔽線保護の計算を行い、地下のケーブルの電磁遮蔽を解析する。叙述を簡潔にするため、金属管と遮蔽線を遮蔽体と総称する。

3.1 計算例と遮蔽効果の解析

図7は500kVの送電系統と地中の通信ケーブルの俯瞰図である。送電線甲と線路乙は変電所の高圧の母線から出たもので、遠くは変電所に通じる。線路甲は図に記された短絡点で故障した場合に相分量法をもちいて線路甲と双回線乙での短絡電流、架空地線での逆流電流 (return current)、電柱から大地に流れ込む電流を計算し^{*12*13}、その結果を通信ケーブルの電磁妨害源とする。遮蔽方式はケーブルを金属管で覆うかケーブル付近にケーブルと同様の深さで平行に遮蔽線の設備する方法があり、施設範囲は図の点1～5である。遮蔽体とケーブル鉛袋の両端は接地抵抗を通してアースに集中する (図7表示)。この系統は50Hzの場合の電界と磁界及び遮蔽規則の解析である。また通信ケーブルを深さ1mのところに埋設する。

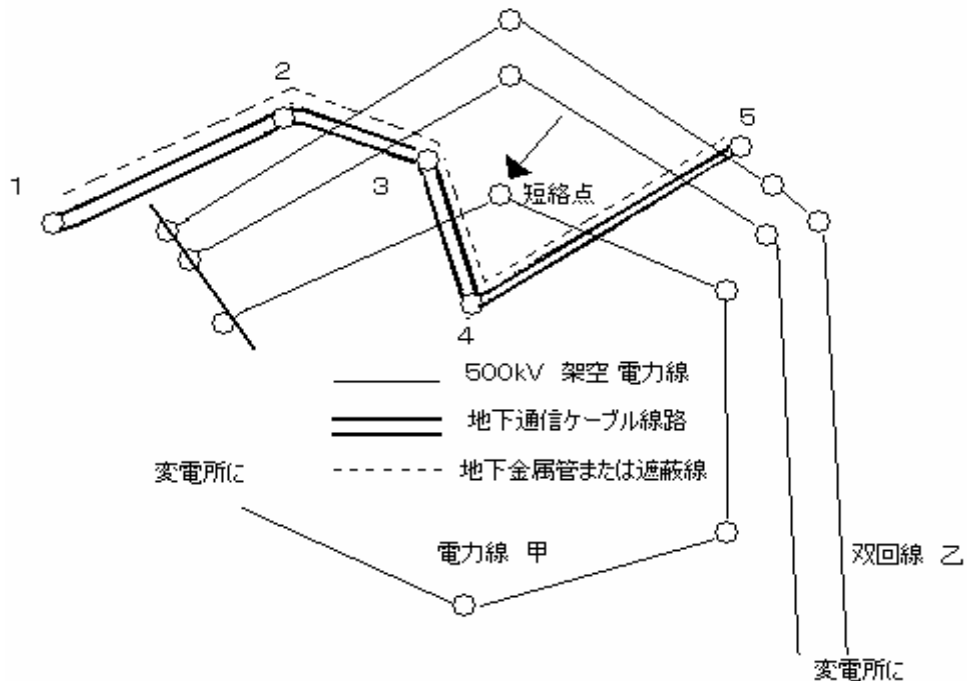


図7 電力線と通信回線の図

3. 2 両端の接地抵抗が遮蔽効果を及ぼす影響

金属管を覆い保護することは接地抵抗の遮蔽効果を及ぼす影響を説明するということになる。

金属管の外直径／厚さ：0.08m／0.02m

鉛袋の外直径／厚さ：0.0296m／0.0016m

大地の導電率：0.02 S/m

表3では異なる接地抵抗でケーブルの総合遮蔽係数を表している。表3からわかるように、接地抵抗の増大によって、遮蔽係数も同様に増大し遮蔽効果を低下する。物理概念より述べると、両端の接地抵抗の増大は遮蔽体と鉛袋中の縦向きの誘導電流を削減する。このため総合的に磁気作用は弱くなる

表3 接地抵抗が遮蔽システムを及ぼす影響

接地抵抗 ()	金属管単独	鉛袋単独	金属管 + 鉛袋
0.1	0.5446	0.8998	0.4916
0.5	0.5626	0.9159	0.5096
1.0	0.5697	0.9309	0.5222
5.0	0.5782	0.9753	0.5547

3. 3 金属管の外直径及び厚さが遮蔽効果を及ぼす影響

金属管の厚さが一定で、直径が増大するとその自己インピーダンスは減少し、また誘導電流が増大することで、ケーブルの遮蔽保護に有利である；もう一方で、管径が増大したあと、ケーブル軸との間の距離が大きくなり、両者の相互インピーダンスは減少し、金属管はケーブルに及ぼす遮蔽は削減する。よって、管径の大きさの遮蔽効果に対する影響は、計算を通して結論を出す必要がある。表4は金属管の厚さ0.002mのとき、異なる管径での遮蔽係数である。表4よりわかるように、一定の範囲内では、管径が大きいほうがケーブル遮蔽に有利である。

表4 金属管外直径の遮蔽係数に対する影響

金属管の外直径 / m	0.080	0.100	0.110	0.125
遮蔽係数	0.5098	0.4704	0.4541	0.4328

3. 4 遮蔽線の材料や付設した位置が遮蔽効果を及ぼす影響

遮蔽線の通信ケーブル保護は遮蔽線の材料、直径、遮蔽線の施設する位置などによって考慮しなければならない。簡単に金属管の遮蔽効果と比較するために、遮蔽線の断面積と金属管の体表面積を同一にして、外直径を0.08m、厚さ0.002mの金属管に対応して、0.02498mの遮蔽線の直径を選ぶ。遮蔽線と鉛袋の接地抵抗は0.5Ωとし、表5と表6に材料と遮蔽線の距離がケーブル軸の距離と異なる場合ケーブルの総合遮蔽係数をここに示す。

表5 遮蔽線材料、直径と配置の場所が遮蔽係数を及ぼす影響
(遮蔽線の直径0.02498m)

距離 / m	銅線	アルミニウム線	鋼線
0.04	0.2100	0.2100	0.5282
0.10	0.2458	0.2518	0.5360
0.50	0.3202	0.3232	0.5490
1.00	0.3497	0.3520	0.5546

表6 遮蔽線材料、直径と配置の場所が遮蔽係数を及ぼす影響
(直径が0.01249m)

距離 / m	銅線	アルミニウム線	鋼線
0.04	0.2675	0.3006	0.5705
0.10	0.2998	0.3268	0.5746
0.50	0.3565	0.3748	0.5816
1.00	0.3803	0.3956	0.5846

表より遮蔽線の材料は遮蔽効果の影響が大きいことがわかり、遮蔽線を敷く位置は遮蔽効果に対する影響が小さいことがわかる。表3と表5の直径0.04mの時の結果をよく見ると同じ材料に対して遮蔽線の距離とケーブルの水平離が金属管の外径と等しい場合、金属管の遮蔽の効果は遮蔽線よりよいが、差は大して変わらない。表5と表6からわかるように遮蔽線が太いほど遮蔽効果がいい。

- (1) 電力線が近接する地中の通信ケーブルに対して危険な影響を与える計算を行うことは**電界**と**磁界**を同時に考慮しなければならない
- (2) 地下通信ケーブル遮蔽保護方式は2種類ある
鋼管を覆うまたは遮蔽線を敷くかの選択は遮蔽効果、防雷、摩擦腐食の防護などの要素を考慮する。
 - ・ **電界**の遮蔽効果より考えるなら、遮蔽線と外側を覆う金属管の効果が拮抗し、遮蔽線の設置がとても簡便であるといえる。
 - ・ **磁界**の影響が防雷影響の遮蔽効果に対して考えるなら、金属管で覆うほうが遮蔽線を設備するより有利であるといえる

4 考察

筆者は本文を翻訳する段階で、専門文章以外にもさまざまな翻訳を試してみた。今回感じたことは一般の翻訳と技術論文の翻訳ではもっとも必要となるものも違うということだ。翻訳の必要な要素は、外国語の読解力、母国語の表現力、専門分野の知識の三つが挙げられるだろう。しかし専門分野と一般文章を翻訳するのではこの要素の重大性が大きく分かれる。一般的に専門文章というのは、その中身を客観的に、正確に周りの人に伝えられるかということが大事である。よって、もっとも大事とされる要素は専門分野の知識である。

筆者も翻訳を始めたとき中国語が出来れば翻訳は簡単と安易に考えていたが、何度も壁にぶつかりそのたびに苦悩し続けた。その背後にこの分野の知識の少なさが大きな原因にあがった。また始めの概要だけを集中して読み内容を理解しようと思ったがこれも上手くいかなかった。実際に一回わからなくても通して読んでみるのが早く読む鍵になるだろうと感じた。

反省する点はこの計算モデルをもとに実験データを取るべきだが、翻訳に時間がかかり実行に移すまでにはいたらなかった。また作者が参考資料に挙げた論文が入手困難であるためその単語の本質的な意味を掴めなかったことも反省のひとつである。

今回参考に用いた邹军 (zou jun) 講師による「**电力线路对通信线路电磁感应危险影响的研究**」の論文だけを挙げると、翻訳に難解な点が4つ残った。まず1つは次から次へと新しい変数が現れる。2つめに原作者の指導教官である教授が考えたと思われる単語を頻繁に用いているところであり、この意味の単語は同様に他の単語もあるとみられ、そちらのほうが頻繁に用いられているのに敢えて使っていたりする。3つ目は肝心な部分が文献参考と書いてあり詳しい内容が書いてないということだ。4つ目は計算よりモデルを表しているが実験による実測値を取ったデータが見当たらないところだ。これらより内容を完全に把握するのにとても難しく感じた。

5 参考資料

* 1 電界

原文では感性結合（耦合）と書かれていたが一般的に用いられている電界と訳した

* 2 磁界

原文では磁界（耦合）と書かれていたが一般的に用いられている磁界と訳した

これらは高攸綱《感性耦合与阻性耦合》北京 人民邮电出版社 1979 を参考に用いている

* 3 遮蔽線（シールド線） Shielding wire 屏蔽线 (ping2 bi4 xian4)

送電線などの誘導雑音源が通信線に影響を与えないように、これらの中に遮蔽線や導体網を置き、これを接地することによって通信線への誘導電圧障害を減少させることを遮蔽効果（屏蔽效应）という

* 4 規定法

- ・ 中華人民共和国 電力工業部

送电线路对电信线路危险影响设计规程

[電力線が電気通信回線に危険な影響を与える設計規定] (簡単に設計規定と呼ぶ)

1994年版

- ・ 能源部东北电力设计院，电力工程高压送电线路设计手册，北京；水利电力出版社

[エネルギー部東北電力設計院、電力工程高压電力線設計ハンドブック] 北京

水利電力出版社、1991

* 5 伝送線路法

- ・ Frazier M, Dunlap J. Power line induced AC potential on natural gas pipelines for complex rights-of-way configurations. EPRIEL-3106, 1983
- ・ Haubrich H J, Flechner B A, Machczynski W. A university model for the computation of the electromagnetic interference on earth return circuits IEEE Trans. On power delivery, 1994, .9(3): 1593~1599
- ・ Taflove A, Dabkowski J. Prediction method for buried pipeline voltages due to 60 Hz AC inductive coupling, Part1-Analysis [J]. IEEE Trans on Power Apparatus and System, 1979, PAS-98(3):788~794.

- * 6 電磁界の数値計算技術
 - Dawalibi F Electromagnetic fields generated by overhead and buried conductors, Part1-single conductor. IEEE Trans. On power delivery, 1986,PWRD-1(4):105~111
 - Dawalibi F Electromagnetic fields generated by overhead and buried conductors, Part2-ground network. IEEE Trans. On power delivery, 1986,PWRD-1(4):112~119

- * 7 Schelkunoff S A. The electromagnetic theory of coaxial transmission lines and cylindrical shields [J] . Bell System Technical Journal , 1934, 13: 532 ~ 579.

- * 8 张秋光 场论 (下册) 北京 地质出版社 1988
[場の理論 (下)]

- * 9 Sunde E D. Earth conduction effects in transmission systems [M] . D. Van Nostrand Company, New York, 1949.

- * 10 Takashima T. Nakae T. Ishibashi R. Calculation of complex fields in conducting media . New York: John Wiley&Sons. 1997

- * 11 Lei Y.Z. Ma X.S. Propagation constant, impedance and admittance calculation of horizontal underground conductors. Proc .of Asia-pacific conference on environmental electromagnetics 1996, 233-236

- * 12 周宇坤 电力线路对金属管道内通讯电缆的影响[毕业设计论文] 北京
清华大学电机系 1998
[電力線が金属管内の通信ケーブルに対する影響]

- * 13 邹军, 袁建生 陆家榆 马信山 架空线路短路电流分布及屏蔽系数的计算
电网技术 2000 , 24 (10); 27-30
[架空電線短絡電流分布及び遮蔽指数の計算]

本章に用いられている記号の紹介を表7に示す

表7 主に使われる記号表

名称	英文	略字	符号位置
大地	Earth	e	下
遮蔽層	Shield	s	下
被覆物	Armor	a	下
金属管	Pipe	p	下
一次場	Primary field	p	上
二次場	Second field	s	上
左端	Left	l	上
右端	Right	r	上
内側	Inner	i	下
外側	Outer	o	下

表8 電磁気変数表

変数名(単位)	日本語	English	中文
Q (C)	点電荷	Charge	电荷 (库)
I (A)	電流	Current	电流 (按)
E (N/C)	電界	Electric field	电场 (牛每库)
D (C/m ²)	電束密度	Electric flux density	电位移 面电荷密度 电通量密度 (库每平方米)
B (T, Wb/m ²)	磁束密度	Magnetic flux density magnetic induction	磁通(量)密度 磁感应强度 (忒斯拉)
H (A/m, N/Wb)	磁界	magnetic field	磁场 (按每米)
Φ _e (C) e	電束	Electric flux	电通量 (库)
Φ _m (Wb)	磁束	Magnetic flux	磁通率 (韦伯)
P (C/m ²)	分極ベクトル	Polarization	极化 (库每平方米)

M (A/m, N/Wb)	磁化ベクトル	Magnetization	磁化 (按毎米)
p (C・m)	電気双極子モーメント	Electric dipole moment	电矩 (库・米)
m (A・m ²)	磁気双極子モーメント	magnetic dipole moment	磁矩 (按・米)
(C/m ³)	電荷密度	Charge density	电荷密度 (库每立方米)
J (A/m ²)	電流密度ベクトル	current density	电流密度 (按每平方米)
χ (C/m ²)	面電荷密度	surface charge density	面电流密度 (库每平方米)
K(A/m)	面電流ベクトル	surface current density	面电流密度 (按毎米)
(C/m)	線電荷密度	Line charge density	(库毎米)
ε ₀ (C ² /N.m ²) (F/m)	真空中の誘電率	dielectric constant in a vacuum	真空介电常数 (法毎米)
μ ₀ (H/m)	真空中の透磁率	magnetic permeability in a vacuum	真空导磁率 (亨利毎米)
ε (C ² /N.m ²) (F/m)	物質中の誘電率	dielectric constant	介电常数 电容率 (法毎米)
μ (H/m)	物質中の透磁率	Magnetic permeability	导磁率 (亨利/米)
ε _r	比誘電率	Relative dielectric constant	相对介电常数
μ _r	比透磁率	relative magnetic permeability	相对导磁率
H (Ω)	固有インピーダンス	intrinsic impedance	固有阻抗 (欧)
χ _e	分極率	polarizability	极化率 极化性
χ _m	磁化率	Magnetic susceptibility	磁化率 磁化性
φ (V)	スカラーポテンシャル	Scalar potential	标势 (伏)
A (A)	ベクトルポテンシャル	Vector potential	矢量势 (按)

V (V)	電圧	Voltage	电压 (伏)
R(Ω)	電気抵抗	registance	阻抗 (欧)
G(S, Ω^{-1})	コンダクタンス	conductance	电导 导率 电导系数 (西门)
σ (S/m)	電気伝導率	conductivity	传导率 传导性 (西门每米)
C(F)	静電容量	capacitance	电容 (法)
L(H)	インダクタンス	inductance	感应系数 自感应 (亨利)
V_m (A · turn)	起磁力	magnetomotive force	磁动势 磁通量
R_m (A/Wb)	磁気抵抗	reluctance	磁阻 (按/韦伯)
S(J/m ²)	ポインティングベクトル	Poynting vector	波印廷矢量
W_e (J)	電氣的エネルギー	electric energy	电能 (焦耳)
W_m (J)	磁氣的エネルギー	magnetic energy	磁能 (焦耳)
P_d (W)	ジュール熱となる電力	Joule heat	焦耳热 (瓦特)
w_e (J/m ³)	電氣的エネルギー密度	electric energy density	电能密度 (焦耳每立方米)
w_m (J/m ³)	磁氣的エネルギー密度	magnetic energy density	磁能密度 (焦耳每立方米)
p_d (W/m ³)	電力密度	power density	电力密度 (瓦特每立方米)

6 参考文献

- | | |
|---------------------|--------------------|
| 斎藤兆古 | 「応用電磁気学入門」 |
| 電気学会 | 「電気磁気学」 |
| オーム社 | 「電子学常用辞書」 |
| 東京堂出版 | 「類語辞典」 |
| 清華大学出版社 | 「電磁学」 |
| 高等教育出版社 | 「日中英 エレクトロニクス用語辞典」 |
| 商務印書出版
(商务印书馆出版) | 「精選日漢・漢日辞典」 |

2005 年度 卒業論文

磁気歯車の与える周辺磁界分布に 関する考察

法政大学 工学部

情報電気電子工学科 4年

学籍番号 02D3059番

須永 高志

指導教員 斎藤 兆古

目次

	ページ数
1. 序章	3
2. 永久磁石を用いた非接触型動力伝達装置	4
2 - 1. 同期速度における動力伝達	4
2 1 1. 磁気歯車の動作原理と特徴	4
2 1 2. 測定実験	5
(A)実験装置	5
(B)実験方法	8
(C)実験結果	9
(D)考察	14
3. 磁気歯車を与える磁界分布	15
3 1. 測定原理	16
3 1 1. 磁界逆問題のシステム方程式導出	16
3 2. 測定実験	19
(A)実験装置	19
(B)実験方法	19
(C)実験結果	20
(D)考察	22
4. まとめ	23
5. 参考文献	23

1. 序章

現在、車をはじめガソリンや軽油を燃料とした内燃機関を動力源とした乗り物は広く世の中に普及している。これらは、速度の変速を行うクラッチの部分には油を用いたトルク・コンバーターが利用されている。このトルク・コンバーターの油は、車が廃車になり、廃油として処分されると、排気ガスと並んで環境破壊の原因になると言われている。現在では自動車リサイクル法により、産業廃棄物の埋め立ては禁止となり、専門業者に委託することで廃油は回収・処理されているため、昔よりも環境について考えられている。しかし、実際廃油が減ったわけではなく、環境を従来より考え、うまく処理されているだけであって根本的な解決にはならない。

ここで挙げた問題の対策として考えられるのは、油を用いずに動力を伝達させ、なおかつ出力の制御ができる装置の開発が必要であると考えられる。従来の接触型の歯車では滑らかな動作および接触部での摩擦損失や音・熱による損失の軽減を考えたら、どうしても油を用いざるを得ない。そこで接触をさせず、磁石の力によると、動力を伝達させる磁気歯車を用いた非接触型動力伝達装置利用すればこの問題が解決できる。磁気歯車を用いた非接触型動力伝達装置の特徴は後に述べるが、主な特徴として、モーターに過負荷がかかった場合には磁気歯車が滑り、動力源を保護する。そして、機械的な摩擦が生じない非接触機構なので効率が良いという2つの特徴が挙げられる。

しかし、実際この磁気歯車が製品化され、一般家庭に入ってきたとすると、磁石を用いるということで、それに伴い、磁界による他の製品に対する影響が少なからず生じると考えられる。

本論文では、この磁気歯車の周囲の磁界分布を測定し、その分布を考察する。

2. 永久磁石を用いた非接触型動力伝達

2 1. 同期速度における動力伝達

2 1 1. 磁気歯車の動作原理と特徴

磁石は、同じ極だと反発力、異なる極では吸引力が生じる。この磁石の吸引反発を利用し、非接触による動力の伝達するのが磁気歯車である。

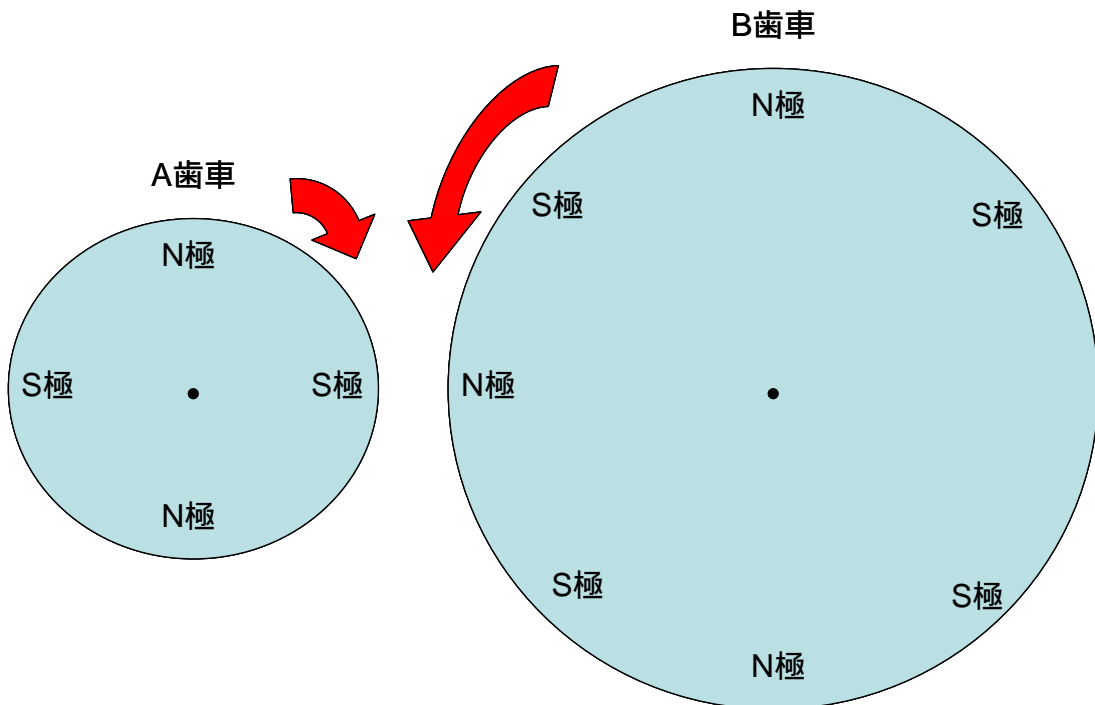


図1 磁気歯車の動作原理

上図のように、円盤状に磁石をN極とS極を交互に設置し、A歯車・B歯車ともに中心を軸にし、接触しないで回転できるように固定する。この時、固定されていないと磁石同士が吸引し、歯車が接触してしまう。しっかり固定した後、A歯車を回転させていくと、磁石の吸引・反発力でB歯車は回転し始める。これが磁気歯車の動作原理である。

磁気歯車には次の 4 つの特徴がある。

メンテナンスフリー

非接触磁気クラッチを使用し、歯車の磨耗・隙間管理の問題が無い。

穏やかな変速動作

歯車の磁気的なバネ効果により、変速時のショック・騒音が小さい。

過負荷保護

過負荷時には磁気歯車が滑り、動力源を保護する。

高効率

機械的な摩擦が生じない非接触機構なので効率が良い。

今回の実験で実際に確認できた特長は の過負荷保護である。モーターの回転を徐々に上げていく分には、滑らないが、急速に高速回転させ、動力を伝達しようと試みても、歯車はすべりを起こし、動力を伝達することは出来ない。

2 1 2 . 測定実験

(A) 実験装置

前述の動作原理を応用して、試作した磁気歯車の両歯車および実験装置を図 2 1、図 2 1 に示す。

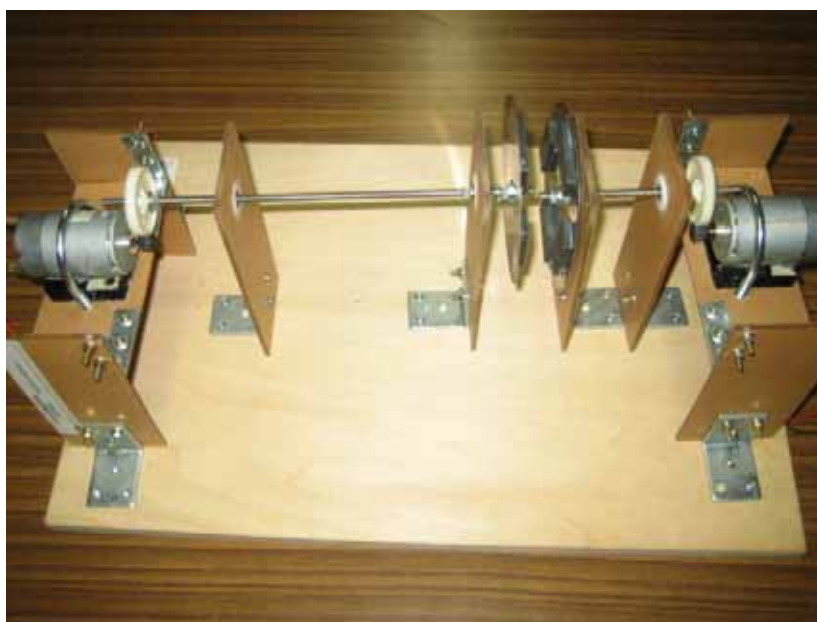


図 2 1 磁気歯車の実験装置



図 2 2 磁気歯車

作成した装置は、磁石を N 極、S 極と交互に配置した同じ 2 枚の歯車を非接触に保つ距離で向かい合わせたもので、動作原理で説明したものととは構造は異なるが、原理的には同じである。

磁気歯車の実験を始める前に、今回の実験で使用しているモーターの特性を知る必要がある。モーターを図 2 3 のように向かい合わせ直結に接続し、その入力電力、出力電力、歯車の回転数などそれぞれのパラメータからモーターを含む装置の特性を測定する。行っ

た実験は外部特性と伝達効率の測定です。



図 2 3 歯車を用いたモーター直結

モーターの外部特性を測定する方法はなにも直結による測定だけではない。例えば、ドライブ・ベルトによる測定である。この測定方法は、図 2 4 のようにモーターを平行に設置し、モーターの歯車をドライブ・ベルトで連結させることによって、動力を伝達する。

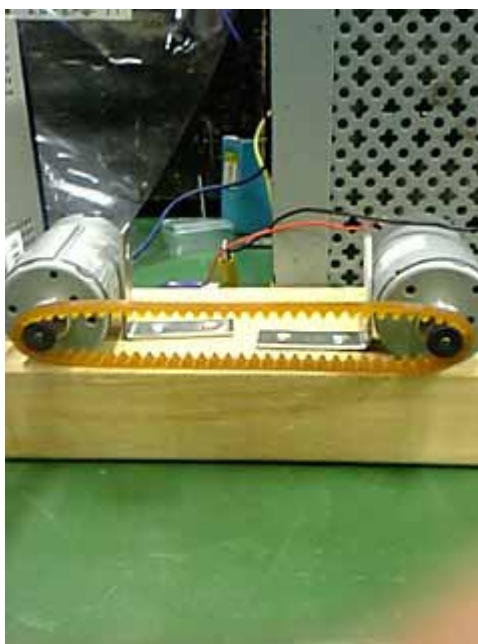


図 2 4 モーターベルトによる連結

しかし、この測定法はモーターをベルトで押さえつけながら動力を伝達するので、ベルトと歯車間の損失が否めない。実際同じ条件で同様の実験を試みたところ、直結による連結の場合より効率が2~3割も低いことがわかる。したがって、モーターの外部特性の測定は、ドライブ・ベルトよりも直結による連結で行った。

(B)実験方法

実験1 外部特性の測定

- 1 - 1 . 駆動側モーターへ印加する電圧を一定として、出力側が無負荷時の場合の歯車の回転数をカウンタを用いて測定する。出力側が無負荷というのは出力側モーターの端子を開放した場合である。
- 1 - 2 . 駆動側モーターへ印加する電圧を一定として、出力側の電機子端子へ抵抗を連続的に変化させた場合の歯車の回転数をカウンタを用いて測定する。この場合、出力側の負荷抵抗を連続的に変化させることは、出力側端子に界磁抵抗器を接続して界磁電流を変化させることである。すなわち、この外部特性の測定は出力側モーターの端子に接続した界磁抵抗器によって出力側モーターの電機子電流を変化させることで装置全体の回転数を制御し、その時の装置の特性を測定する。

実験2 伝達効率の測定

出力側モーターの電機子に界磁抵抗器を接続して、その抵抗値を変化させることによって出力側モーターの電機子電流により装置全体の回転数を制御し、その時の入力側モーターと出力側モーターのトルクの比率より、装置の効率を測定する。

実験3 入出力特性の測定

出力側モーター端子に 1000[]の抵抗を接続し、その抵抗に消費される電力を求める。このとき、入力側のモーターを 1~16[V]と 1[V]ごとに電圧を上げていきながら測定する。

今回の測定に用いた計算式を次に示す。

回転数 n [rad/sec] を機械的角速度 ω_m :

$$\omega_m = \left(\frac{2\pi}{60} \right) \times n \text{ [rad/sec]} \quad \dots 2$$

電力 W [W] は

$$W = V \cdot I - R \cdot I^2 \text{ [W]} \quad \dots 2$$

ただし、 V : モーターの端子間の電圧、 I : モーターの端子間の電流、 R : 入力側もしくは出力側のモーターの内部抵抗。

トルク T [N · m] は

$$T = \left(\frac{W}{\omega_m} \right) \text{ [N · m]} \quad \dots 2$$

効率 η [%] は

$$\eta = \left(\frac{W_o}{W_i} \right) \times 100 = \left(\frac{T_o}{T_i} \right) \times 100 \text{ [%]} \quad \dots 2$$

ただし、 W_o : 出力電力、 W_i : 入力電力、 T_o : 出力トルク、 T_i : 入力トルク。

(C) 実験結果

実験 1 外部特性の測定

(B) 実験方法の節で述べたように、機械的角速度と入力側モーターのトルクを測定し、縦軸に機械的角速度、横軸に入力側モーターのトルクをとってグラフ化したもの（外部特性曲線）を示す。入力電圧は連結装置、磁気歯車ともに 16V で一定とする。

この外部特性曲線を描く事により、曲線上の点と X 軸、Y 軸で囲まれた面積が伝達される電力であることが式よりわかる。

なお、再現性を得るために同じ測定を 3 回ずつ繰り返し行った。磁気歯車の外部特性は、3 回の測定において再現性は見られるがデータにかなりのばらつきがあるのは、正確に目盛りが読み取れなかったことや、計測器自体の目盛りの不正確さ、によって誤差が生じたためだと考えられる。

直結によるモータの外部特性

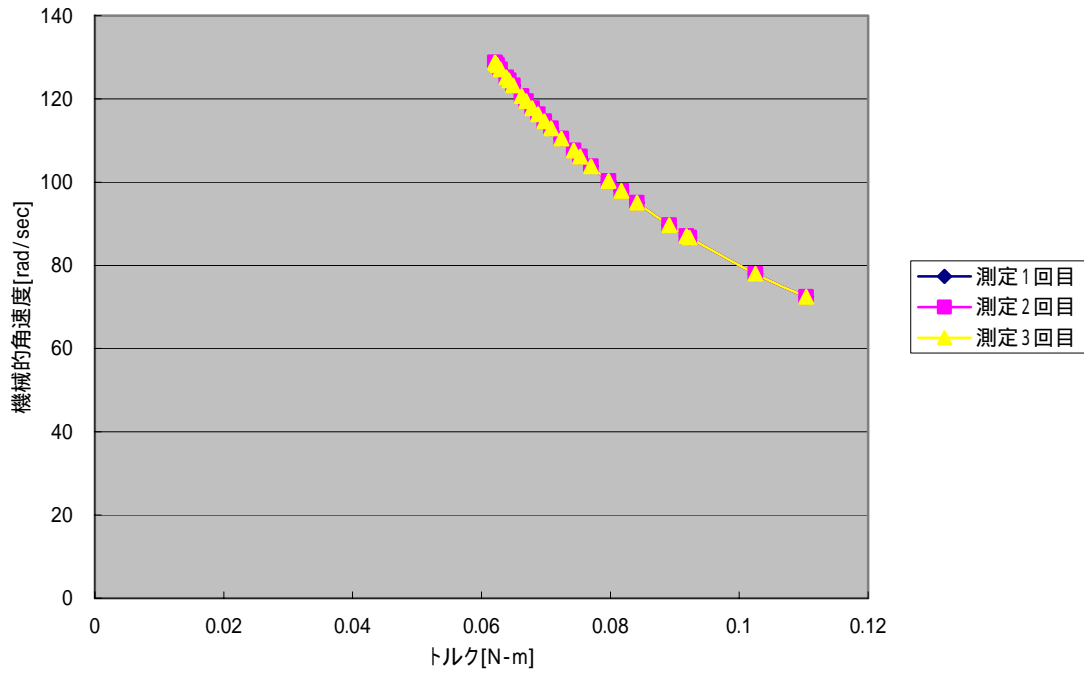


図 2 5 直結によるモータの外部特性

磁気歯車の外部特性

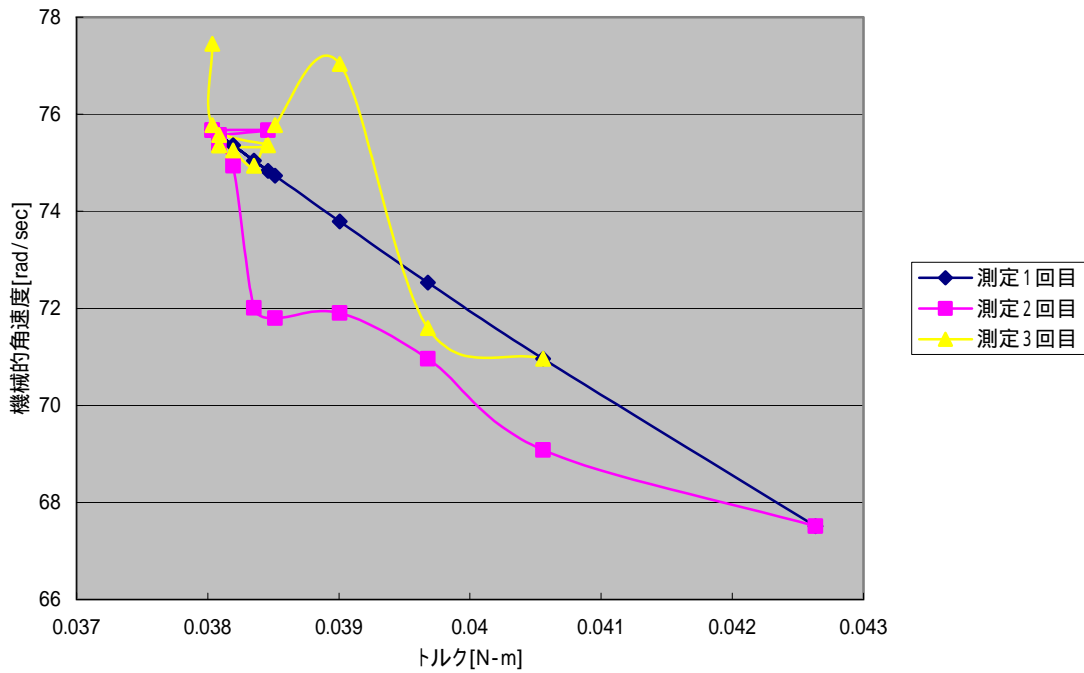


図 2 6 磁気歯車の外部特性

実験 2 伝達効率の測定

(B)実験方法の節で述べたように、入力側モーターと出力側モーターのトルクの比率と出力トルクを測定し、縦軸に効率、横軸に出力トルクをとってグラフ化したもの（トルクの伝達効率曲線）を連結装置は入力電圧 16V 一定、磁気歯車は入力電圧 16V 一定の場合でそれぞれ示す。

なお、これも外部特性の測定同様、再現性を得るために同じ測定を 3 回ずつ繰り返し行った。

トルクの伝達効率特性については、上の外部特性の測定のところで述べた原因により誤差が生じ、実験結果にばらつきが出たと考えられる。

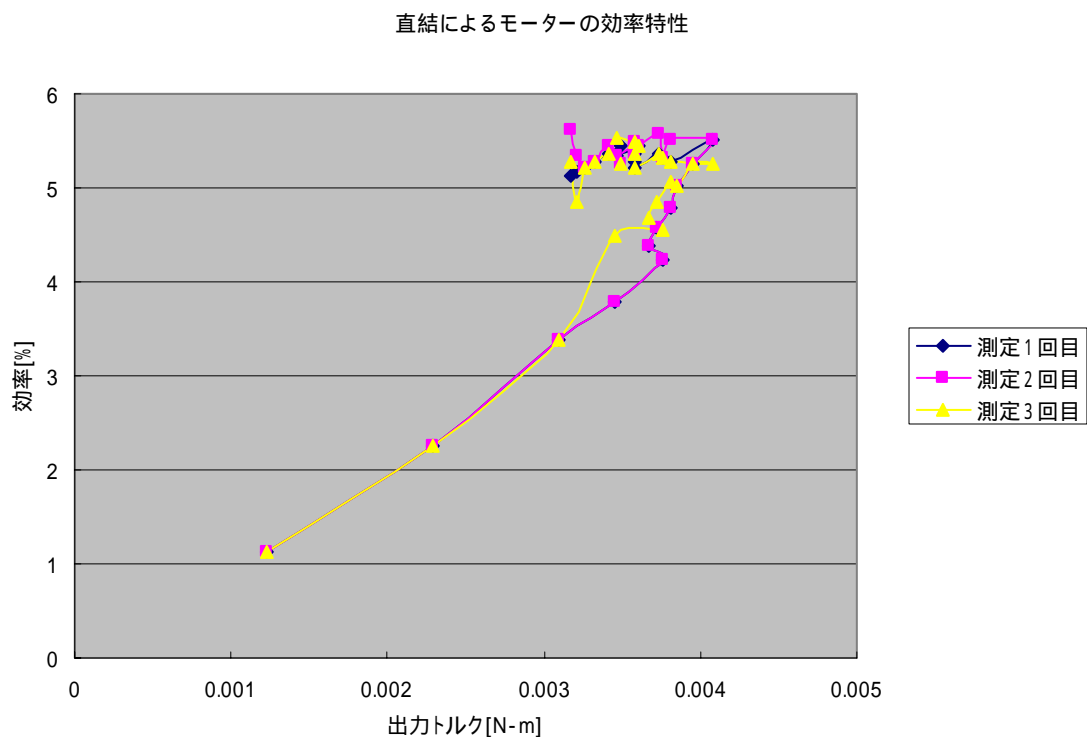


図 2 7 直結によるモーターの効率特性

磁気歯車の効率特性

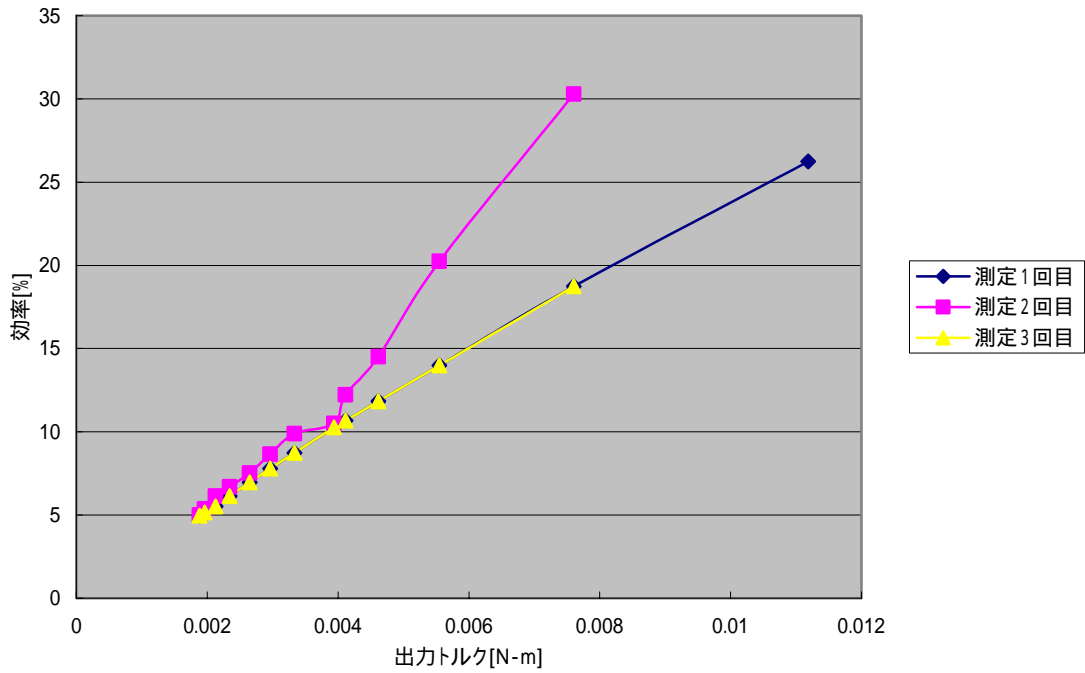


図 2 8 磁気歯車の効率特性

実験 3 入出力特性の測定

(B)実験方法の節で述べたように、出力側電機子端子に接続されている抵抗の電力を測定し、縦軸に出力電力、横軸に入力電力をとってグラフ化したもの（効率曲線）をそれぞれ示す。なお、これも外部特性の測定同様、再現性を得るために同じ測定を 3 回ずつ繰り返し行った。

直結によるモーターの入出力特性

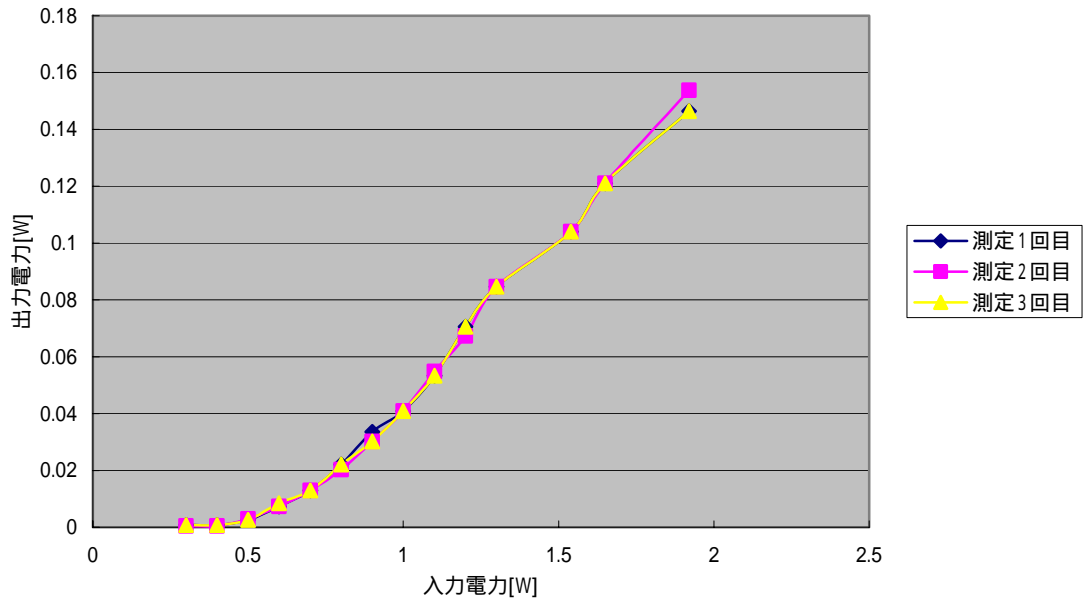


図 2 9 直結によるモーターの入出力特性

磁気歯車の入出力特性

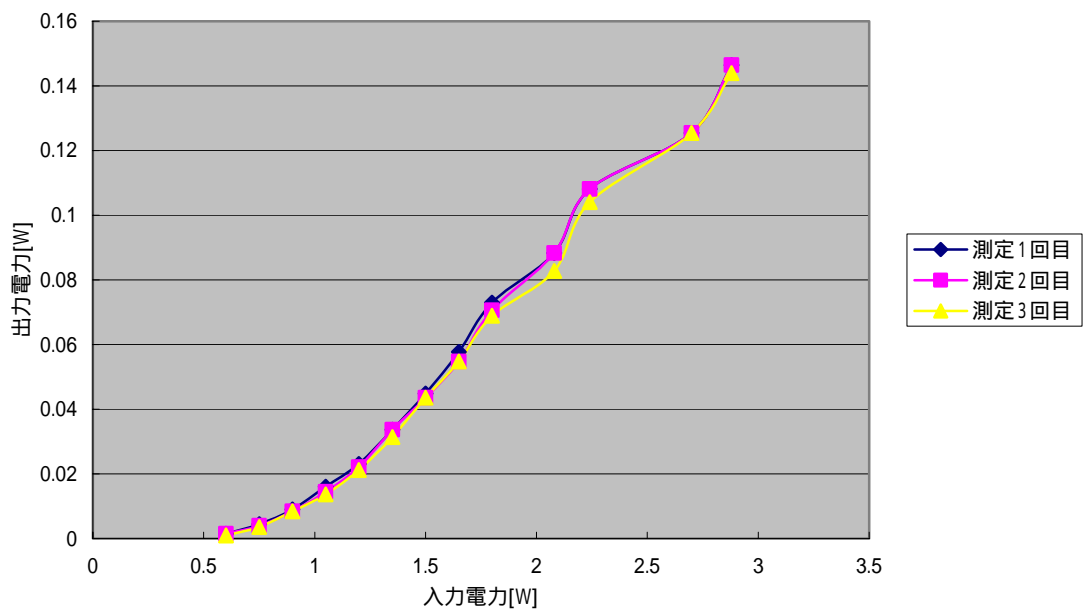


図 2 10 磁気歯車の入出力特性

(D)考察

今回、直結による連結装置と磁気歯車の外部特性と効率特性、入出力特性を測定した。この両者の外部特性と効率特性、入出力特性を比較してみて、実験的にわかった磁気歯車の長所と短所はいかに示す通りである。

長所

1. 両者の伝達効率の測定を行った結果、磁気歯車に印加する入力電圧を大きくしていくにつれて伝達効率も上昇していくことが図 2-7、2-8 を見るとわかる。また、モーターを直結で連結した装置に比べ、磁気歯車の方の効率が約 3 割高いことがわかった。
2. 両者の外部特性の測定結果より、直結での連結装置と磁気歯車の外部特性曲線を見比べると、磁気歯車の方はほぼ 1/2 の入力電圧で連結装置の外部特性曲線とほぼ等しい曲線が得られた。したがって、磁気歯車を用いれば、従来の直結接触型の動力伝達装置に印加した入力電圧の 1/2 の入力電圧でそれと同等の電力を伝達させることができると考えられる。今回作成した磁気歯車では直結での装置と同じ入力電圧を印加できる程の強度がないため、同じ入力電圧を印加した時の外部特性曲線の比較はできないが、仮に同じ入力電圧を印加したとすれば、磁気歯車の方が伝達される電力が大きくなると考えられる。
3. 上の 2 つで述べたように磁気歯車の方が、従来の接触型の動力伝達装置よりも伝達電力が大きい。すなわち伝達効率が上がっているということは非接触型の動力伝達装置である磁気歯車を用いたため、接触部分の摩擦損失が減少したということであり、実験的にだが、磁気歯車の長所が証明できたと考えられる。

短所

1. 今回作成した磁気歯車の実験装置では、出力側の歯車にすべりが生じ入力側の歯車との同期が外れるとその歯車は回転をせずにそのまま止まってしまった。
2. 磁気歯車の実験装置の両歯車は常時完全にロックした状態である。言い換えれば、常に入力側モーターに負荷がかかっている状態であるため、始動するには大電圧を印加しなければ始動しない。

以上の長所と短所が、磁気歯車の特性の測定によって確認できた。

3. 磁気歯車が与える周辺磁界分布

磁気が働く空間の状態を“磁界”という。磁気ははたらく空間の状態を「磁界」といい、磁界は磁石の周りだけでなく、電気的には電流が流れている物の周りには必ず磁界が発生する。私たちの身の周りの洗濯機やテレビなどの家庭電化製品や送電線からも磁界は発生している。これらのものからどれくらいの磁界が発生しているのかを図1に示す。

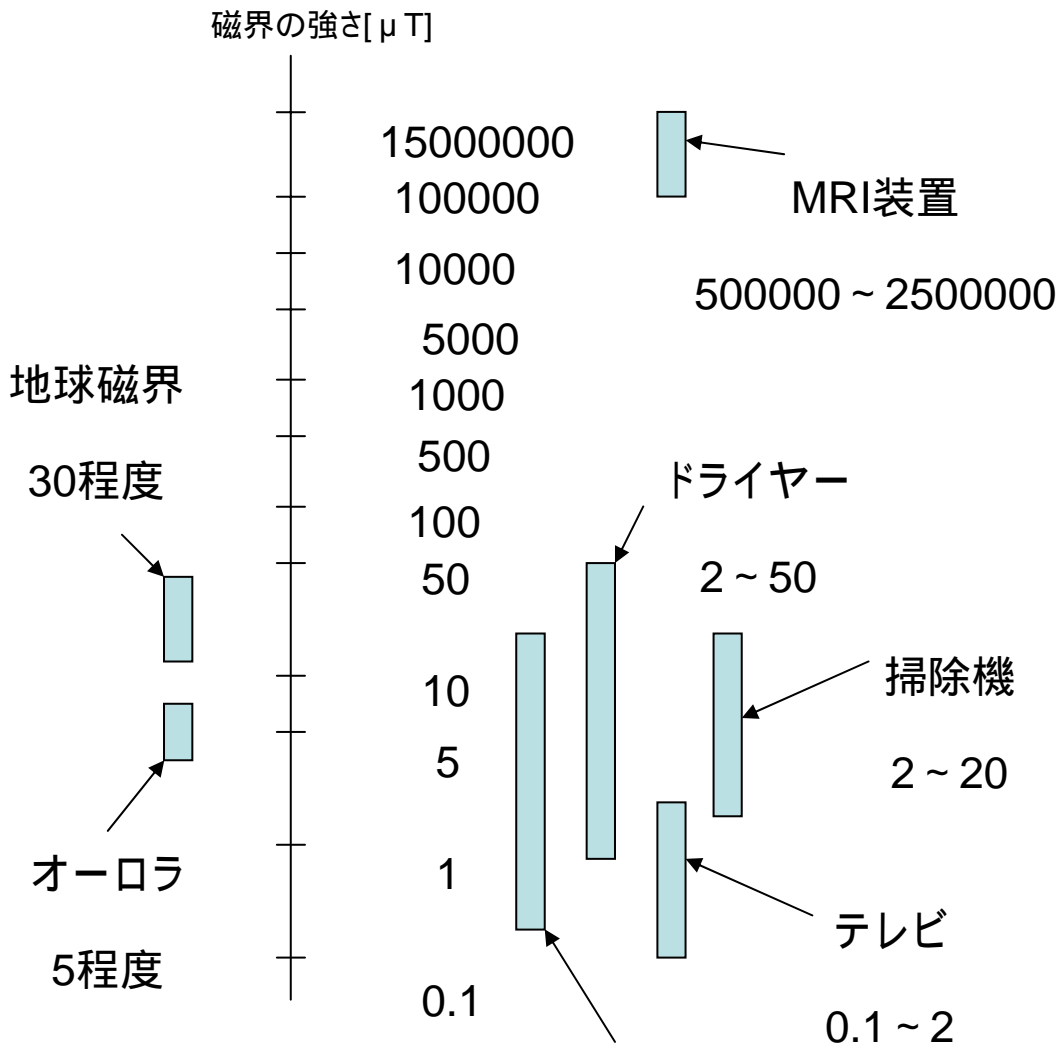


図3 1 身の周りの磁界の強さ

磁界の大きさは、「電界」と同じく発生源から離れると急激に小さくなる。磁界の強さ(磁束密度)を表す正式な単位は [T] (テスラ) ですが、従来よく用いられていた [G] (ガウス) が使われることもある。

ガウスとテスラの関係は

$$1[G] = 0.1[T]$$

のようになる。

ここでは、今まで述べてきた磁気歯車の周囲にはどの程度の磁界が発生するのかをサーチコイルを用いて測定しその分布について考察する。

3 1 . 測定原理

3 1 1 . 界逆問題のシステム方程式導出

図 3 2 で示すように、ある平面導体に分布する電流密度を平面導体上の磁界を測定することによって求める問題を考える。

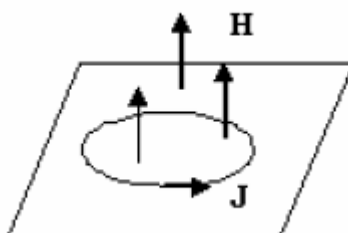


図 3 2 ループ電流と磁界

変位電流密度 $d D / d t$ が無視できる系において電流密度は磁界と次式のような関係になる。

$$\nabla \times H = J \quad \dots 3$$

従って、電流密度 J は必ず次式を満足しなければならない。

$$\nabla \times \nabla \times H = \nabla \times J \quad \dots 3$$

この式の関係はキルヒホッフの節点則と等価であり、電流密度を面積について積分したある点に流入する電流 i は流出電流と等しいことを意味する。

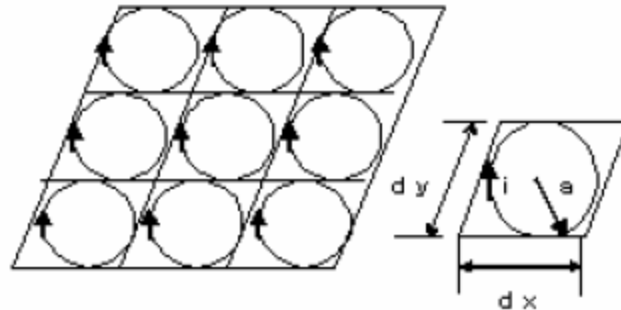


図 3 4 微小ループ電流モデル

図 3 4 の微小長方形中を循環する電流を同図中に示す円形ループで置き換える。円形ループの半径は式 3 によって決める。式 3 で決まるループ半径 a は微小長方形で縦横比が等しい正方形に内接する円を意味する。この場合、問題対象全体を通して各正方形の中心とモデルループの中心が一致し、最も良い近似となる。縦横比が異なる微小長方形では、各長方形の中心と式 3 で決まる半径を持つループの中心を一致させることから問題全体の形状を正方形に変形して考えることになる。従って、微小ループ電流モデルを適用する場合、分割された微小領域は正方形とすべきである。これは、例えば問題対象が長方形の場合、縦と横で分割数を変えることで実現される。

$$a = \sqrt{\left(\frac{dx}{2}\right)\left(\frac{dy}{2}\right)} \quad \dots 3$$

ループ電流モデルでは対象領域を微小長方形で分割し、その微小長方形の内接円を考えているのでどうしても対象領域全体を満足することが出来ない。よってシステム的には不安定なシステムとなってしまう可能性がある。そこで式 3 で示すように内接円の半径を倍にすることで外接円となり、円ループが重複するようにすることで対象領域を満足し、より安定なシステムを得ることができると考えられる。これを修正ループ電流モデルとする。

3 2 . 実験方法、実験装置

(A)実験装置

今までの実験で使用してきた磁気歯車の周囲に図 3 5 のようにサーチコイルを設置し、磁界を測定していく。

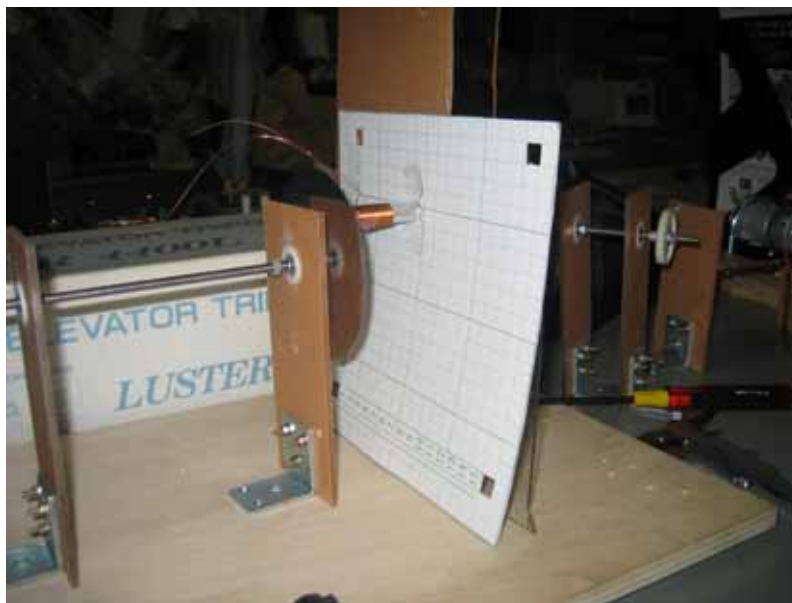


図 3 5 サーチコイル設置

(B)実験方法

実験 1 磁気歯車の周囲の磁界の測定

- 1 1 磁気歯車から 1.5[cm]のところの方眼紙を固定します。次に、半径 0.5[cm]のサーチコイルを準備し、オシロスコープと接続する。中心から上下右左に 1[cm]間隔で計 5[cm]をサーチコイルで誘導起電力を測定する。つまり、10×10 の 100 点の誘導起電力を測定する。
- 1 2 磁気歯車と包含用紙の距離を 2.0[cm]、2.5[cm]とし、同様の実験を行う。

今回の測定に用いた計算式を次に示す。

$$H = \frac{NI}{L} \text{ [A/m]} \quad \dots 3$$

N : 巻き数、 I : ループ電流、 L : 磁路長

(C)実験結果

実験 1 磁気歯車の周囲の磁界の測定

- 1 1 磁気歯車から 1.5[cm]のところの方眼紙を固定する。次に、半径 0.5[cm]のサーチコイルを準備し、オシロスコープと接続する。中心から上下右左に 1[cm]間隔で計 5[cm]をサーチコイルで誘導起電力を測定する。つまり、 10×10 の 100 点のサーチコイルの誘導起電力を測定する。
- 1 2 磁気歯車と方眼紙の距離を 2.0[cm]、2.5[cm]とし、同様の実験を行う。

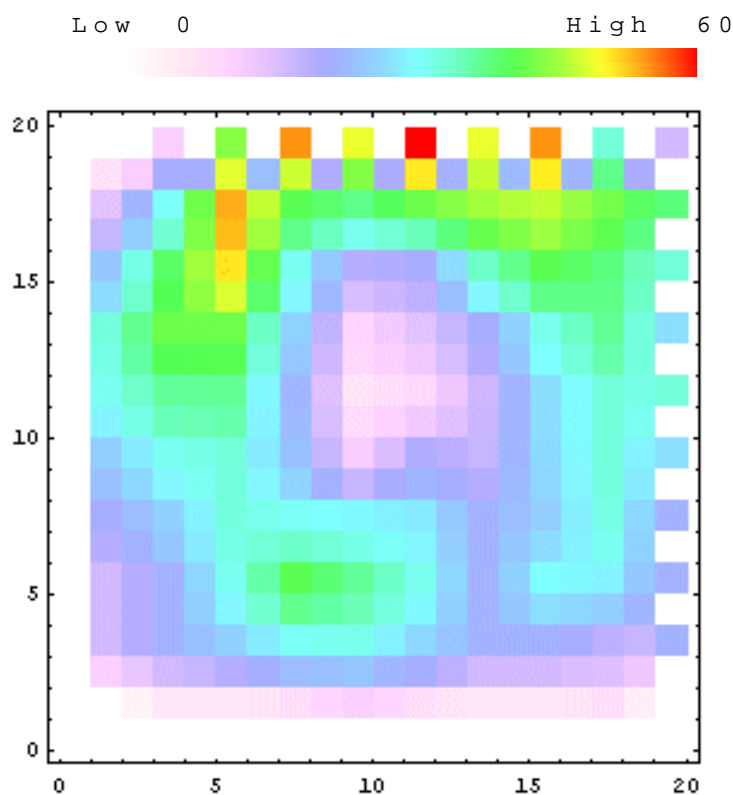


図 3 6 磁気歯車との距離 1.5[cm]

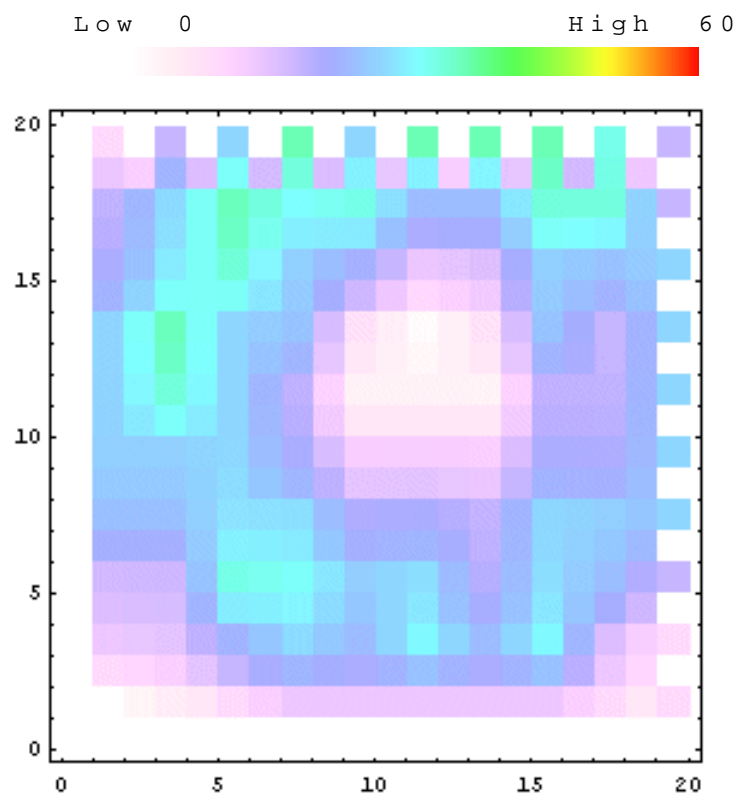


図 3 7 磁気歯車との距離 2.0[cm]

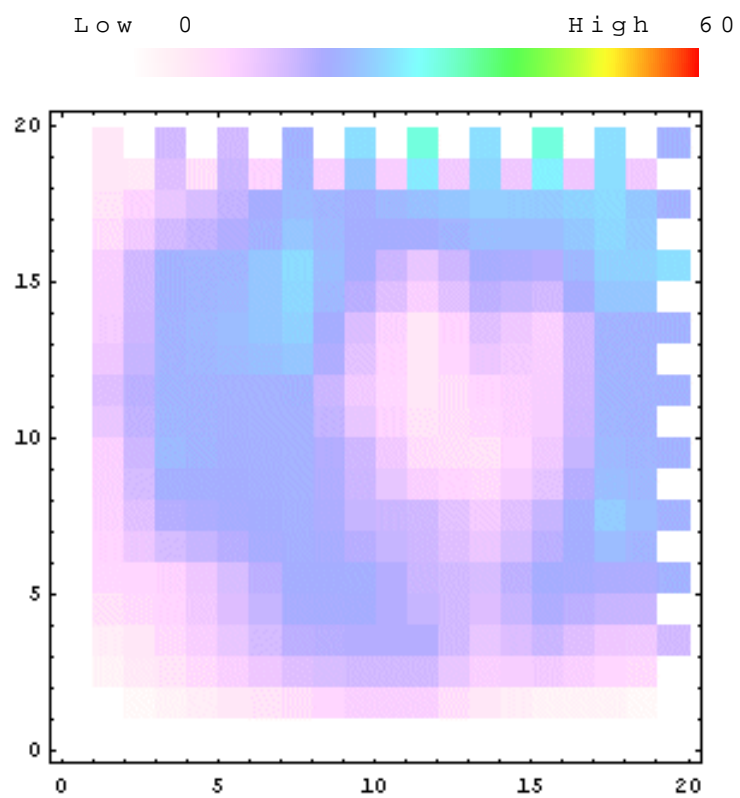


図 3 8 磁気歯車との距離 2.5[cm]

(D)考察

ここまでで、磁気歯車の周辺の磁界分布の状況が測定できた。

磁界は、式 3 、 3 からわかるように磁石に近ければ近いほど磁界は強くなっていく。作成した磁気歯車は磁石をドーナツ形に配置したため、値にばらつきがあるもののその特徴が出ているため、実験は成功だと考えられる。

磁気歯車は、家庭電化製品に比べて多くの磁界が発生してしまうため、身近な電化製品や、医療機器の使用は困難である。しかし、磁気歯車を用いれば、従来の直結接触型の動力伝達装置に印加した入力電圧の 1/2 の入力電圧でそれと同等の電力を伝達させることができるので、発電所や工場の大型設備の歯車としては十分利用できると考えられます。

きれいなドーナツ形にならなかったのは、正確に目盛りが読み取れなかったことや、計測器自体の目盛りの不正確さによって誤差が生じたためだと考えられます。

4. まとめ

以上、ここまでいろいろな装置を作成し実験を行ってきたが、装置などはすべて自作であるためその精度の面での伝達効率のロスなどは無視できるほど小さいものではないと考えられる。また、どの実験結果においても実験のみで、シミュレーションを行っていないため理論上の外部特性や効率特性と比較していないためどこまで信頼性がある実験結果かどうか分からない。

しかし、今までの実験を通して理論的には、非接触型の動力伝達機構いろいろな場所の利用が可能で、最初に述べたエンジンのクラッチ以外にも様々なものに応用できるものだと考えられる。

5. 参考文献

- 社団法人 電気学会著、“電気磁気学(第二次改訂版) ”、オーム社、1950 年初版
- 応用電磁著工学研究室、“応用電磁工学入門 ”、参図テクス、1996 年初版
- 斎藤兆古・早野誠治、“電気機器通論 ”、参図テキスト、2003 年

2005 年度

ウェーブレット変換を用いたデータ・ハインディング

Data Hiding By Wavelets Transform

法政大学工学部電気電子学科 4 年

学籍番号 c01d3088

高木信郎

指導教授 齋藤 兆古

第1章	序論	-2-
第2章	ウェーブレット変換の基礎	-2-
	2.1 概要	
	2.2 離散値系ウェーブレット変換の考え方	
	2.3 まとめ	
第3章	画像の圧縮	-14-
	3.1 概要	
	3.2 提案手法	
	3.3 数値実験	
	3.4 考察	
第4章	データハインディング	-18-
	4.1 概要	
	4.2 提案手法	
	4.3 数値実験	
	4.4 考察	
第5章	まとめ	-24-
第6章	参考文献	-24-

第1章 序論

データハインディングは透かし技術やステガノグラフィの総称であり、原画像に表れないよう隠し入れる事である。今回データハインディングの手法は原画像の周波数空間上において隠し入れる手法を使う。

ここで必要なのは、ハインドする為に原画像の周波数帯の決定が必要となる。

RGB カラー画像を用いたデータハインディングの場合、非カラー画像に比べ情報量が多いためハインディング能力が高い。

既存方法で、原画像のある1バンドに別の画像を隠し入れる方法がある。例えばG成分に画像を隠し入れる場合、G・B成分の情報は使用しないという具合に行う。

また音声データも一次元の数値データではあるが、データハインディングが当然できる。原音声データを周波数空間上に置き、ある部分をカットしその部分に別の音声データを隠し入れ、元の時間空間に戻しハインディングを行う。

本論文では、2次元データの画像の3バンドすべてに、一次元データの音声データを隠し入れる。

一次元のデータハインディングをうまく使い、二次元データの中に1次元データを隠し入れる技法を提案し、有効性を確認したので、ここに報告する。

第2章 ウェーブレット変換の基礎

2.1 概要

2.1.1 関数系の変換

古典的な関数変換の目的は、解析的に扱いにくい関数系を解析的に扱いやすい関数系へ変換することである。例えば、ラプラス(Laplace)変換は微積分演算を単純な掛け算や割り算へ化す変換である。また、関数系の変換とは、ある関数、例えば時間変化する関数を解析が容易な周期関数の和で表現することにも使われる。具体的な例としてフーリエ(Fourier)変換を考えよう。フーリエ変換は解析的に扱えない関数を解析的に扱える角周波数の異なる正弦波と余弦波の和で表現する変換である。換言すれば、フーリエ変換は解析的に扱いにくい関数系を解析的に扱いやすい関数系へ分解する変換と考えてもよい。

ウェーブレット変換は、ラプラス変換のように解析的な計算を容易な形へ変換する方法でなく、フーリエ変換のように与えられた関数を他の関数系の和で表現するために用いられる。いま、ある任意の時間 t をパラメータとする関数 $f(t)$ を一定値 a_0 、正弦波および余弦波の和で表現できるとしよう。すなわち、
を角周波数として、

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin(i\omega t) + \sum_{i=1}^{\infty} b_i \cos(i\omega t) \quad (1)$$

と仮定する。

問題は(1)式の係数、 a_0 、 a_i 、 b_i を決める方法である。いま、関数 $f(t)$ が時間 $t=0$ から $t=T$ の区間で(1)式の係数を決めることを考えれば、(1)式は、 $\omega = 2\pi/T$ であるから、

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i\frac{2\pi}{T}t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i\frac{2\pi}{T}t\right) \quad (2)$$

とも書ける。

(2)式の両辺を時間 $t=0$ から $t=T$ の区間について積分すると、

$$\begin{aligned} & \int_0^T f(t) dt \\ &= \int_0^T \left\{ a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i\frac{2\pi}{T}t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i\frac{2\pi}{T}t\right) \right\} dt \quad (3) \\ &= a_0 \end{aligned}$$

が成り立つ。

したがって、定数項 a_0 は

$$a_0 = \frac{1}{T} \int_0^T f(t) dt \quad (4)$$

となる。

次に、(2)式の両辺に正弦波 $\sin[j(2\pi/T)t]$ 、 $j=1,2,3\dots$ を掛け算し、時間 $t=0$ から $t=T$ の区間について積分する。

$$\begin{aligned} & \int_0^T f(t) \sin\left(j \frac{2\pi}{T} t\right) dt \\ &= \int_0^T \left\{ a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i \frac{2\pi}{T} t\right) \right\} \times \sin\left(j \frac{2\pi}{T} t\right) dt \quad (5) \\ &= a_i \frac{2}{T}, i=j \text{ or } 0, i \neq j \end{aligned}$$

(5)式から係数 a_i は

$$a_i = \frac{2}{T} \int_0^T f(t) \sin\left(i \frac{2\pi}{T} t\right) dt \quad (6)$$

として得られる。

同様に、(2)式の両辺に余弦波 $\cos[j(2\pi/T)t]$ 、 $j=1,2,3,\dots$ を掛け算し、時間 $t=0$ から $t=T$ の区間について積分することで

$$b_i = \frac{2}{T} \int_0^T f(t) \cos\left(i \frac{2\pi}{T} t\right) dt \quad (7)$$

として係数 b_i が得られる。

したがって、関数 $f(t)$ は、区間 $t=0$ から $t=T$ で、

$$\begin{aligned} f(t) &= \frac{1}{T} \int_0^T f(t) dt + \\ & \sum_{i=1}^{\infty} \left\{ \frac{2}{T} \int_0^T f(t) \sin\left(i \frac{2\pi}{T} t\right) dt \right\} \sin\left(i \frac{2\pi}{T} t\right) + \sum_{i=1}^{\infty} \left\{ \frac{2}{T} \int_0^T f(t) \cos\left(i \frac{2\pi}{T} t\right) dt \right\} \cos\left(i \frac{2\pi}{T} t\right) \quad (8) \end{aligned}$$

と書ける。これがいわゆるフーリエ変換の原型となるフーリエ級数であり、左辺の関数 $f(t)$ を右辺の計算が簡単な定数項と三角関数の和に変換している。

2.1.2 関数の直交性と線形性

フーリエ級数の考え方の中に重要な関数間で成り立つ性質、すなわち、関数の直交性(orthogonality)が使われている。

まず(4)式の係数 a_i が計算される過程を考えよう。(3)式は(2)式の両辺に定数値1を掛け算し積分する演算である。このとき、

$$\int_0^T 1 \cdot \sin\left(i \frac{2\pi}{T} t\right) dt = 0, \int_0^T 1 \cdot \cos\left(i \frac{2\pi}{T} t\right) dt = 0, \quad (9)$$

$i = 1, 2, \dots$

の関係が成り立つために、(4)式の係数 a_0 がうまく計算できた。この関係を、定数値1と正弦波 $\sin[j(2\pi/T)t]$ 、および余弦波 $\cos[j(2\pi/T)t]$ 間の直交性と呼ぶ。同様に、(6),(7)式で計算される a_i 、 b_i は

$$\int_0^T \sin\left(i \frac{2\pi}{T} t\right) \sin\left(j \frac{2\pi}{T} t\right) dt = 0,$$
$$\int_0^T \cos\left(i \frac{2\pi}{T} t\right) \cos\left(j \frac{2\pi}{T} t\right) dt = 0, \quad (10)$$

$i \neq j$

なる直交性が成り立つことに基づいている。

(9),(10)式から、直交性とは与えられた関数を他の関数の和で表現しようとする場合、和となる関数の大きさ(係数)を一意的に決める条件であることがわかる。言い換えれば、与えられた関数を他の関数の和で表現できる条件である。

さて、ある任意の時間 t をパラメータとする関数 $f(t)$ は、区間 $t=0$ から $t=T$ で、一定値、正弦波および余弦波の和で表現できることがわかった。この変換は、一定値、正弦波および余弦波間で直交性が成り立つことが条件であった。この結果に至る過程を考えてみると、まず、展開される関数の和でもとの関数が表現されるとする大前提があることに気づくであろう。ある関数が他の関数の和で表現できる性質を線形性と呼ぶ。では、(6)式で計算される正弦波の係数 a_i が

定数 c_i の n 個の和で表現されるとしよう。すなわち、

$$a_i = c_i + c_i + \dots = nc_i \quad (11)$$

である。(11)式から、和を前提とする系では比例関係が成り立つことを意味することがわかる。すなわち、線形系とは比例関係が成立する系である。

2.1.3 連続系ウェーブレット変換

フーリエ級数で、和となる正弦波や余弦波を基底関数と呼ぶ。この基底関数に正弦波や余弦波のように直交性が成り立たない関数を採用することを考える。例えば、図1に示す、

$$\varphi(t) = (1-t^2)\exp\left(-\frac{t^2}{2}\right) \quad (12)$$

で与えられる関数である。

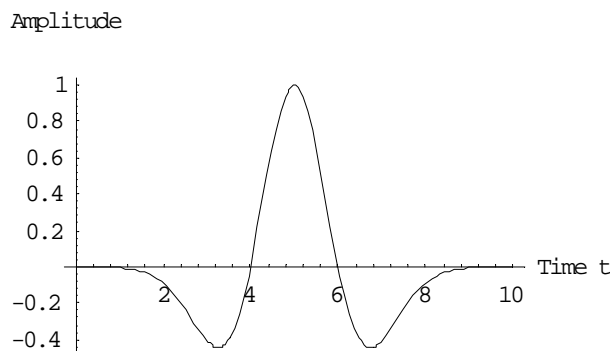


図1：非直交基底関数の例(メキシカンハット型)

この基底関数を用いて、

$$\Psi(t) = \frac{1}{\sqrt{|a|}} \varphi\left(\frac{t-b}{a}\right) \quad (13)$$

なる関数を考えよう。

(13)式で、パラメータ a は基底関数の広がり、すなわち、図1の時間 $t=1$ から $t=9$ までの周期 T を表し、 b は基底関数のピーク $t=5$ の時間位置を示す。フーリエ級数では、周期 T の異なる正弦波と余弦波を基底関数として考えた。(13)式では、周期 T に対応する a と時間位置 b をパラメータとする変換を行う。

(6), (7) 式と同様な形式の内積演算

$$T(a,b) = \frac{1}{\sqrt{c_\phi}} \int_{-\infty}^{\infty} \Psi^{(a,b)}(t) f(t) dt \quad (14)$$

を行い、 a , b をパラメータとする係数 $T(a,b)$ を計算する。

さて、この係数 $T(a,b)$ はどのような物理的な意味を持つであろうか。まず、パラメータ b で図1に示す波形のピーク時間位置が決まる。つぎに、パラメータ a で波形の周期 T が決まる。したがって、係数 $T(a,b)$ は、関数 $f(t)$ が周期 $T=a$ 、時間 $t=b$ でピークを持つ図1の波形と共通部分を取り出した値に相当することを意味する。言い換えれば、仮に関数 $f(t)$ が図1の波形と全く同じ波形であれば、係数 $T(a,b)$ は最大値をとる。これは、ある任意波形から、図1の波形と同じ波形成分を取り出す演算として(14)式がきわめて有効であることを意味する。この変換を連続ウェーブレット変換と呼ぶ。残念ながら、(12)式の基底関数はパラメータ a , b が異なっても直交性を持たない。このため、線形性も成り立たない。

2.1.4 離散値系ウェーブレット変換

離散値系(discrete)とはどのような系であろうか。離散値なる用語が生まれたのは計数型計算機を用いて数値計算を行う場合、連続関数をサンプリングして得られる数値で代表したことに起因する。例えば、関数 $f(t)$ を時間 $t=0$ から $t=T$ の区間で、 t ごとに n 個サンプリングして離散値系で表すと、

$$F = [f(\Delta t), f(2\Delta t), f(3\Delta t), \dots, f(n\Delta t)]^T \quad (15)$$

なるベクトルとなる。

同様に、(1)式を離散値系で書けば、

$$\begin{pmatrix} f(\Delta t) \\ f(2\Delta t) \\ f(3\Delta t) \\ \vdots \\ f(n\Delta t) \end{pmatrix} = a_0 \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \begin{pmatrix} \sin\left(\frac{2n}{T}\Delta t\right) & \sin\left(2\frac{2n}{T}\Delta t\right) & \sin\left(3\frac{2n}{T}\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}\Delta t\right) \\ \sin\left(\frac{2n}{T}2\Delta t\right) & \sin\left(2\frac{2n}{T}2\Delta t\right) & \sin\left(3\frac{2n}{T}2\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}2\Delta t\right) \\ \sin\left(\frac{2n}{T}3\Delta t\right) & \sin\left(2\frac{2n}{T}3\Delta t\right) & \sin\left(3\frac{2n}{T}3\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}3\Delta t\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sin\left(\frac{2n}{T}n\Delta t\right) & \sin\left(2\frac{2n}{T}n\Delta t\right) & \sin\left(3\frac{2n}{T}n\Delta t\right) & \cdots & \sin\left(n\frac{2n}{T}n\Delta t\right) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} \cos\left(\frac{2n}{T}\Delta t\right) & \cos\left(2\frac{2n}{T}\Delta t\right) & \cos\left(3\frac{2n}{T}\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}\Delta t\right) \\ \cos\left(\frac{2n}{T}2\Delta t\right) & \cos\left(2\frac{2n}{T}2\Delta t\right) & \cos\left(3\frac{2n}{T}2\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}2\Delta t\right) \\ \cos\left(\frac{2n}{T}3\Delta t\right) & \cos\left(2\frac{2n}{T}3\Delta t\right) & \cos\left(3\frac{2n}{T}3\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}3\Delta t\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \cos\left(\frac{2n}{T}n\Delta t\right) & \cos\left(2\frac{2n}{T}n\Delta t\right) & \cos\left(3\frac{2n}{T}n\Delta t\right) & \cdots & \cos\left(n\frac{2n}{T}n\Delta t\right) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \quad (16)$$

or

$$\mathbf{F} = a_0 \mathbf{I} + \mathbf{A}\mathbf{S} + \mathbf{B}\mathbf{C}$$

となる。ただし、 \mathbf{I} は n 次の単位列ベクトルである。

また、(16)式右辺の係数 a 、ベクトル \mathbf{S} と \mathbf{C} は、

$$a_0 = \frac{1}{n} \mathbf{I}^T \mathbf{F}, \mathbf{S} = \frac{2}{n} \mathbf{A}^T \mathbf{F}, \mathbf{C} = \frac{2}{n} \mathbf{B}^T \mathbf{F} \quad (17)$$

で与えられる。

さて、離散値系でフーリエ係数を計算する過程で、(16)式を

$$\begin{aligned} \mathbf{F} &= a_0 \frac{1}{\sqrt{n}} \mathbf{I} + \sqrt{\frac{2}{n}} \mathbf{A} \mathbf{S} + \sqrt{\frac{2}{n}} \mathbf{B} \mathbf{C} \quad (18) \\ &= a_0 \mathbf{I}^T + \mathbf{A}' \mathbf{S} + \mathbf{B}' \mathbf{C} \end{aligned}$$

と書き直すと、

$$a_0 = \frac{1}{\sqrt{n}} \mathbf{I}'^T \mathbf{F}, \mathbf{S} = \sqrt{\frac{2}{n}} \mathbf{A}'^T \mathbf{F}, \mathbf{C} = \sqrt{\frac{2}{n}} \mathbf{B}'^T \mathbf{F} \quad (19)$$

の関係が成り立つ。 \mathbf{I} は n 次の単位行列である。この結果は、離散値系で正弦波や余弦波のフーリエ係数を求める場合、(19)式の係数行列 \mathbf{A}' 、 \mathbf{B}' の逆行列がそれぞれの転置行列で与えられることを意味する。言い換えれば、変換行列の逆行列が変換行列の転置行列で与えられることでフーリエ係数の直交性が満足される。

離散値系のウェーブレット変換は、連続系ウェーブレット変換の基底関数に対応する変換行列が常に直交性を満足する線形変換の一種である。ただし、ウェーブレット変換に用いる変換行列は、正弦波や余弦波のように連続な基底関数をサンプリングして構築されず、最初から離散値系で与えられる変換行列である。これは、従来の離散値系が、サンプリング個数を無限大にした場合、連続系に化することを前提にしているのに対し、サンプリング個数を無限大にしても連続系に化さない離散値系を前提とすることを意味する。換言すれば、離散値系のウェーブレット変換は、線形変換を行う演算が計数型計算機で行われることを大前提としている。この意味で離散値系ウェーブレット変換は、従来の離散値系と一線を画する離散値系である。

2.2 離散値系ウェーブレット変換の考え方

2.2.1 ウェーブレット変換行列

いま、 a と b なる数値を考えてみると、この数値の線形的な組合せは、

$$\begin{aligned} s &= a + b \\ d &= a - b \end{aligned} \quad (20)$$

が考えられる。さて、逆に s と d が与えられ a と b を求めようとするれば、

$$\begin{aligned} a &= \left(\frac{1}{2}\right)(s + d) \\ b &= \left(\frac{1}{2}\right)(s - d) \end{aligned} \quad (21)$$

となる。

(20)式の関係を行列を使って書くと、

$$\mathbf{X}' = C \cdot \mathbf{X} \quad (22)$$

と書ける。

ここで、 \mathbf{X}' , C , \mathbf{X} はそれぞれ、

$$\mathbf{X}' = \begin{bmatrix} s \\ d \end{bmatrix}, \mathbf{X} = \begin{bmatrix} a \\ b \end{bmatrix}, C = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (23)$$

である。他方、(21)式は、

$$\mathbf{X} = D \cdot \mathbf{X}' \quad (24)$$

と書ける。ただし、行列 D は、

$$D = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (25)$$

である。行列 C と D との関係は、

$$D \cdot C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (26)$$

となる。すなわち D は C の逆行列になっている。しかし、(25)式の行列の係数

(1/2)をCとDに平等に分配して、

$$C = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, D = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (27)$$

とすれば、行列DはCの転置行列であり、かつ逆行列となる。すなわち、

$$C^{-1} = C^T \quad (28)$$

である。次に a 、 b の2個の数値だけでなく、 a 、 b 、 c なる数値の組合せを考える。この場合、

$$\begin{aligned} s_1 &= a+b, s_2 = b+c, s_3 = c+a \\ d_1 &= a-b, d_2 = b-c, d_3 = c-a \end{aligned} \quad (29)$$

等の組合せが考えられる。この関係を行列で書くと、

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad (30)$$

となる。(30)式右辺の係数行列は6行3列の長方形行列であるため、(28)式のようにうまく逆行列を得られない。これは、組み合わせる数値の個数が奇数であると(24)-(28)式のような線形変換が適用できないことを意味する。では、 a 、 b 、 c 、 d の4個の数値の組みを考える。最初に考えられる組合せは、

$$\begin{bmatrix} s_1 \\ d_1 \\ s_2 \\ d_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad (31)$$

となる。これには、当然、

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} s_1 \\ d_1 \\ s_2 \\ d_2 \end{bmatrix} \quad (32)$$

が成り立つ。しかし、 a と b 、 c と d それぞれの組合せはできるが、 a と c 、 b と c 等の組合せはできない。このため、(29)式の左辺を並べ換えて、次の組合せを得る。

$$\begin{bmatrix} S_1 \\ D_1 \\ d_1 \\ d_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ d_1 \\ d_2 \end{bmatrix} \quad (33)$$

(31)式で、 a 、 b 、 c 、 d の4個の数値の組合せがすべて揃ったこととなり、結局、もとの $\{a, b, c, d\}^T$ のベクトルは(29)と(31)式の2回の線形変換で、和 S_1 と差 D_1 、 d_1 、 d_2 を要素とするベクトル $\{S_1, D_1, d_1, d_2\}^T$ に変形された。(31)式の逆変換は

$$\begin{bmatrix} s_1 \\ s_2 \\ d_1 \\ d_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} S_1 \\ D_1 \\ d_1 \\ d_2 \end{bmatrix} \quad (34)$$

によって行われ、(32)式の左辺を(30)式右辺のベクトルの形に並べ換えて、(30)式に代入することで完全な逆変換ができる。(29)式から(31)式までの変換をまとめて書くと、

$$\begin{bmatrix} S_1 \\ D_1 \\ d_1 \\ d_2 \end{bmatrix} = \left(\frac{1}{\sqrt{2}} \right)^2 \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \quad (35)$$

または、

$$\mathbf{S} = \mathbf{W} \cdot \mathbf{X} \quad (36)$$

となる。 \mathbf{W} はウェーブレット変換行列であり、この変換行列を構成する基底関数をハール(Haar)基底とよぶ。

(34)式の逆変換は \mathbf{W} の転置行列 \mathbf{W}^T が \mathbf{W} の逆行列になるから、

$$\mathbf{X} = \mathbf{W}^T \cdot \mathbf{S} \quad (37)$$

で与えられる。

次に、6個の数値の組合せを考えると、(29)式と同じ形で第1回の変換は、和が3組と差が3組となり、(28)式と同様に第2回の変換はできない。したがって、数値の組を和と差に分解する線形変換は、組を構成する数値の個数が、偶数でかつ、2、4、8、16、32等、2のべき乗でなければならない。

2.3 まとめ

本章では、フーリエ解析の考え方から出発し、離散値系ウェーブレット変換の具体的な例題までを述べた。フーリエ級数から、関数の変換、直交性、さらに線形性の概念を説明し、フーリエ解析の拡張として連続系ウェーブレット変換の概念を紹介した。つぎに、連続関数をサンプリングして得られる離散値系について、フーリエ級数の離散値形を導き、離散値系における直交性や線形変換の概念を紹介した。

第3章 画像の圧縮

3.1 概要

ウェーブレット変換によるデータ圧縮の原理は、もとのデータの特徴を表す絶対値の大きなスペクトラムのみを残し、他をゼロとすることにある。本章はウェーブレット変換を使ったイメージデータの圧縮方法について述べる。サンプルデータは、512×512 bmp 画像サイズ、256階調×3の3バンド画像を使う。以下にスペクトラムの特定部分を残す領域法について述べる

3.2 提案手法

原画像をそれぞれRGB成分別に取り出す
各成分をウェーブレット変換しデータをカットする
カットした部分すべてに0を埋め込む
逆ウェーブレット変換して圧縮完成

3.3 数値実験

3.3.1 使用データ

本実験で用いるサンプルイメージを図1に示す



Fig.1 原画像

3.3.2実験

図1を使用し、RGB別にする各それぞれのデータ数262144個。

図2aにウェーブレット変換しデータ数65536個の5%にカット、カットした95%に0を埋めたものを示す。

図2bにウェーブレット変換しデータ数65536個の25%にカット、カットした75%に0を埋めたものを示す。

図2cにウェーブレット変換しデータ数65536個の50%にカット、カットした50%に0を埋めたものを示す。

図2dにウェーブレット変換しデータ数65536個の75%にカット、カットした25%に0を埋めたものを示す。

図2eにウェーブレット変換しデータ数65536個の約88%にカット、カットした約12%に0を埋めたものを示す。



Fig.2a 圧縮率5%の画像



Fig.2b 圧縮率25%の画像



Fig.2c 圧縮率50%の画像



Fig.2d 圧縮率75%の画像



Fig.2e 圧縮率約88%の画像

3.3.3 相関係数

原画像データと圧縮画像,圧縮画像をbmpに保存後の相関係数を表3に示す。

圧縮率 [%]	圧縮画像との相関係数[%]	圧縮画像との相関係数 (bmp 保存後) [%]
5	81.2542	81.2522
25	90.6951	90.6935
50	94.6287	94.6274
75	98.7362	98.7353
88	99.5141	99.5136

Table.3
原画像との相関係数
及び bmp 形式に保存後の原画像との相関係数

3.4 考察

目視、相関係数からみても圧縮率5%は画像の再現率は低い。圧縮率約88%は原画像を再現できる

圧縮率75%は相関係数が原画像に近く高い再現率が望めた、また圧縮率50%を見ると目視した結果ほぼ圧縮率75%と同じである。

圧縮率25%は目視からは、完全に原画像に近くはなく画像は荒れるが確認はできる程度である。

画像の再現性から考察すると、圧縮率50%より高い圧縮率が適していると考えられる

Bmpに保存しても画像は壊れない事も考察できる。

第4章 データハインディング

4.1 概要

透かし技術は、

外に表れた情報が重要なもの

少量の目印情報(シグニチャ)を埋め込む
埋め込んだ情報(シグニチャ)は壊れない
という要求があり、

ステガノグラフィ技術は

外に表れない情報が重要なもの
埋め込むことができる容量は多いことが望ましい
場合により、埋め込んだ情報は壊れても構わない
という条件がある。

透かし技術およびステガノグラフィ技術として

第3者にハインディングされたものを意識させない
当事者はハインディングされたものを復元できる
第3者はハインディングされたものを復元しがたい
という要求がある。

さらにシグニチャの情報量が多いと第3者にハインディングされたものを意識され易くなる、第3者に意識させないためにはハインディングの劣化をふせぐ必要がある。

4.2 提案手法

原データを2次元画像、シグニチャに1次元の音声を使用する

3バンドからなる原画像の各成分ごとに分け解析する

2次元の原データを、1次元データにする

1次元化した原データをウェーブレット変換する

その後原画像をシグニチャ分のデータだけカットし、シグニチャを埋め込む
ウェーブレット再構成により3バンドからなるデータ・ハインディング画像
を生成する

問題はシグニチャの音声を、256階調の原データに埋め込むため整数化する必要がある。

4.3 数値実験

4.3.1 使用データ

使用データは図1に示す画像を原画像として用い、図2に示すハインディング音声とする。図1は512×512の画像サイズであり、画素あたり256階調×3の3バンド画像である。図2のハインディング音声は8000Hz mono、データ数95754個である。

4.3.2 実験

音声を各バンドに埋め込むため、音声データを 1~31918, 31919~63836, 63837~162000 の3つに分ける。
分けた音声をハイディングしやすい様にそれぞれ100倍し10を加え小数点以下を切り捨てたものを図4に示す

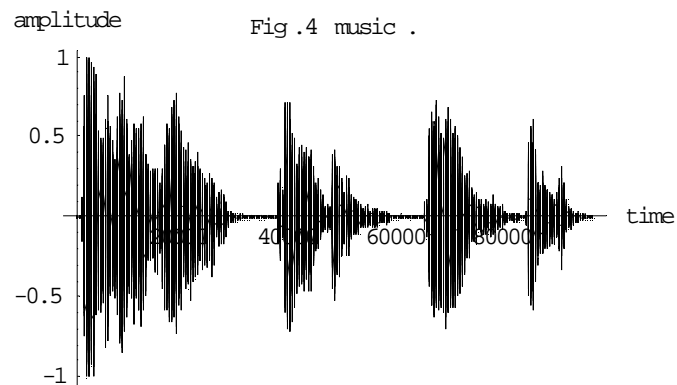


Fig.4 原音声データ

原画像データを1次元化し、各バンドそれぞれウェーブレット変換した後、データ数を230229個にカットする。(約88%に圧縮)

ウェーブレット変換後カットしたデータに音声データを各バンドに埋め込む、

画像のR成分には音声の1~31918個を

画像のG成分には音声の31919~63836個を

画像のB成分には音声の63837~95744個をとるという具合である。

各バンド総データ数262144

これを図5a,図5b、図5cに示す

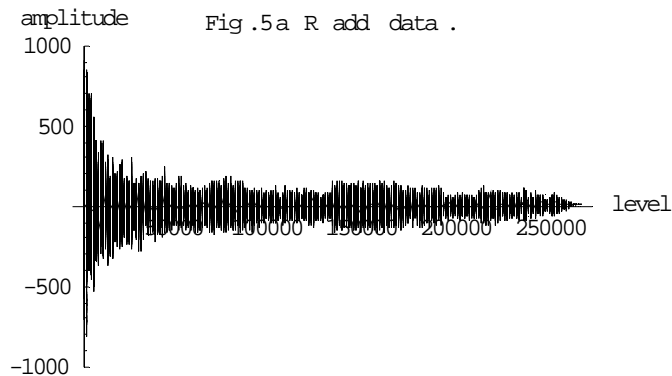


Fig.5a ウェーブレット変換後1~31918個の音声データ

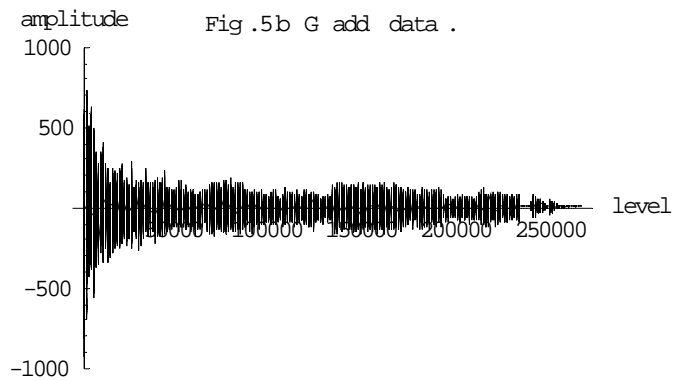


Fig.5b ウェーブレット変換後 31919~63836 個の音声データ

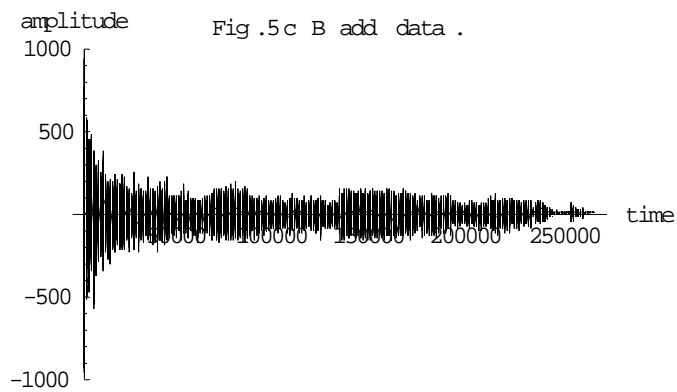


Fig.5c ウェーブレット変換後 63837~95744 個の音声データ

各バンドそれぞれ逆ウェーブレットしデータ・ハインディングをする
 各バンドを合成し2次元配列に戻しカラー画像に戻した画像を図6に示す



Fig.6 ハイディング画像

図6をbmp画像にし、図7に示す



Fig.7 ハイディング画像をbmp形式にした画像

図6をウェーブレット変換した後、音声を取り出した結果を図8に示す。

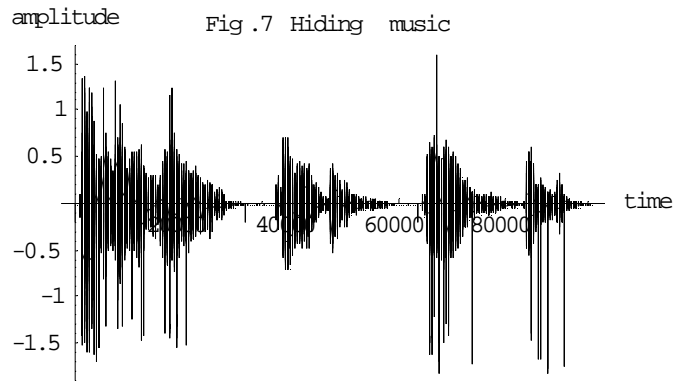


Fig.8 ハイディング画像から取り出した音声データ

4.3.3 相関係数

原画像とハイディング画像、ハイディング画像（bmp 保存後）との相関係数を表.8 に示す。

原画像とハイディング画像との相関係数[%]	原画像とハイディング画像との相関係数(bmp 保存後) [%]	原音声とハイディング画像から取り出した音声(bmp 保存後) [%]
98.8769	98.7588	90.8314

Table.8

原画像とハイディング画像との相関係数
及び

原音声データとハイディング後取り出した音声データとの相関係数

4.4 考察

原画像を 88%に圧縮しその部分に音声をハインドした。ハイディング画像の再現率は相関係数からは高いが、目視した感じでは完全に原画像と同じではないことがわかる。今回は一曲を三分割し 3 バンドにハインドしたが、とうぜんながら 3 バンドに別々の曲を入れることも可能である。

第5章 まとめ

データハインディングに適する画像の圧縮レベルは 50%以上が適していると考察でき、データハインディングによりハインドする。音声データの完全再現はできないが、相関係数・音声を聞いた結果、認識ができ画像の中に音声を入れるデータ・ハインディングは可能である。また原画像とハインディングされた画像とのデータサイズは同じになったので、第三者にはわからないようハインディングできた。

しかし今回使用した画像の形式が bmp で 768kbyte とデータの容量が大きくなってしまふことが難点で、さらにデータ数を圧縮するのが次の課題である。

参考文献

- 1) 齊藤兆古：MATHEMATICA による WAVELET 変換
- 2) 齊藤兆古：ウェーブレット変換の基礎と応用
- 3) 齊藤兆古：MATHEMATICA による画像処理入門
- 4) 齊藤兆古：ウェーブレット変換の基礎と応用
- 5) 新井康平、瀬戸要：固有値展開による情報の偏りを利用した多重解像度に基づくデータ・ハインディング
- 6) DCI-86503 NHK 効果音ライブラリー 動物と鳥