

2007 年度 修士論文

可視化動画像の固有情報抽出

と

その応用に関する研究

VISUALIZED DYNAMIC IMAGE CHARACTERISTICS EXTRACTION
AND
ITS APPLICATIONS

指導教授 齋藤 兆古

法政大学大学院
システムデザイン研究科
システムデザイン専攻 修士課程
学籍番号 06U0003

コウ ショウ
氏名 高 翔

Abstract

In order to realize a machine visual inspecting system, this paper tries to work out one of the dynamic image cognition methodologies. The method proposed in this paper employs Fourier cosinusoidal transform to extract the geometrical characteristics of the target. Thus an obtained geometrical characteristic is called the geometrical Eigen pattern.

The geometrical Eigen pattern of a frame image is extracted by taking the dominant major spectra from Fourier cosinusoidal transform of the frame image. Further , rearrangement of this geometrical Eigen pattern into a vector form gives the frame Eigen vector. Summation of entire frame Eigen vectors extracted from each of the frame images comprising a dynamic image gives a dynamic image Eigen vector.

To implement the dynamic image cognition, it is essential to set up the database of targets. A series of distinct dynamic image Eigen vector constitutes a system matrix. When an input dynamic image is given, a dynamic Eigen vector of the input dynamic image is can be obtained in much the same way as the database one, and then this Eigen vector becomes an input vector of the dynamic image cognition system of equations. Solving for this dynamic image cognition system of equations by least squares reveals a cognized dynamic image while the input test dynamic image should be registrant as a database. Uniqueness of solution element in the least square solution vector suggests a validity of the cognition result.

The method of dynamic image Eigen vector is applied to the color dynamic images as well as infrared dynamic images over 200 targets. Since color dynamic images are composed of three independent color components (red, green and blue) depending on wavelength of visual light, then the color dynamic Eigen vector has three times information compared with those of monochrome. Thereby, this leads that most of the color dynamic images are cognized by the proposed method exactly. On the other side , infrared dynamic images are composed of the thermal information which is a typical monochrome image and also easily contaminated by background thermal source. As a result, the infrared dynamic images are cognized by the proposed method having 90 % accuracy.

Thus, this paper has proposed the method of dynamic image cognition for machine visual inspecting system development, and revealed a fairly good result as a first step of development.

目次

Abstract	i
第1章 序論	1
第2章 理論	2
2.1 概要	2
2.2 フーリエ変換	2
2.2.1 関数系の変換	2
2.2.2 関数の直交性と線形性	4
2.2.3 離散値系フーリエ変換	5
2.2.4 一次元フーリエ変換の実際	7
2.2.5 偶関数と奇関数	8
2.3 フーリエ余弦変換による動画像の特徴抽出	9
2.3.1 画像の不変量抽出	9
2.3.2 幾何学的固有パターン	9
2.3.3 固有パターンの一致性	10
第3章 実験	12
3.1 画像認識	12
3.1.1 モノクロ静止画像認識	12
3.1.2 モノクロ動画像認識	13
3.1.3 カラー静止画像認識	15
3.1.4 カラー動画像認識	17
3.2 手話動画像識別	19
3.2.1 手話動画像収録	19
3.2.2 モノクロ手話動画像識別	20
3.2.3 カラー手話動画像識別	22
3.3 赤外線動画像識別	23
3.3.1 赤外線について	23
3.3.2 赤外線動画像識別	24
3.3.3 赤外線手話動画像識別	25
第4章 考察	27
第5章 結論	28
参考文献	30
研究業績	31
謝辞	32

第1章 序論

現代社会システムの高度情報化に伴い、音声認識や画像認識をはじめとするパターン識別技術の工業的な応用が注目されている。人間は極めて高度な画像認識能力を有している。例えば、近づいてくる小動物、それが犬であるのか、猫であるのか、あるいはその他の動物なのか、瞬時に判別できる。未知の小動物の特徴をつかみ、過去から現在まで記憶の中に蓄積された膨大なデータを参照し、一瞬にして対象を識別し、認識を完了する。この人間の視覚情報処理機能を計算機で代行せんとするマシンインテリジェンス研究において、画像認識・識別技術は重要な基幹技術の一つである[1]。

筆者の所属する研究室では、機械には機械に適した学習・判断能力があると考え、その第一段階として静・動的画像認識を可能とする人工眼球計画を発足し、研究・開発を行ってきた。筆者らの画像認識手法は、セキュリティ、工程管理システムなどに於ける機械監視、機械診断、機械識別、機械検査などを可能とする画像の固有パターンを用いるものである。固有パターンの一つは画像を構成する画素の濃淡分布情報に基づいている。このため、対象画像が変形しても対応可能な特徴を持つ反面、画像の幾何学的固有情報を失う問題点がある。本研究では、画像を構成する画素の空間的配置情報、すなわち、画像の幾何学的固有情報を用いた画像認識法に関して考察する。具体的には、考察対象である画像認識法を手話画像など動きが主体である動画像認識へ適用し、人工視覚情報処理機能の一応用技術としての妥当性と有用性を検証する。その結果、手話の計算機による全自動翻訳機開発の可能性が示唆され、人工視覚情報処理機能の一応用技術の実現可能性が検証される。さらに、赤外線は完全な暗闇でも画像情報を提供しており、可視光が存在しない環境下に於けるセキュリティなどへの適用性を検討するため、赤外線動画像識別をおこなう。その結果、考察対象とする画像認識法は赤外線画像へ対しても有効であることを検証する。

本論文では、デジタルビデオカメラを使って撮影した動画像からフーリエ余弦変換を用いて動画像の幾何学的特徴抽出法について述べ、次にテスト画像とデータベース画像の一致性評価を最小自乗法で行う画像認識法について述べ、その妥当性を具体的な手話動画像や赤外線画像へ適用し、検証する。すなわち、本論文では、画像を構成する画素の空間的配置情報を持つ画像の固有パターン抽出の原理を述べた後、具体的な検証例を通して、フーリエ余弦変換を用いた動画像の幾何学的特徴抽出法が優れた画像の幾何学的固有パターン抽出法であり、動画像識別・認識の有力な一方法であることを報告する。その結果、画像の幾何学的固有パターンを用いた手話識別では96%の識別率を得た。さらに、ほとんど暗闇である微弱な可視光下で得た赤外線動画像に対して90%の識別率を得た。

第2章 理論

2.1 概要

データを直交するデータ（線形独立，すなわち，互いに重複する情報を持たないデータ）の線形和へ並べ直す演算の代表に図 2.1 に示すようにフーリエ級数がある．フーリエ変換やフーリエ級数は関数が連続関数であるため，基準座標の選び方で原点に対して線対称か点对称かで偶関数か奇関数がそれぞれ決まる．このため，離散化された数値の並びで与えられる計算機中の一連の離散値データも原点に対して線対称か点对称かで偶関数か奇関数かそれぞれ仮定できる．図 2.1 はフーリエ級数の例を示すが，この波形をフーリエ級数へ展開すると，この波形は奇関数であるから，基底関数は奇関数である正弦波のみとなり，高調波次数は奇数次のみとなる．これをフーリエ正弦変換と呼ぶ．他方，図 2.1 の縦軸の原点を $x=0.5$ とすれば，この関数は偶関数となるため，この関数のフーリエ級数は偶関数である余弦波と次数は偶数次のみとなる．本研究では，与えられたデータの境界で偶関数の対称性を仮定してフーリエ変換を行う[2]．

ここでは，フーリエ級数から離散値系のフーリエ変換までの理論的背景を整理して述べる．

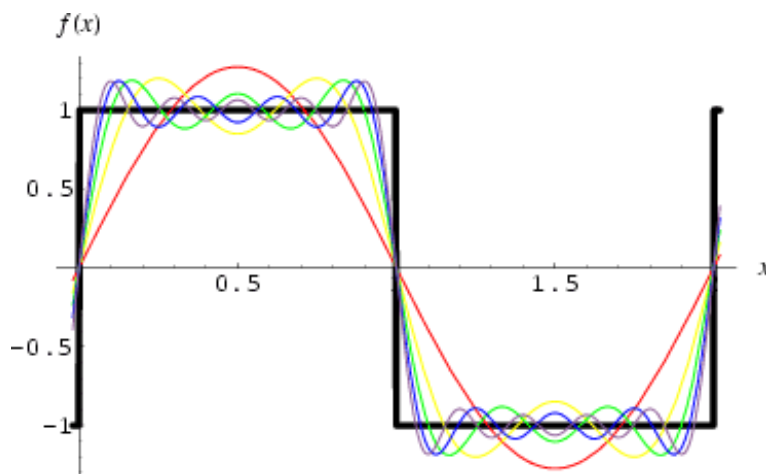


図 2.1 フーリエ級数の例

2.2 フーリエ変換

2.2.1 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)変換を取り上げる．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．

今，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する．すなわち， ω を角周波数として，

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin(i\omega t) + \sum_{i=1}^{\infty} b_i \cos(i\omega t) \quad (2.1)$$

と仮定する．

問題は式(2.1)の係数， a_0 ， a_i ， b_i を決める方法である．今，関数 $f(t)$ が $t=0$ から $t=T$ の区間で式(2.1)の係数を決めることを考えれば，式(1)は， $\omega = 2\pi/T$ であるから，

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i\frac{2\pi}{T}t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i\frac{2\pi}{T}t\right) \quad (2.2)$$

とも書ける．

式(2.2)の両辺を時間 $t=0$ から $t=T$ の区間について積分すると，

$$\begin{aligned} \int_0^T f(t)dt &= \int_0^T \left(a_0 + \sum_{i=1}^{\infty} a_i \sin\left(i\frac{2\pi}{T}t\right) + \sum_{i=1}^{\infty} b_i \cos\left(i\frac{2\pi}{T}t\right) \right) dt \\ &= a_0 T \end{aligned} \quad (2.3)$$

が成り立つ．

したがって，定数項 a_0 は

$$a_0 = \frac{1}{T} \int_0^T f(t)dt \quad (2.4)$$

となる．

次に，式(2.2)の両辺に正弦波 $\sin\left[j\left(\frac{2\pi}{T}t\right)\right]$ ， $j=1,2,3,\dots$ を掛け算し，時間 $t=0$ から $t=T$ の区間について積分する．

$$\int_0^T f(t) \sin(j \frac{2\pi}{T} t) dt = \int_0^T \left(a_0 + \sum_{i=1}^{\infty} a_i \sin(i \frac{2\pi}{T} t) + \sum_{i=1}^{\infty} b_i \cos(i \frac{2\pi}{T} t) \right) \times \sin(j \frac{2\pi}{T} t) dt$$

(2.5)

$$= a_i \frac{T}{2}, \quad i = j \quad \text{or} \quad 0, \quad i \neq j$$

式(2.5)から係数 a_i は

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

(2.6)

として得られる .

同様に , 式(2.2)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 ,

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

(2.7)

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 ,

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

(2.8)

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

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 .

2.2.2 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 .

まず式(2.4)の係数が計算される過程を考える . 式(2.3)は式(2.2)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 ,

$$\int_0^T 1 \cdot \sin(i \frac{2\pi}{T} t) dt = 0, \quad \int_0^T 1 \cdot \cos(i \frac{2\pi}{T} t) dt = 0, \quad i = 1, 2, 3, \dots \quad (2.9)$$

の関係が成り立つために、式(2.4)の係数 a_0 が計算できた。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、式(2.6)、(2.7)で計算される a_i 、 b_i は

$$\int_0^T \sin(i \frac{2\pi}{T} t) \sin(j \frac{2\pi}{T} t) dt = 0, \quad \int_0^T \cos(i \frac{2\pi}{T} t) \cos(j \frac{2\pi}{T} t) dt = 0, \quad i \neq j \quad (2.10)$$

なる直交性が成り立つことに基づいている。

また、式(2.9)、(2.10)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。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。式(2.6)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する。すなわち、

$$a_i = c_i + c_i + \dots = n c_i \quad (2.11)$$

である。式(2.11)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、線形性とは比例関係が成立する系である。

2.2.3 離散値系フーリエ変換

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、

$$F = [f(\Delta t), f(2\Delta t), f(3\Delta t), \dots, f(n\Delta t)]^T \quad (2.12)$$

なるベクトルとなる。

同様にして，式(2.1)を離散値系で書けば，

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

or

(2.13)

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

となる．ただし， \mathbf{I} は n 次の単位列ベクトルである．

また，式(2.13)右辺の係数 a_0 ，ベクトル \mathbf{S} と \mathbf{C} は，

$$a_0 = \frac{1}{n} \mathbf{I}^T \cdot \mathbf{F}, \quad \mathbf{S} = \frac{2}{n} \mathbf{A}^T \cdot \mathbf{F}, \quad \mathbf{C} = \frac{2}{n} \mathbf{B}^T \cdot \mathbf{F} \quad (2.14)$$

で与えられる．

ここで，離散値系でフーリエ係数を計算する過程で，式(2.13)を

$$\begin{aligned} \mathbf{F} &= a_0 \frac{1}{\sqrt{n}} \mathbf{I} + \sqrt{\frac{2}{n}} \mathbf{A}\mathbf{S} + \sqrt{\frac{2}{n}} \mathbf{B}\mathbf{C} \\ &= a_0 \mathbf{I}' + \mathbf{A}'\mathbf{S} + \mathbf{B}'\mathbf{C} \end{aligned} \quad (2.15)$$

と書き直すと,

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

$$\mathbf{A}'^T \cdot \mathbf{A}' = \mathbf{I}, \quad \mathbf{B}'^T \cdot \mathbf{B}' = \mathbf{I} \quad (2.16)$$

の関係が成り立つ。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、式(2.16)の係数行列 A' , 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。

2.2.4 一次元離散値系フーリエ変換の実際

簡単のため、式(2.1)で n 次の 1 次元ベクトル F をフーリエ変換することを試みる。

$$\mathbf{F} = [f(0), f(\Delta x), f(2\Delta x), \dots, f(\overline{n-1}\Delta x)] \quad (2.17)$$

変換行列は次式で与えられる。

$$\mathbf{C} = \frac{1}{\sqrt{n}} \begin{pmatrix} e^0 & e^0 & e^0 & \dots & e^0 \\ e^0 & e^{i\Delta x} & e^{i2\Delta x} & \dots & e^{i\overline{n-1}\Delta x} \\ e^0 & e^{i2\Delta x} & e^{i4\Delta x} & \dots & e^{i2\overline{n-1}\Delta x} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e^0 & e^{i\overline{n-1}\Delta x} & e^{i\overline{n-1}2\Delta x} & \dots & e^{i\overline{n-1}\overline{n-1}\Delta x} \end{pmatrix} \quad (2.18)$$

$$e^{i\alpha} = \cos(\alpha) + i\sin(\alpha), \quad i = \sqrt{-1}$$

式(2.17)のフーリエ変換は、式(2.18)の変換行列を使って式(2.19)のように行われる。

$$\mathbf{F}' = \mathbf{C}\mathbf{F}$$

$$= \frac{1}{\sqrt{n}} \begin{pmatrix} e^0 & e^0 & e^0 & \dots & e^0 \\ e^0 & e^{i\Delta x} & e^{i2\Delta x} & \dots & e^{i(n-1)\Delta x} \\ e^0 & e^{i2\Delta x} & e^{i4\Delta x} & \dots & e^{i2(n-1)\Delta x} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e^0 & e^{i(n-1)\Delta x} & e^{i(n-1)2\Delta x} & \dots & e^{i(n-1)(n-1)\Delta x} \end{pmatrix} \begin{pmatrix} f(0) \\ f(\Delta x) \\ f(2\Delta x) \\ \vdots \\ f(n-1\Delta x) \end{pmatrix} \quad (2.19)$$

フーリエ逆変換は，上添え字「*」が複素共役演算を示すこととして，

$$\mathbf{F} = (\mathbf{C}^T)^* \mathbf{F}' \quad (2.20)$$

で行われる．式(2.20)が成り立つためには，

$$\mathbf{C}^{-1} = (\mathbf{C}^T)^* \quad (2.21)$$

の条件が必要である．このように，もとの行列の転置行列の複素共役が逆行列となる行列をエルミート (Hermite) 行列と呼ぶ．

2.2.5 偶関数と奇関数

ここで強調して置かねばならないことは，奇関数性や偶関数性は本質的な図の性質でなく，座標系の軸に依存することである．たとえば，図2.2のように，もし線分AA'が垂直軸として選ばれたら，図2.2の関数は偶関数となり，そのフーリエ係数は余弦波の項だけを含むこととなる．これをフーリエ余弦変換という．

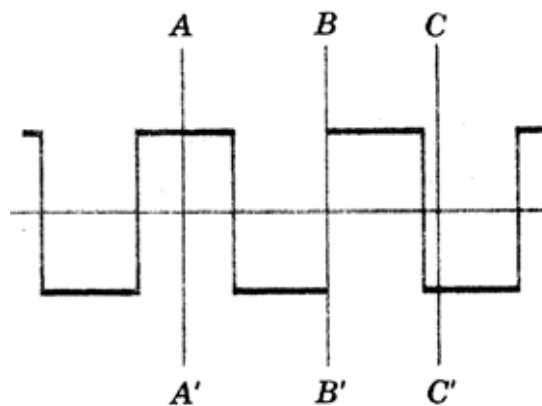


図 2.2 偶関数と奇関数

他方，もし線分 BB が垂直軸として選択されたならば，図 2.2 は奇関数を表し，フーリエ展開には正弦波の項しか出てこないであろう．

最後に，もし CC のような一般的な線分が垂直軸に選ばれたら，図 2.2 は奇関数や偶関数の何れでも無く，そのフーリエ級数には正弦波と余弦波の項が出てくるであろう．これをフーリエ正弦変換と言う．

2.3 フーリエ余弦変換による動画像の特徴抽出

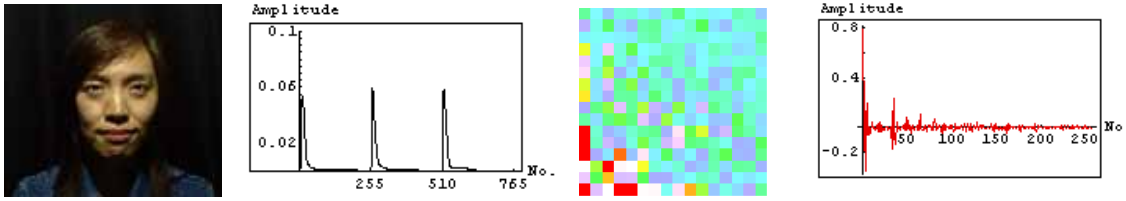
2.3.1 画像の不変量抽出

計算機のスクリーン上へ可視化された画像は 2 次元平面上の画素 (Pixel) で構成されており，画素の幾何学的配置と濃淡若しくは色情報によって表現される．すなわち，カラー画像の場合，各画素は可視光の波長によってそれぞれ赤，緑，青（以下それぞれ R, G, B と略記）成分の情報からなる．このため，各画像は固有の濃淡情報を持ち，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画像の不変量を本論文では画像の色彩もしくは濃淡固有パターン (RGB or Tone Eigen Pattern) と呼ぶ [3] [4]．

2.3.2 幾何学的固有パターン

フーリエ余弦変換を画像の空間周波数成分抽出に用いる．フーリエ余弦変換は，与えられた画像データを偶関数と見なし，画素数に等しいフーリエスペクトラム全てを取り扱うことを可能とする．ここでは，数値の整数化に伴う桁落ちを削減するため，空間周波数スペクトラムから直接 1 次元ベクトル化した固有パターンを採用する．

本論文で考える幾何学的情報とは点，線，面の繋がり情報である．点は画像を構成する最小限要素でゼロ次元の要素であり，線は点を連続して配置することで得られる 1 次元要素である．そして，面は線を隙間無く並べることで得られる 2 次元要素である．これらの周波数情報を比較すれば，線の空間周波数は面の空間周波数よりも高く，点の空間周波数は線の空間周波数よりも高いため，可視化情報の幾何学的に固有な特徴量が算出される．また，画像のフーリエ余弦変換は主要な空間周波数分布を原点近傍領域に集中することを勘案し，フーリエ余弦スペクトラム中で原点を含む 16×16 の空間周波数スペクトラムを 1 次元化し画像の幾何学的固有パターンとする．



(a) サンプル画像 (b) RGB 固有パターン (c) 周波数パワースペクトラム (d) 周波数固有パターン

図 2.3 固有パターンの例

図2.3はカラー画像の(a)サンプル画像とその(b) RGB固有パターン(c)サンプル画像の空間周波数余弦パワースペクトラムと，(d)原点を含む16x16 画素領域の周波数パワースペクトラムを切り出して作成された幾何学的固有パターンを示す[5][6]．

2.3.3 固有パターンの一貫性

本論文における画像認識は，予め複数の画像から固有パターンを抽出し，データベースを構築しておく．その後，認識処理対象として与えられる任意の入力画像から固有パターンを抽出し，データベース画像と入力画像間の固有パターンの一貫性を線形システム方程式の解から評価し，入力画像をデータベース画像のいずれかの画像と認識するものである．

データベースに n 個の固有パターンが得られているとすれば，システム行列 C は，

$$C = [E_1, E_2, E_3, \dots, E_n] \quad (2.22)$$

で与えられる．ここで，任意の入力画像の固有パターンを E_x とすると，式(2.23)の線形システム方程式が得られる．

$$E_x = C \cdot X \quad (2.23)$$

式(2.23)における X は，データベース画像の各固有パターンの重みを要素とする n 次のベクトルである．固有パターンの次数を m とすると， $n = m$ でない限りシステム行列 C は m 行 n 列の長方形列となるので，式(2.23)は不適切な線形システム方程式となる．本論文における固有パターンの次数 m は，カラー画像の場合 $m = 768$ ，モノクロ画像の場合 $m = 256$ となる．従って，カラー画像の場合はデータベース数 $n < 768$ ，モノクロ画像の場合はデータベース数 $n < 256$ であれば，解ベクトル X の算出に式(2.24)で示す最小自乗法を適用することができる[7]．

$$\mathbf{X} = [C^T C]^{-1} C^T \mathbf{E}_x \quad (2.24)$$

図 2.4 に示すように，式(2.24)で得られた解ベクトル \mathbf{X} の第 j 番目の要素が 1 で，他の要素が全て 0 であった場合を考える．解ベクトル中の横軸はデータベース番号と対応しているので，入力画像の固有パターン \mathbf{E}_x はデータベース画像の第 j 番目の固有パターン \mathbf{E}_j に等しい．よって入力画像をデータベースの第 j 番目の画像と認識できる．なお，本論文では得られた解ベクトル中で最大値をとる要素を認識された対象とする．

しかし，実際の問題では数値誤差やノイズが伴うため，必ずしも解ベクトル \mathbf{X} の第 j 番目の要素が 1 で他の要素が全て 0 とはならない．このため，本論文では得られた解ベクトル中で最大値をとる要素を認識された対象とする． $n > m$ となる不適切な線形システム方程式の解法として，最小ノルム解やベクトル型サンプルパターンマッチング法 (GVSPM) 等が知られている．

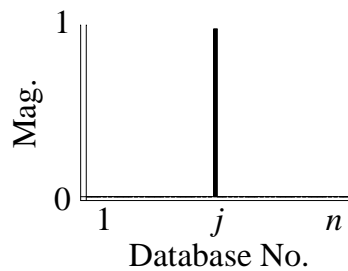


図 2.4 解ベクトルの例

第3章 実験

3.1 画像認識

3.1.1 モノクロ静止画像認識

本章ではモノクロ静止画像認識実験を行う．背景の色彩情報を可能な限り削減するために白熱電球の下，黒幕で構成された簡易暗室で撮影を行った．図 3.1 に，モノクロ供試画像に対する空間周波数パワースペクトラムと，幾何学的固有パターンを示す．図中のモノクロ画像とは反射光ベクトルの大きさ分布である[1]．

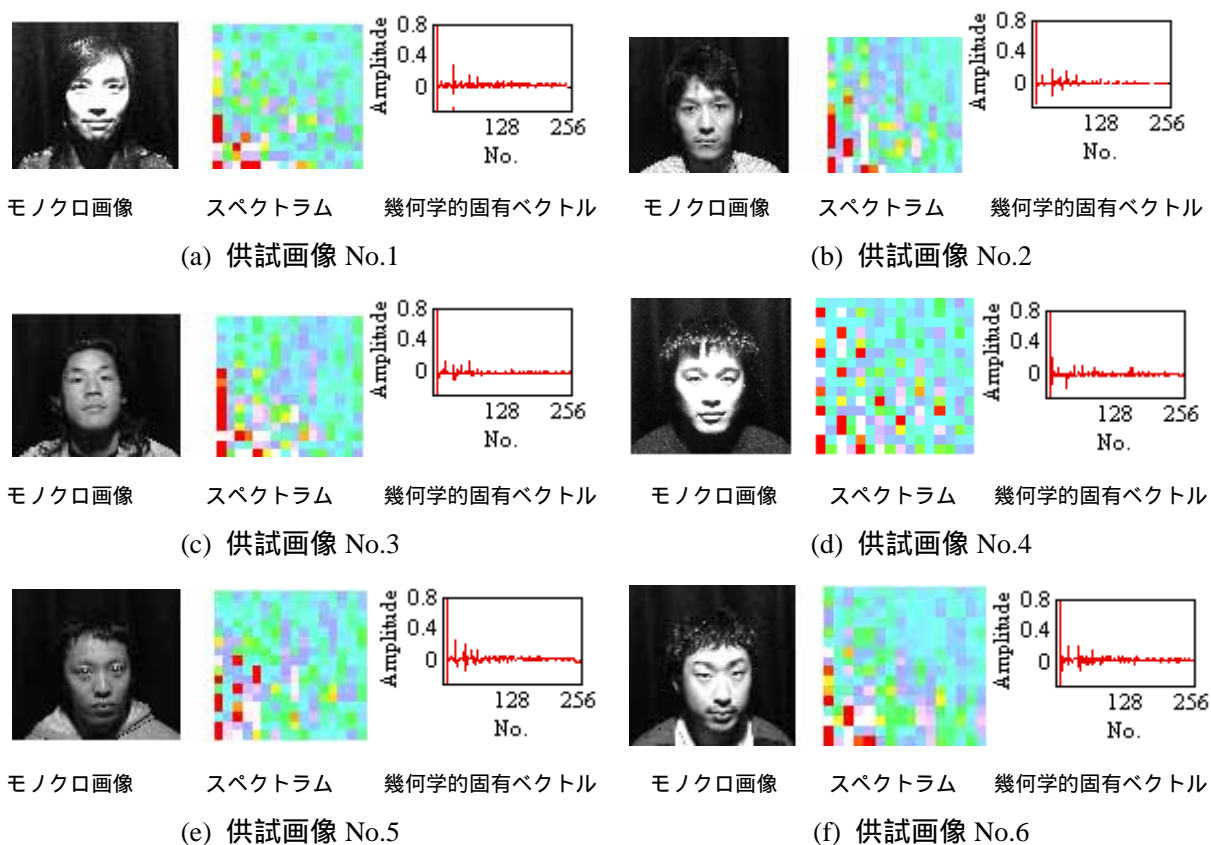


図 3.1 モノクロ画像と幾何学的固有パターン

本章で取り扱う幾何学的固有パターンは，フーリエ余弦変換のスペクトラムを切り出して作成されているため，データベースと入力間の再現性を伴うことが原則であり，それを十分に満たすよう簡易暗室を用いて撮影された画像をデータベース構築と供試画像とする．図 3.2 に認識結果を示す．

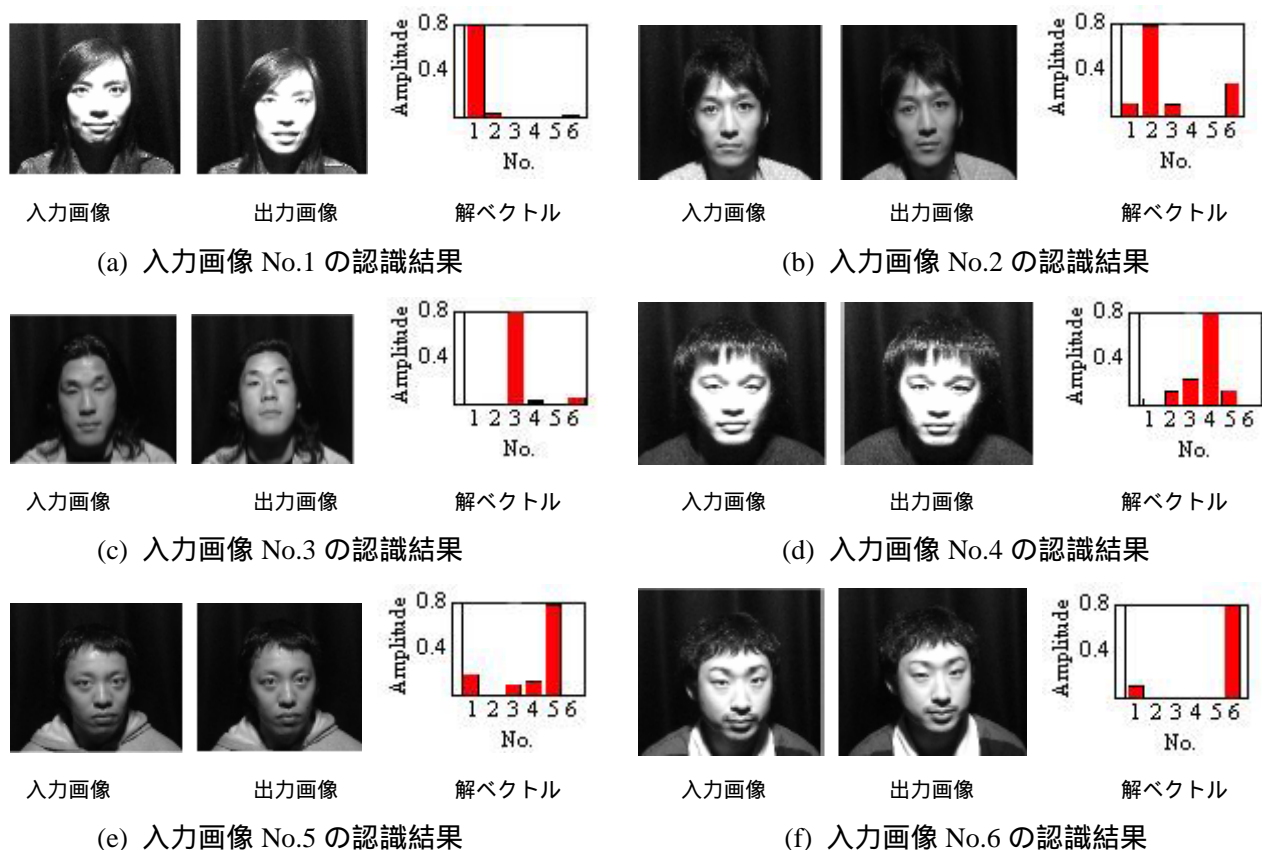


図 3.2 モノクロ静止画像の認識結果の例

3.1.2 モノクロ動画画像認識

空間周波数成分による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画像認識へ一般化する．動画画像は複数のフレーム画像データによって構成されるので，フレーム画像データ全体を通して固有パターンとなる特徴量が抽出される必要がある．単純なフーリエ変換と異なり，フーリエ余弦変換は画像中の対象物の空間位相情報に依存して異なるスペクトラムを与える性質がある．このため，式(3.1)に示すように各フレーム画像から空間周波数情報を算出し，全フレームのスペクトラム情報を時系列方向に加算したものを動画画像の固有パターンとして用いる．

$$\mathbf{E}_{\text{geo,mono}} = \sum_{i=1}^n (\text{frame spectrum})_i \quad (3.1)$$

図 3.3 にデータベース動画画像とフーリエ余弦スペクトラムと幾何学的固有パターンを示す．ここでは紙面の都合上動画画像を構成する全フレーム画像の中で 100 枚目のフレームを

示している．またスペクトラムは時間軸方向に加算されたものであり，原点の一定値を含む 16×16 領域を切り出したものであり，幾何学的固有パターンはそのスペクトラムを 1 次元に並べ直したものである．

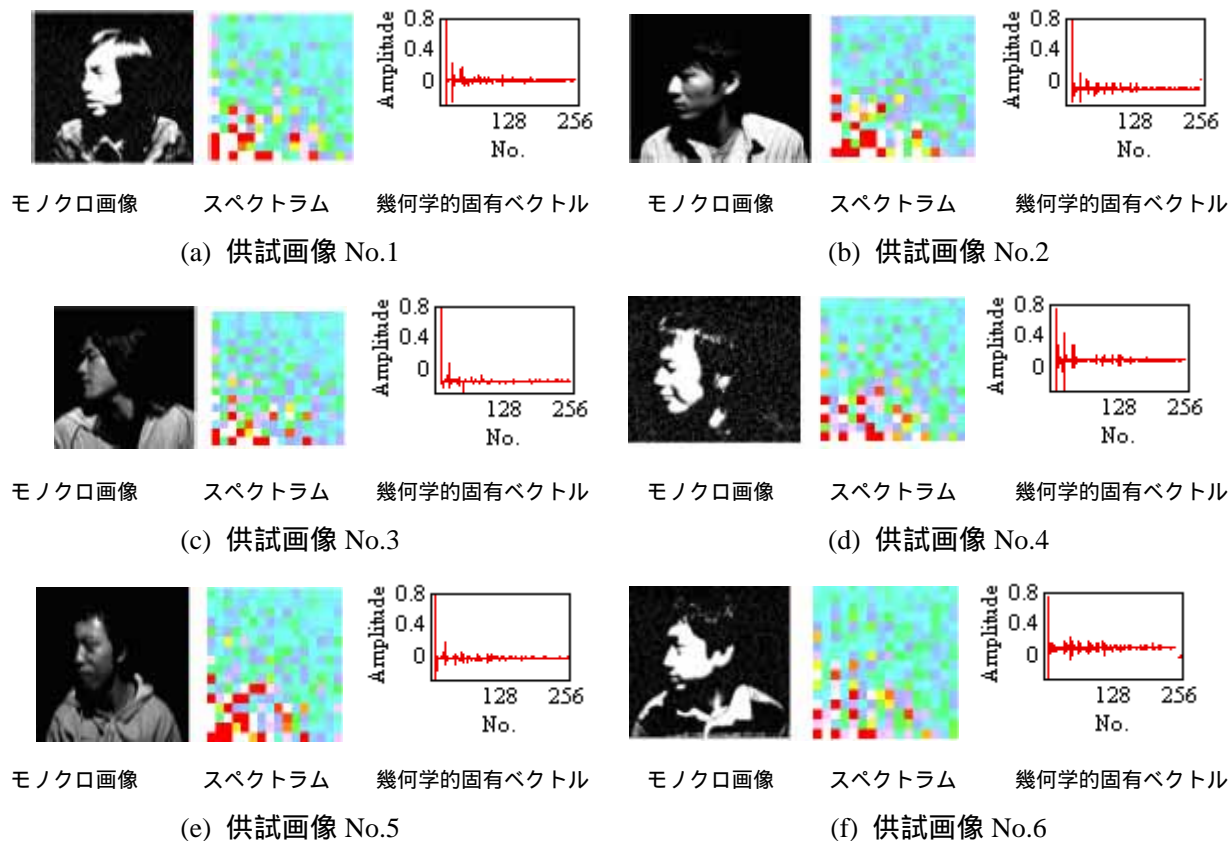


図 3.3 モノクロ動画像と幾何学的固有パターン

モノクロ動画像から算出された幾何学的固有パターンを用いて，動画像認識を行う．ここでは，画像中を人物が大きく移動するような動画像ではなく，カメラの前で固定されている動画像を用いる．換言すれば，幾何学的固有パターンはフーリエ余弦変換の 16×16 の低周波領域のみを取り扱うため，変化の大きな移動体よりも局所的動きを有する動画像は全体としての空間位相変化がないために有効な幾何学的固有パターンが得られる．図 3.4 に認識結果を示す．

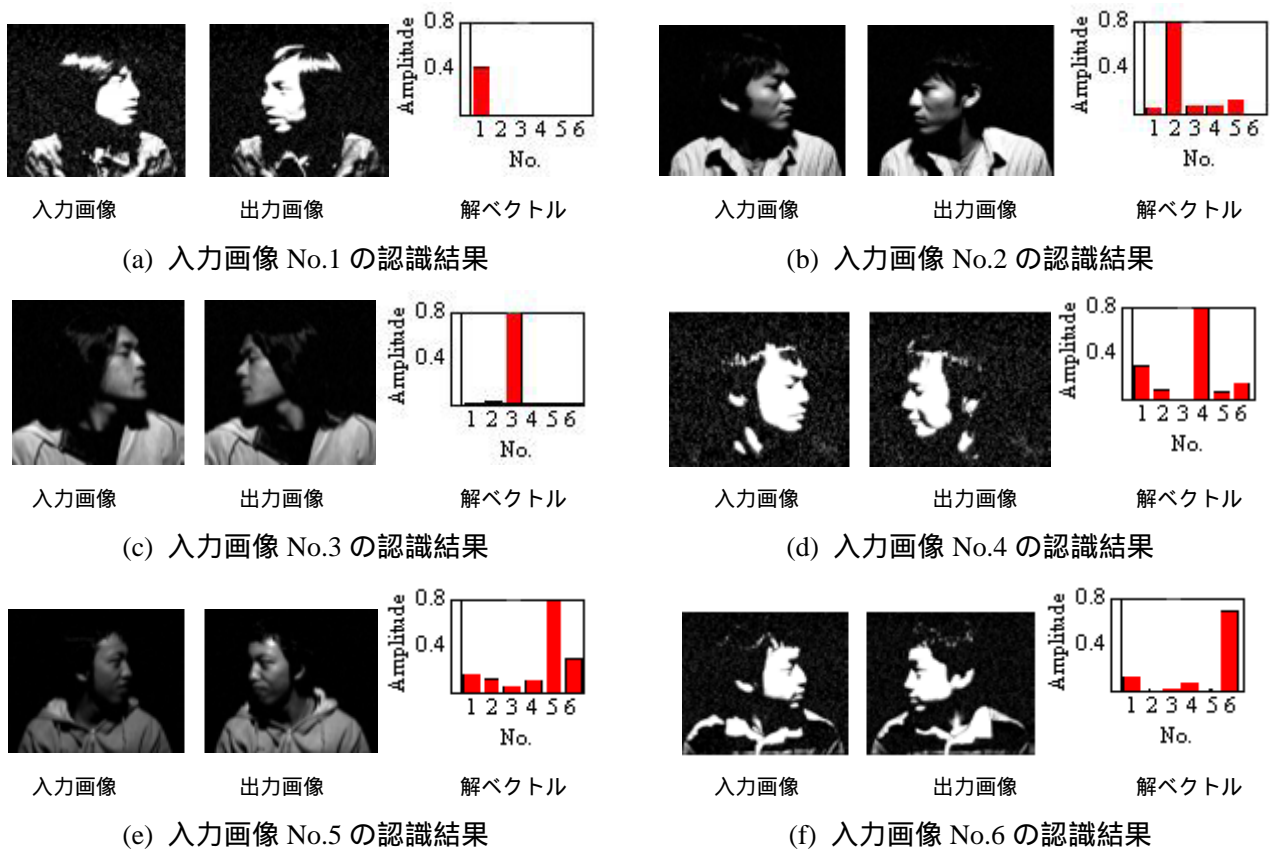


図 3.4 モノクロ動画画像の認識結果の例

3.1.3 カラー静止画像認識

3.1.1 節で記したモノクロ静止画像からの幾何学的固有パターンをカラー静止画像へ一般化する．カラー画像は R, G, B の 3 要素によって構成されるため，カラー画像の幾何学的固有パターンは各成分をフーリエ余弦変換し，それぞれのスペクトラムから 16×16 領域を 1 次元に結合したものより算出される．カラー画像の幾何学的固有パターン \mathbf{E}_{geo} は R 成分ベクトル $\mathbf{E}_{\text{geo}, R}$, G 成分ベクトル $\mathbf{E}_{\text{geo}, G}$, B 成分ベクトル $\mathbf{E}_{\text{geo}, B}$ によって構成され，式(3.2)で与えられる．式(3.2)の上添え字 T は転置を表す．

$$\mathbf{E}_{\text{geo,color}} = [\mathbf{E}_{\text{geo}, R}, \mathbf{E}_{\text{geo}, G}, \mathbf{E}_{\text{geo}, B}]^T \quad (3.2)$$

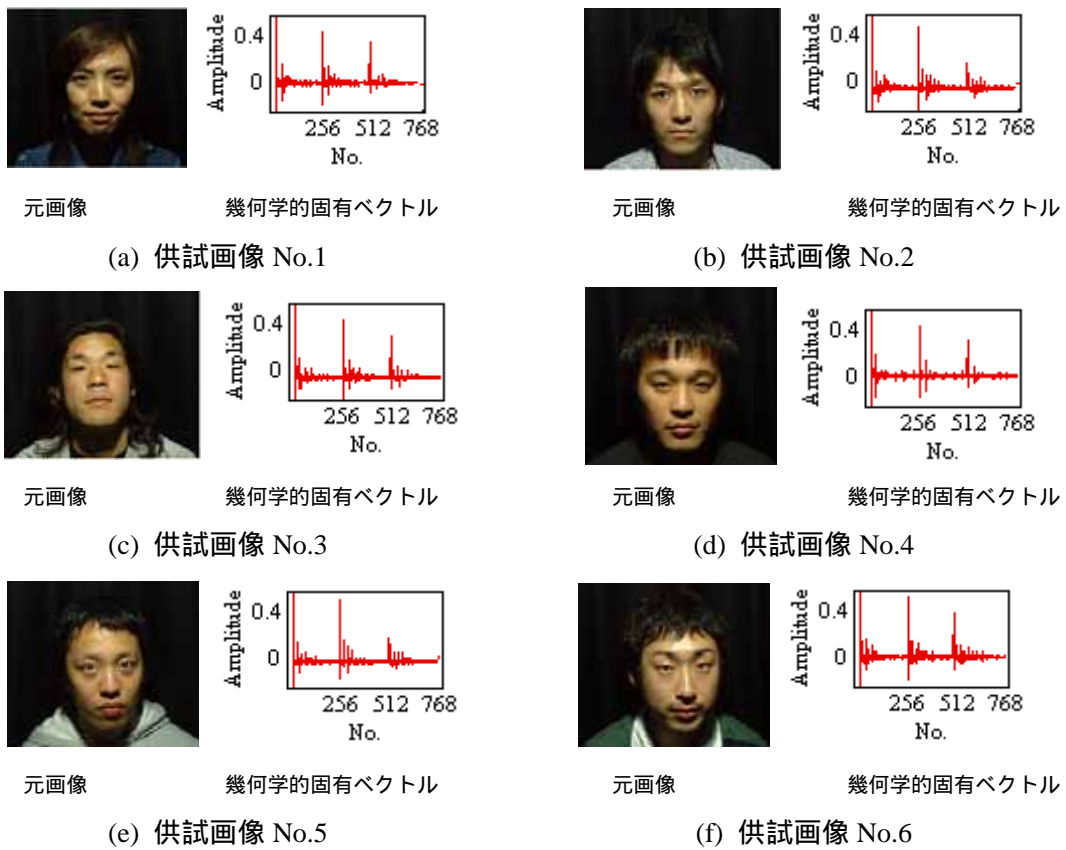


図 3.5 カラー画像と幾何学的固有パターン

図 3.5 に示したカラー画像の幾何学的固有パターンを用いて画像認識を行い，図 3.6 に認識結果を示す．3.1.1 節で記した，モノクロ画像の幾何学的固有パターンと比較して色情報の意味も含んだ幾何学的固有パターンとして認識が行われる．このため，モノクロ画像認識結果よりも精度の高い解を得られる．



(a) 入力画像 No.1 の認識結果



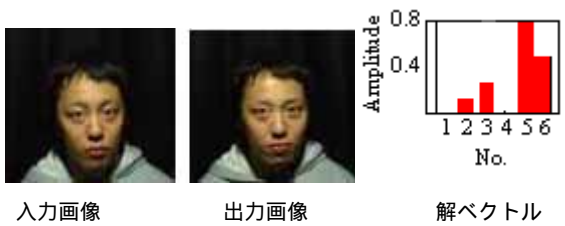
(b) 入力画像 No.2 の認識結果



(c) 入力画像 No.3 の認識結果



(d) 入力画像 No.4 の認識結果



(e) 入力画像 No.5 の認識結果

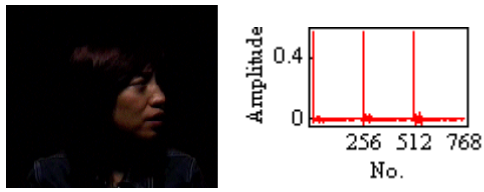


(f) 入力画像 No.6 の認識結果

図 3.6 カラー静止画像の認識結果の例

3.1.4 カラー動画認識

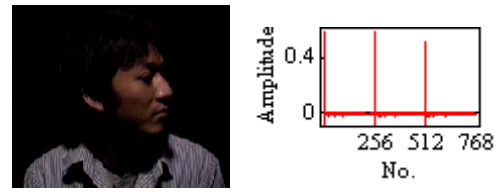
3.1.3 節で記したカラー静止画像からの幾何学的固有パターンをカラー動画へと拡張する。カラー動画は R, G, B 成分で時系列に描かれているものであるから、フレーム画像の各成分でフーリエ変換し、スペクトラムを時間軸方向に加算したものを幾何学的固有パターンとする。したがって、カラー動画の幾何学的固有パターンはモノクロ動画の要素の 3 倍となる。図 3.7 にカラー動画と幾何学的固有パターンの例を示す。図 3.8 はカラー動画の幾何学的固有パターンの認識結果の例を示す。



元動画像

幾何学的固有ベクトル

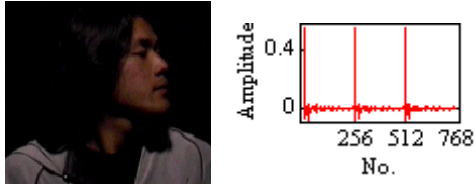
(a) 供試画像 No.1



元動画像

幾何学的固有ベクトル

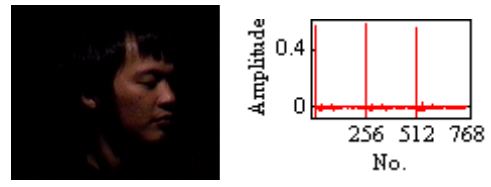
(b) 供試画像 No.2



元動画像

幾何学的固有ベクトル

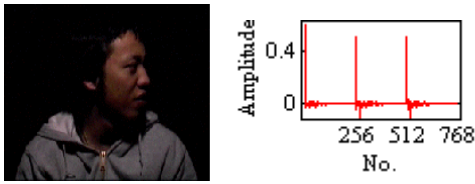
(c) 供試画像 No.3



元動画像

幾何学的固有ベクトル

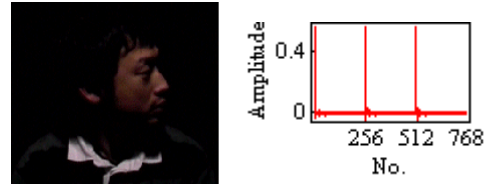
(d) 供試画像 No.4



元動画像

幾何学的固有ベクトル

(e) 供試画像 No.5



元動画像

幾何学的固有ベクトル

(f) 供試画像 No.6

図 3.7 カラー動画像と幾何学的固有パターン

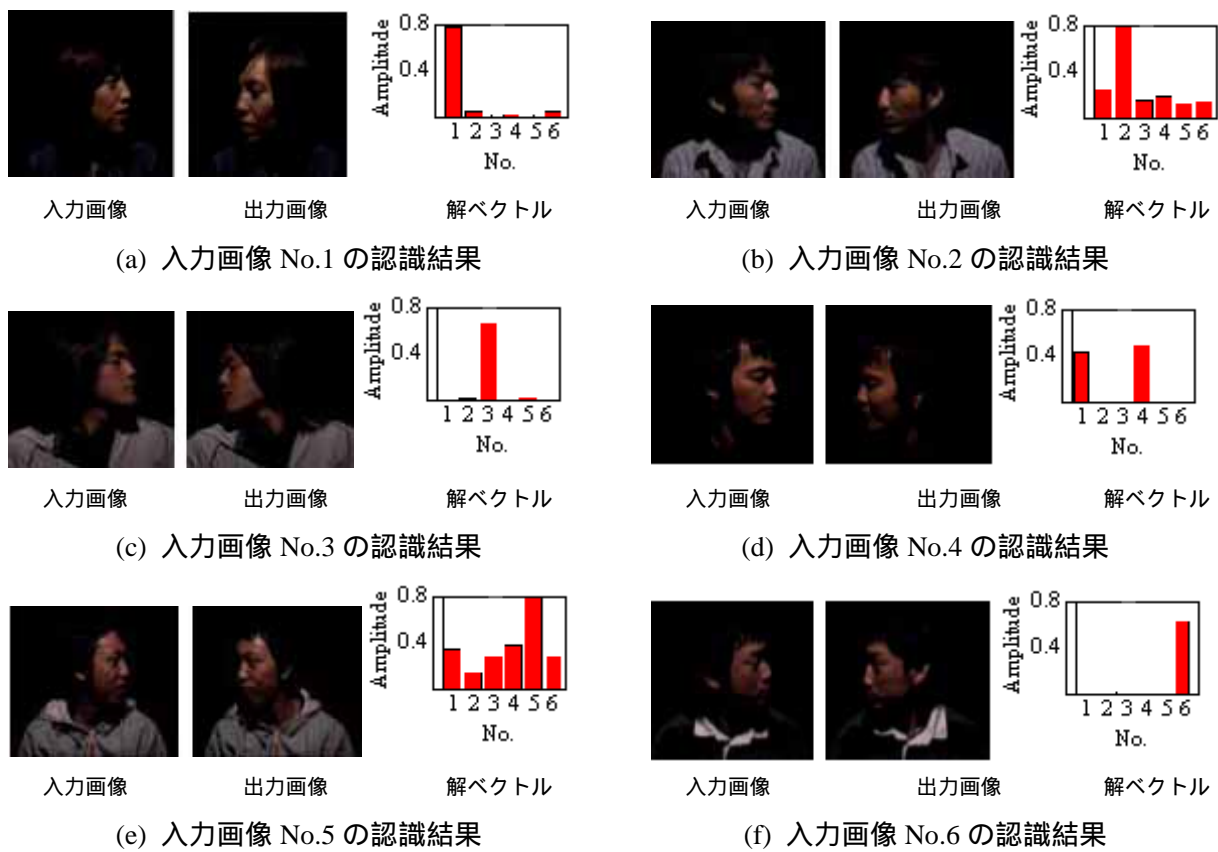


図 3.8 カラー動画画像の認識結果の例

3.2 手話動画画像識別

3.2.1 手話動画画像の収録

本節は提案する方法をモノクロ、カラー動画画像の手話動画画像へ適用し、手話識別を行なう。5人が20通りと他の5人が9通りの手話（全体で145通り）を行っている動画画像を収録した。図3.9は5人が9通りのデータベースとして収録した手話動画画像例のフレーム画像を示す[8]。

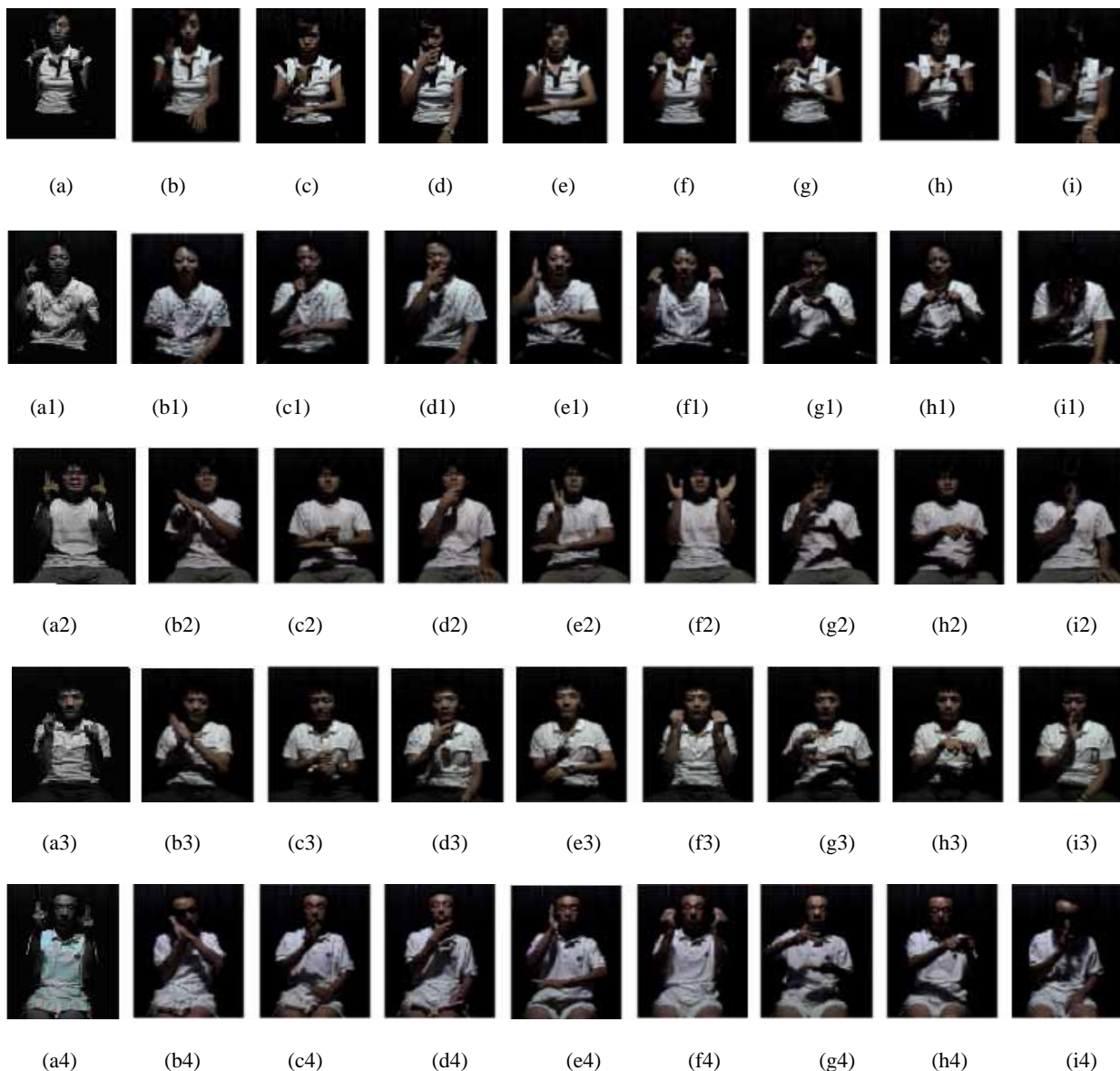


図 3.9 手話動画像の例

3.2.2 モノクロ手話動画像識別

カラー画像とモノクロ画像では画像を構成する色彩情報が異なるため、当然であるが識別結果も異なることが想定される。これを確認するため、ここではカラー収録動画像からモノクロ動画像を生成し識別を行った。

図 3.10 に 3.2.1 節の手話動画像の内 4 人のモノクロデータベース動画像の 1 フレーム、および幾何学的固有パターンを示す。ここでは紙面の都合上動画像を構成する全フレーム画像の中で 80 枚目のフレームを示している。図 3.11 はモノクロ手話動画像の幾何学的固

有パターンを使った認識結果で識別に成功した1例である。図3.12はモノクロ手話動画の幾何学的固有パターンを使った認識結果で識別に失敗した1例である。

全体を纏めると145手話モノクロ動画に対して135通りの手話が正確に識別され、10通りの手話の識別に失敗し、全体として識別率は約93%となった。

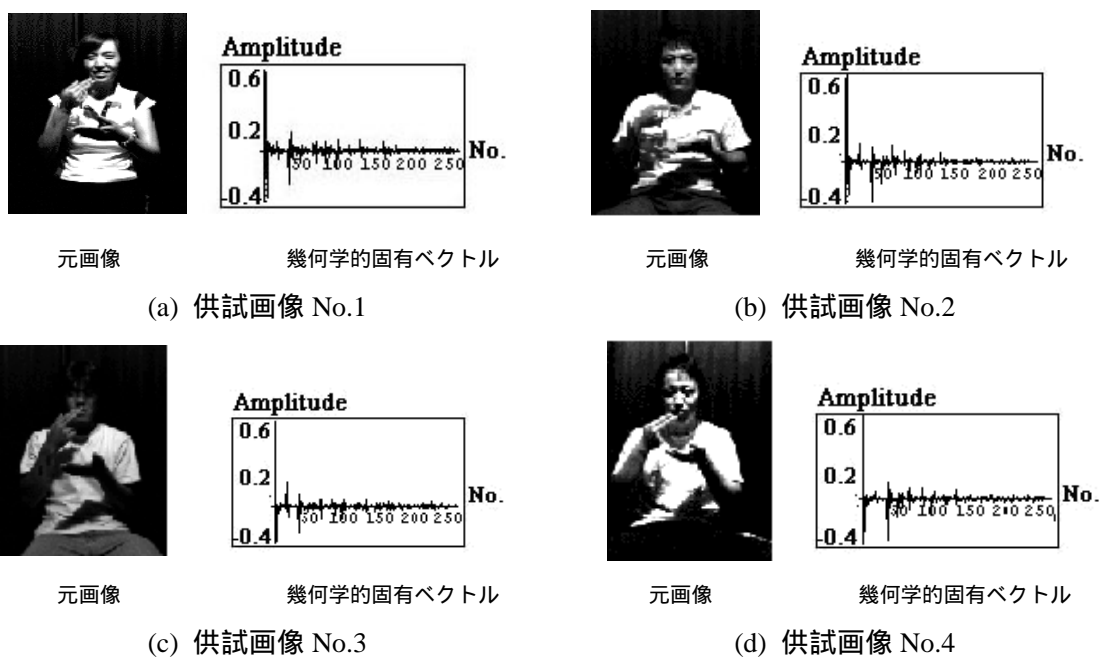


図 3.10 モノクロ手話動画と幾何学的固有パターン

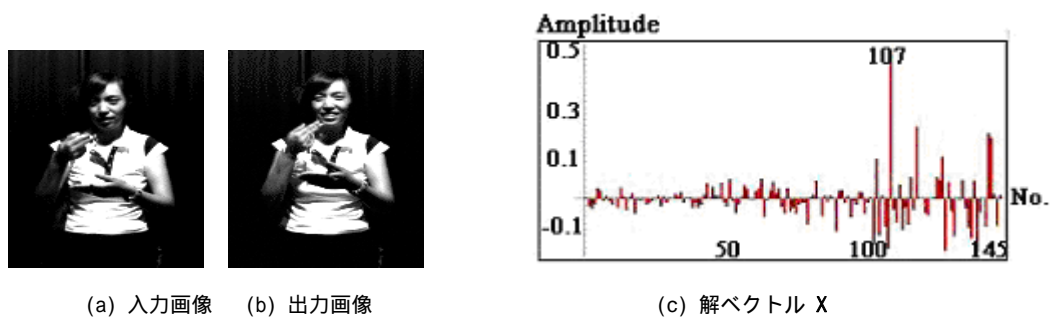


図 3.11 モノクロ手話動画認識成功の例

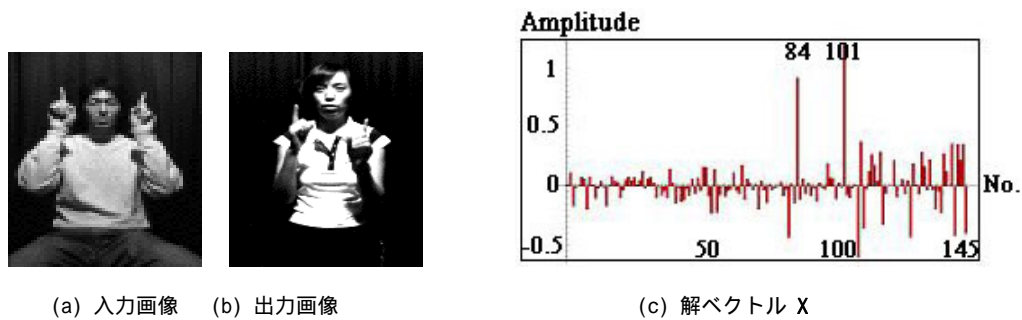


図 3.12 モノクロ手話動画認識失敗の例

3.2.3 カラー手話動画認識

カラー動画は R, G, B 成分で構成されたフレーム画像を時系列に並べて描かれるので, 各フレーム画像の各成分をフーリエ変換し, 得られたそれぞれのスペクトラムを時間軸方向に加算して幾何学的固有パターンとする. カラー動画は R, G, B 成分で構成されるため, カラー動画の幾何学的固有パターンの情報量 (要素数) はモノクロ動画の要素の 3 倍となる.

図 3.13 にカラー手話動画の 1 フレームと幾何学的固有パターンの例を 4 個示す. 図 3.14 はカラー手話動画の幾何学的固有パターンを使った認識結果で識別に成功した 1 例である. 図 3.15 はカラー手話動画の幾何学的固有パターンを使った認識結果で識別に失敗した 1 例である.

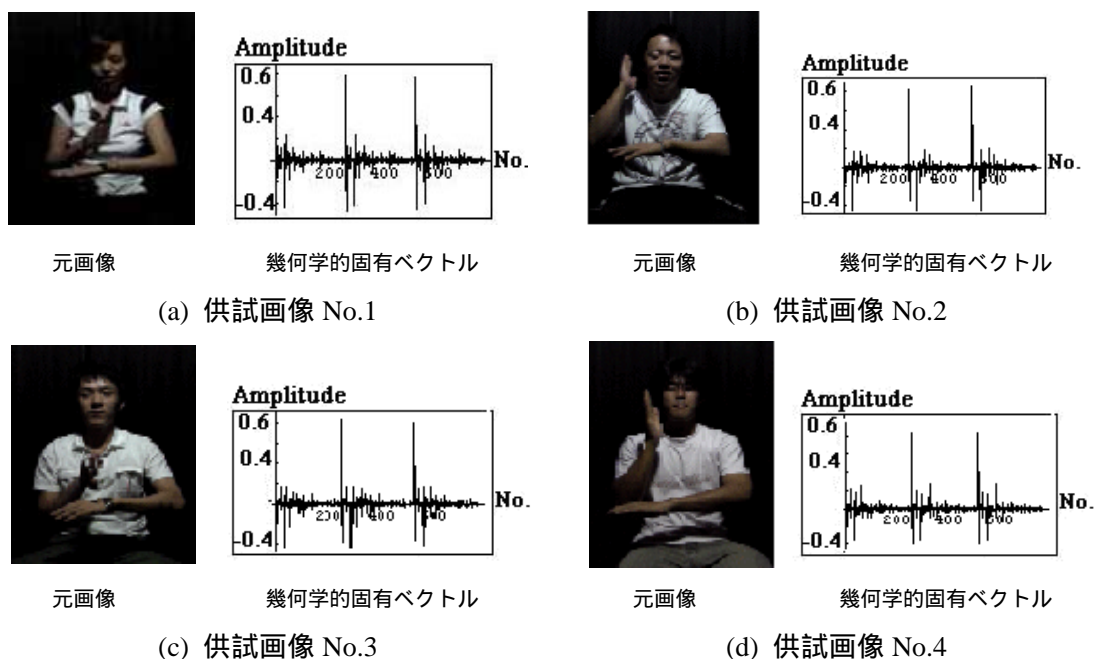


図 3.13 カラー手話動画と幾何学的固有パターン

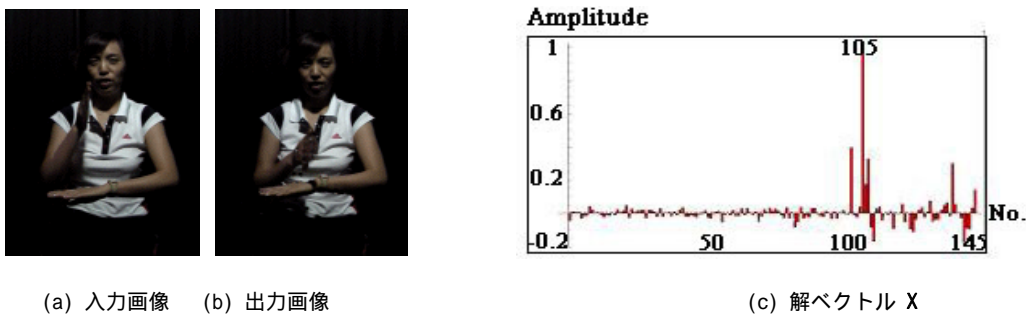


図 3.14 カラー手話動画像認識成功の例

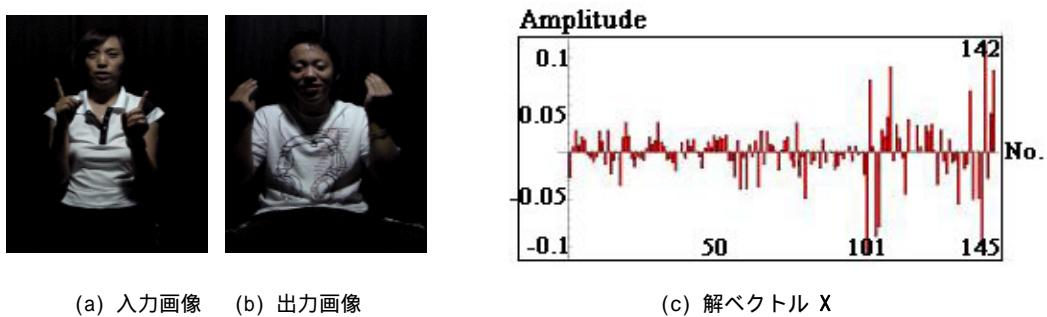


図 3.15 カラー手話動画像認識失敗の例

全体を纏めると 145 手話カラー動画像に対して 139 通りの手話が正確に識別され、6 通りの手話の識別に失敗し、全体として識別率は約 96% となり、モノクロ画像に対しカラー情報の増加は確実に識別率の向上に繋がることが確認された。9 人が行う 145 通りの手話動画像へ提案した方法で識別を行った結果、モノクロ手話動画像で約 93%、カラー手話動画像で約 96% の識別率を得た。

3.3 赤外線動画像識別

3.3.1 赤外線について

赤外線は可視光外の長波長領域に分布し、すべての温度エネルギーを有する物体から放射される光である。図 3.16 は赤外線の波長領域を示す模式図である。一般に、赤外線は目に見えないため、可視光の明るさ、暗さに依存しない。また、物体が放射する光の全エネルギー、波長と物体の温度の間には相関関係があり、温度が高い物体は、放射される赤外線の全エネルギーが多く、波長が短い赤外線を多く放射する。従って、物体から放射される特定波長の赤外線の強度を測定することによって物体の温度が測定できる [9]。

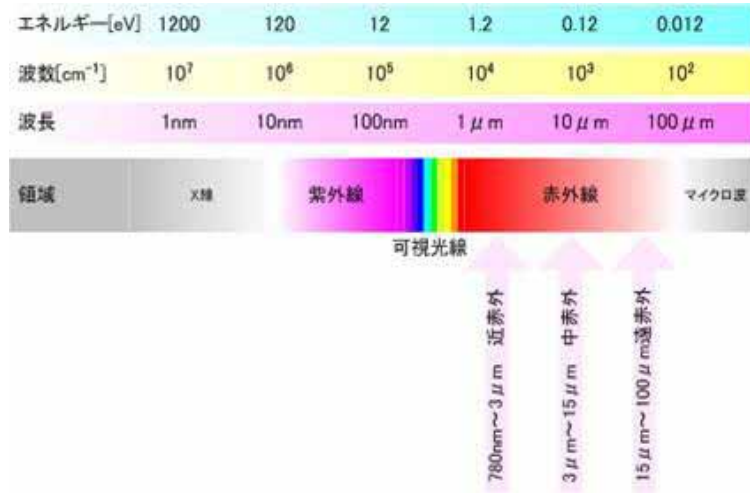
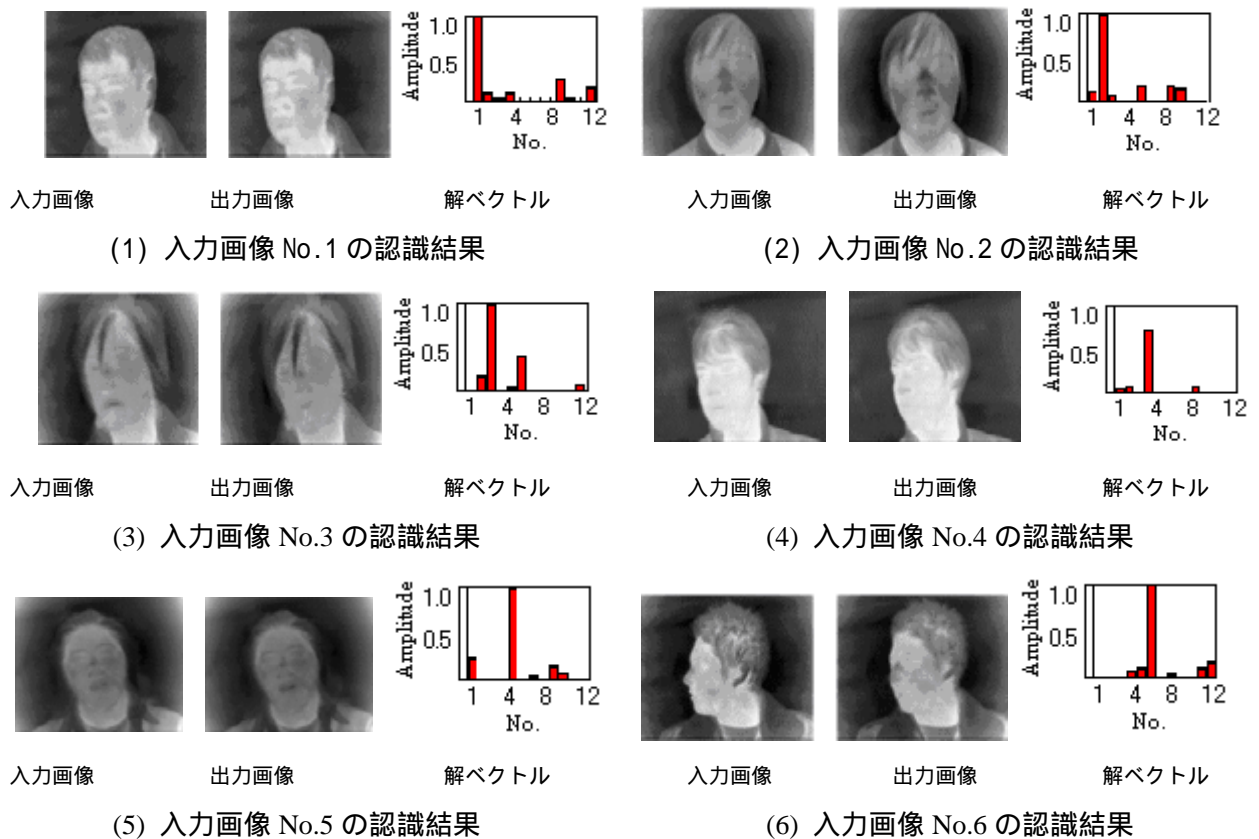


図 3.16 赤外線の波長領域を示す模式図

3.3.2 赤外線動画の識別

通常，人の体温は摂氏 36 度程度であり，赤外線カメラによって人の検知が可能である。本節では，理論の検証実験として，まず，モノクロ動画と同様にして，単純な規則性を持つ赤外線動画の識別実験を行った。図 3.17 は認識結果を示す。



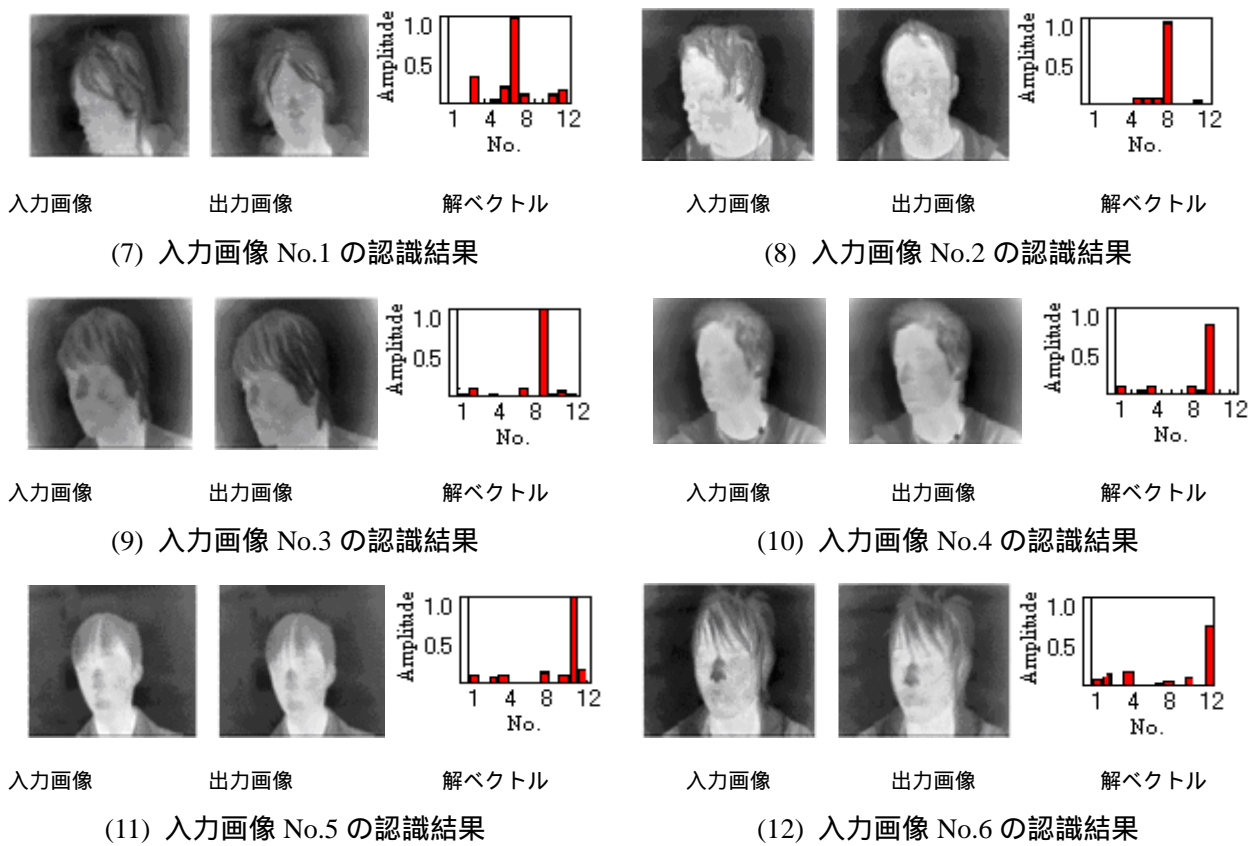


図 3.17 赤外線動画画像認識結果の例

図 3.17 にそれぞれテスト画像と識別されたデータベース中の画像とベクトル X の要素の大きさを示す. 図 3.17 の最初の結果は, 係数行列 C は 12 個の幾何学的固有パターンからなり, 入力ベクトル E_x と最も一致する固有ベクトルは式(2.22)の係数行列 C の第 1 列であることを意味する. 赤外線動画画像の動きに制限が伴うが, ほぼ 100% 近く識別された.

3.3.3 赤外線手話動画画像の識別

2.3.3 節で述べた動画画像認識の一致度評価を赤外線動画画像手話に適用し, 規則的であるがより複雑な動作の認識を試みる. 常温のもとで 7 人が 10 通りと暖房環境のもとで 6 人が 10 通り (全体で 130 通り) を収録した. 図 3.17 は赤外線手話動画画像の幾何学的固有パターンを使った認識結果で, 図 3.18 は識別に成功した例である. 図 3.19 は識別に失敗した例を示す. 全体で 130 個の赤外線手話動画画像に対して 106 通りの手話が正確に識別され, 24 通りの手話の識別に失敗し, 全体として識別率は約 82% となった. これは, 暖房環境下において, 部屋の温度が上昇し, 識別対象である人間の熱情報が削減されるためである. 換言すれば, 赤外線はエネルギーを有する対象が生ずる熱分布を可視化するため, 他の熱源の存

在が識別率を低下するのは当然であることを意味する．部屋に熱源が無い環境下では 90% 以上の識別率を得た．

以上の結果から，本手法は熱源を対象物のみが有する場合良好な識別率を得ることが可能である．

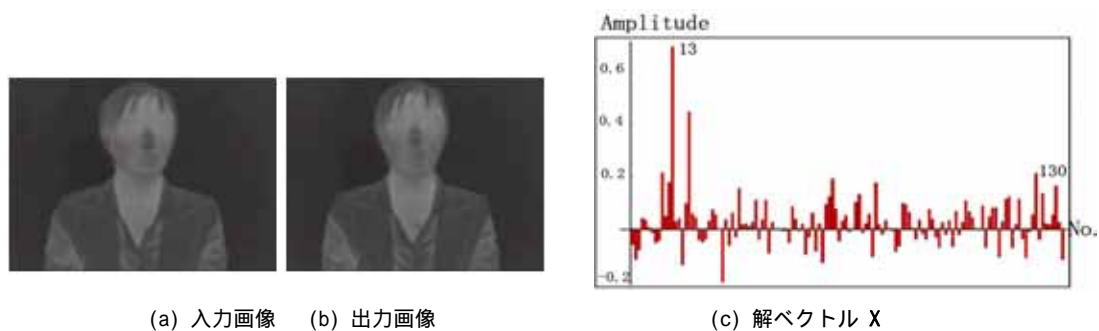


図 3.18 赤外線動画画像認識成功の例

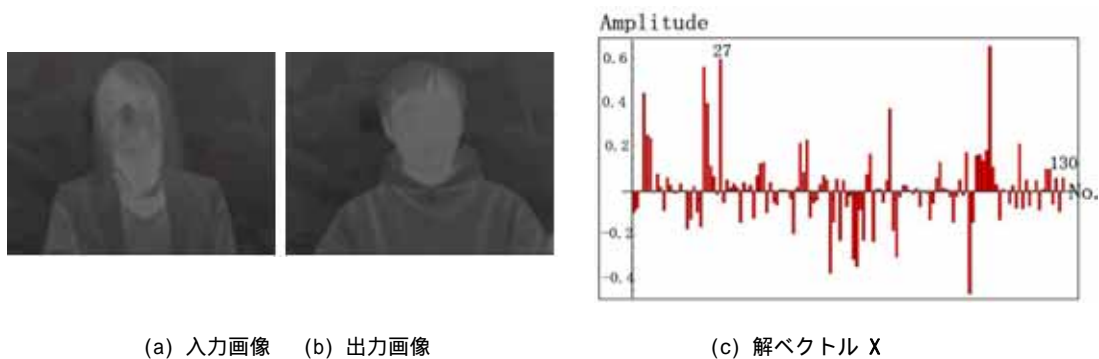


図 3.19 赤外線動画画像認識失敗の例

第4章 考察

図 3.4 と図 3.8 に示すそれぞれモノクロ動画像とカラー動画像に関する認識結果では、動画像は 100%認識され、信頼性の高い結果が期待できることが確認された。この結果により、フーリエ余弦変換を用いた画像認識技術の妥当性が検証されたと考えられる。画像認識技術が確立されれば、人間の視覚情報処理機能が機械により実現でき、工場ラインの検査等、いわゆる機械監視システム等の無数のアプリケーション、すなわち、セキュリティや工程管理システムなどにおける機械による監視・診断・識別・検査の完全な自動化が可能となる。

次に、本手法を手話識別へ応用した場合について述べる。図 3.11 と図 3.14 に示すように同一人物が行う同一手話がそれぞれに識別されるか否かを検討した。9 人が行う 145 通りのモノクロ手話動画像で 93%、カラー手話動画像で約 96%の識別率を得た。以上の結果により、手話の計算機による全自動翻訳機開発の可能性が示唆された。これは、聴覚障害者が日常生活の中で健聴者と情報のやりとりを自動化することが出来ることを意味し、人工視覚情報処理機能の一応用技術として手話自動識別機開発の可能性が明らかにされた。

最後に、赤外線カメラを使った赤外線動画像識別に関する考察を述べる。100%の識別率が得られた単純な規則性を持つ赤外線動画像の識別結果を図 3.17 は示す。図 3.18 に示す赤外線手話動画像認識結果で、識別対象以外に熱源が存在する 130 ケースの識別では 81%の識別率が得られている。他方、識別対象以外に熱源が存在しない 78 ケースの識別では 90%以上の識別率が得られた。

以上の結果から、赤外線動画像を使った以下の応用が考えられる。1)赤外線を使った状態監視システム、特に可動機器の熱分布のダイナミックな変化を利用した事故・破損予知を究極の目的とした機械監視システム構築の実現。2)温度較正された赤外線画像や熱画像を作成するサーモグラフィ技術は多様な分野での応用が可能であり、この赤外線応用技術に付加機能を可能とする。例えば、熱源となる物体や生物の存在を検知することができる性質を利用して、森の中での動物の探知、暗闇での動物の種類や固体別識別など、主として動物の管理・観察分野での活用が期待できる。さらに使い方を工夫することによって人命救助などへの活用も考えられる。

第5章 結論

社会の高度情報化に伴い、音声認識や画像認識をはじめとするパターン識別技術の工業的な応用が注目されている。人間は極めて高度な画像認識能力を有している。例えば、近づいてくる小動物、それが犬であるのか猫であるのか、あるいはその他の動物なのか、瞬時に判別できる。未知の小動物の特徴をつかみ、過去から現在まで記憶の中に蓄積された膨大なデータを参照し、一瞬にして対象を識別し、認識・識別を完了する。この人間の視覚情報処理機能を計算機で代行せんとするマシンインテリジェンス研究において、画像認識・識別技術は重要な基幹技術の一つである。

マシンインテリジェンス開発の第一段階として静・動的画像認識・識別が必須である。本論文で検討している画像認識・識別方法は、セキュリティ、工程管理システムなどに於ける機械監視、機械診断、機械識別、機械検査などを可能とする画像の固有パターンを用いるものである。固有パターンの一つは画像を構成する画素の濃淡分布情報に基づいている。このため、対象画像が変形しても対応可能な特徴を持つ反面、画像の幾何学的固有情報を失う問題点がある。本研究では、画像を構成する画素の空間的配置情報、すなわち、画像の幾何学的固有情報を用いた画像認識法に関して考察を行った。具体的には、考察対象である幾何学的固有情報を用いた画像認識法を手話画像など動きが主体である動画像認識へ適用し、人工視覚情報処理機能の一応用技術としての妥当性と有用性を検証した。その結果、手話の計算機による全自動翻訳機開発の可能性が示唆され、人工視覚情報処理機能の一応用技術の実現可能性が検証された。さらに、赤外線は完全な暗闇で画像情報を獲得可能とする。このため、可視光が存在しない環境下に於ける赤外線動画像識別をおこなった。その結果、考察対象とする画像認識法は赤外線画像へ対しても有効であることを検証した。

以下、本論文で得られた成果を中心として章別の内容を要約する。

第2章では、画像認識の理論について述べた。まず、フーリエ変換とフーリエ余弦変換について述べ、具体的には関数系の変換・関数の直交性と線形性・離散値系フーリエ変換・偶関数と奇関数に関して述べた。また、フーリエ余弦変換による幾何学的固有パターンを抽出する手法を述べ、テスト画像とデータベース画像のパターン一致性を最小自乗法で評価する手法を述べた。

第3章では、本論文で提案した手法を検証する実験について述べた。先ず、モノクロ・カラー静止画像、動画像認識を試みた結果、動画像の動きに制限が伴うが、比較的良好な結果が得られることを明らかにした。

次に、モノクロ、カラー、それぞれの手話動画像別を試みた。9人が行う145通りの手話動画像へ提案手法を適用した結果、モノクロ手話動画像で約93%、カラー手話動画像で約96%の識別率が得られ、良好な精度で動画像識別が可能であることを示した。

最後に、暗闇環境中で誰かどのような動作をしたかの識別可能かを検討するため、赤外

線動画像認識をおこない、その結果、環境下に識別対象以外に熱源が存在しない場合、90%以上の高い識別率が得られることを検証した。

以上、フーリエ余弦変換を用いた画像の幾何学的周波数情報が静止画像のみならず動画に対しても極めて有効な画像固有の情報を提供することを検証し、マシンインテリジェンス研究において、画像認識・識別技術は重要な基幹技術となりえる可能性を明らかにした。

参考文献

- [1] 高翔, 小杉山格, 齋藤兆古, 堀井清之, "フーリエ余弦変換による動画像の特徴抽出", 可視化情報学会誌 Vol.26, No.1 (2006) pp.141-144.
- [2] 齋藤兆古: 場の理論と可視化, 研究室内部資料, 2006.
- [3] 佐藤隆紀, 早野誠治, 齋藤兆古, 堀井清之, "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.
- [4] 丸山和夫, 早野誠治, 齋藤兆古, 堀井清之, "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.
- [5] 小杉山格, 早野誠治, 齋藤兆古, 堀井清之, "可視化画像の幾何学的複雑さ抽出の試み", 可視化情報学会誌, Vol.25, No.1(2005)pp.67-70.
- [6] 小杉山格, 早野誠治, 齋藤兆古, 堀井清之, "画像の固有パターンに関する一考察", 可視化情報学会誌, Vol.25, No.2(新潟 2005)pp.135-136.
- [7] G.Strang, "Linear Algebra and its Applications", Academic Press, 1976 年
- [8] 高翔, 青木亮一, 齋藤兆古, "画像の幾何学的固有パターンを用いた手話識別", 第十五回 MAGDA コンファレンス, (2006) pp.66-70.
- [9] 高翔, 小杉山格, 齋藤兆古, 堀井清之, "赤外線動画像の識別", 可視化情報学会誌, Vol.27, No.1 (2007) pp.225-226.

研究業績

- [1] 可視化情報学会第34回可視化情報シンポジウム
高翔, 小杉山格, 齋藤兆古, 堀井清之, “フーリエ余弦変換による動画像の特徴抽出”
- [2] 法政大学情報メディア教育研究センター研究報告
高翔, 小杉山格, 齋藤兆古, “可視化画像の固有パターンを用いた画像認識”
- [3] 2006年度 第15回 MAGDA コンファレンス in 桐生
高翔, 青木亮一, 齋藤兆古, “画像の幾何学的固有パターンを用いた手話識別”
- [4] 可視化情報学会第35回可視化情報シンポジウム
高翔, 齋藤兆古, 堀井清之, “赤外線動画像の識別”

謝辞

本研究を進めるに当たり，齋藤兆古教授には数多くのご指導，ご支援を賜りました．深く感謝致します．

また，齋藤兆古研究室の皆様，公私にわたりご助言，ご支援を賜りました．ありがとうございました．