

2004 年度

修士論文

画像の RGB 情報とその応用に関する研究

A STUDY OF RGB COLOR INFORMATION AND ITS APPLICATION

法政大学大学院工学研究科

電気工学専攻修士課程

丸山 和夫

学籍番号 : 03R3237

指導教授 : 齋藤 兆古

Abstract

This paper concerns with a method of dynamic image cognition. Our image cognition method has two distinguished features. One is that the image is handled as its Eigen (characteristic) pattern and time series frequency characteristic. The other is that the cognition is carried out by solving for the ill-posed linear system of equations comprising the Eigen patterns and time series frequency characteristic as a database. The Eigen pattern consists of histograms in RGB color components of the image. This Eigen pattern of color components yields a numerical data set which is hard to be changed by the differences in spatial position of target as well as resolution of the image. The time series frequency characteristic consists of Fourier power spectra of time changing Eigen patterns. This time series frequency characteristic is a distinct frequency characteristic to each of the objects. The image cognition method proposed in this paper is composed of the Eigen pattern and time series frequency characteristic.

As a result, it is revealed that the method proposed in this thesis is one of powerful and effective methods of dynamic image cognition.

- 目次 -

第 1 章 緒論	3
第 2 章 可視化情報の特徴量	4
2.1 空間情報	4
2.1.1 固有パターン	4
2.1.2 固有パターンの動画像への適用	5
2.2 空間情報の特徴	6
2.2.1 固有パターンの検証	6
2.2.2 人物画像の固有パターン	7
2.3 時間軸情報	8
2.3.1 従来のゆらぎ画像解析	8
2.3.2 時系列周波数特性による画像解析	10
2.4 時間軸情報の特徴	11
2.4.1 光源のゆらぎ検証	11
2.4.2 自然界のゆらぎ画像解析	13
第 3 章 可視化情報の特徴量を利用した画像認識	15
3.1 最小自乗法による一貫性評価	15
3.2 固有パターンを利用した動画像認識	17
3.2.1 1 フレーム画像認識	17
3.2.2 動画像認識	19
3.2.3 3 CCD 動画像認識	22
3.3 固有パターンと時系列周波数特性を用いた動画像認識	24
3.3.1 固有パターンと時系列周波数特性による風景認識	24
3.3.2 固有パターンと時系列周波数特性による 3 CCD 動画像認識	28
第 4 章 結論	31
参考文献	32

第1章 緒 論

本論文は、CCD(Charge Coupled Device)カメラを装着した計算機に人間と同等の視覚情報処理能力を与えることを究極の目標とする研究にかかわる。画像を計算機によって人間と同等に認識させることが可能であれば、人間の脳が行う視覚情報処理を計算機が代行可能であることを意味し、産業のオートメーション化や建築物などのセキュリティシステムなど、人間の視覚情報処理を必要とする分野を全て機械に代行させることが可能となる[1]。

従来の大部分の画像認識技術は、対象の幾何学的特異点抽出に基づいた手法が一般的で、対象の幾何学的変形に対応不可能なため、人間の情報処理能力には到底及ばない問題点があった。この問題点を鑑み、本論文は動画像の空間情報と時間軸情報を利用した全く新しい動画像認識手法を提案し良好な精度で動画像認識が可能であることを報告する。

空間情報として本論文では、計算機に取り込まれた画像から、画素数、配置、変形等に関係となる画像の特徴量を抽出する。本稿では、この空間情報の特徴量を固有パターンと呼ぶ[2]。

時間軸情報として本論文では、計算機に取り込まれた動画像から対象物固有の周波数特性を抽出する。自然界の動きは、天体の自転・公転や振り子の周期などに代表されるように、物体はそれぞれ固有の周波数特性をもっている。この周波数特性はロウソク炎、人間の動き、風景(波や雲など)にも存在し「 $1/f$ ゆらぎ」や「 $1/f^2$ ゆらぎ」などと大まかに分類されている[3][4]。本論文ではこのゆらぎ解析の概念を導入し動画像の時間軸情報の特徴量を抽出し、動画像認識へ応用する。この時間軸情報の特徴量を時系列周波数特性と呼ぶ。

本論文では動画像から固有パターンと時系列周波数特性を抽出し、固有パターンと時系列周波数特性を合わせた画像の特徴量を用いて動画像認識を試みる。

第 2 章 可視化情報の特徴量

可視化情報は空間情報と時間軸情報とに分けることができる。本章では可視化情報の空間情報と時間軸情報の特徴量について述べる。

2.1 空間情報

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、画像データを $m \times n$ 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。

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

2.1.1 固有パターン

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、

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

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。

$$I_R^D \in Round \left[D \times \frac{R_k}{Max[I_R]} \right] \tag{3}$$

$$k = 1, 2, 3, \dots, m \times n$$

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。

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

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固有ベクトルを算出する。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。モノクロ画像の場合, モノクロ画像の固有パターンは 255 次のベクトルとなる[3]。

図 1 に静止画像の固有パターンの 1 例を示す。図 1 に示すように画像はそれぞれ独自の固有パターンを持つことがわかる。

$$E = [E_R, E_G, E_B] \quad (5)$$

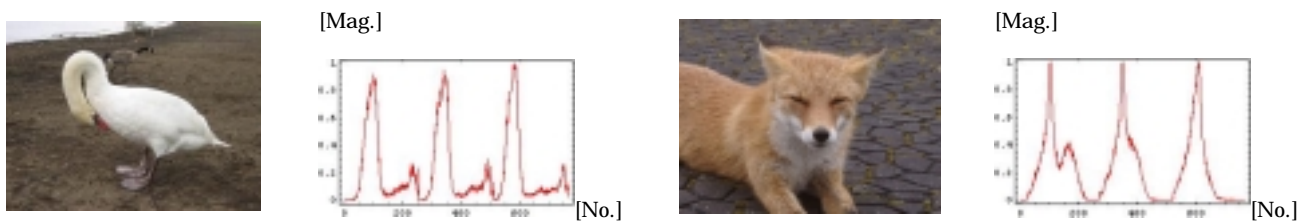
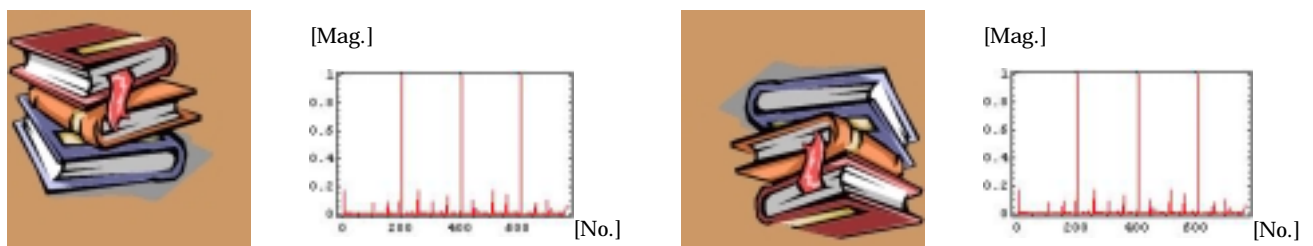


図 1 サンプル画像と固有パターン

図 2 は, 画像中の対象物(本)の位置や角度, 及び画像の解像度が変化した例である。抽出された固有パターンは位置・角度・画素などの情報を削除した特徴量であることがわかる。



(a) 画像 1 (300×300 画素)

(b) 画像 2 (100×100 画素)

図 2 サンプル画像と固有パターン 2

2.1.2 固有パターンの動画画像への適用

次に RGB 固有パターンの抽出方法を動画画像に一般化する。本論文では, 対象物の位置や角度のみならず動きに依存しない特徴量を動画画像の固有パターンとして抽出する。現在一般的に使われているビデオカメラや CCD 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り, 動画画像は静止画像を時系列に並べた一連の画像である。そのため, 動画画像 RGB 固有パターンは, すべてのフレーム画像から画素値分布を算出し画像の特徴量を抽出する。この方法を合成画像法と呼ぶ。

固有パターンはスクリーン上の可視化画像の位置や角度に依存しない量であるため, 動画画像固有パター

ンは、動画像中のフレーム移動情報を削除することが可能となる。合成画像法を用いた RGB 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。

2.2 空間情報の特徴

2.2.1 固有パターンの検証

1 フレーム画像から抽出した固有パターンは、図 1、図 2 に示したように画像固有の特徴であり、位置・角度などが変化しても影響を受けていないことが分かった。ここでは、動画像全フレームから抽出した固有パターンについて述べる。図 3、図 4 に動画像を示す。同一物体であるが動きの異なる動画像である。

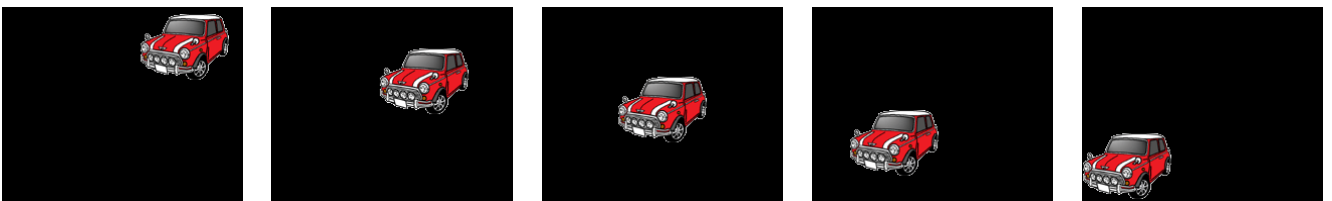


図 3 動画像 1



図 4 動画像 2

図 3、図 4 の動画像から固有パターンを抽出した。それぞれの固有パターンを図 5、図 6 に示す。図 5、図 6 の固有パターンは同値であり動画像から抽出された固有パターンは動きの情報を削除した特徴量であることがわかる[5]。

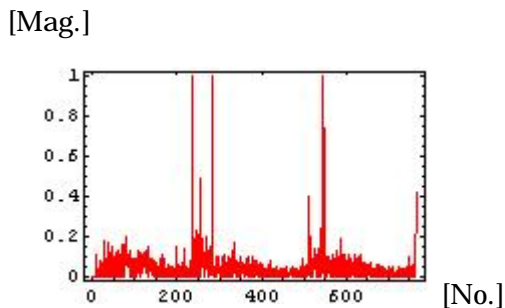


図 5 動画像 1 の固有パターン

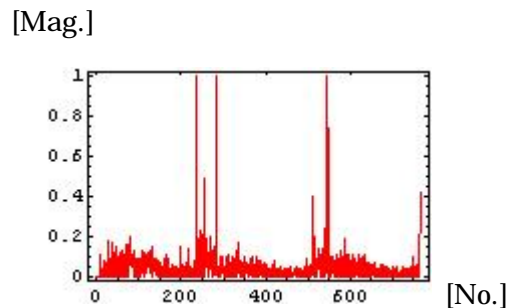


図 6 動画像 2 の固有パターン

2.2.2 人物画像の固有パターン

動画画像の固有パターンが対象物の動きに影響されないことがわかった．ここでは人物の顔画像から固有パターンを抽出し表情が変化しても固有パターンに影響を与えないことを示す．図7，図8に人物動画画像の例を示す．同一人物であるが表情(動き)が異なる動画画像である．

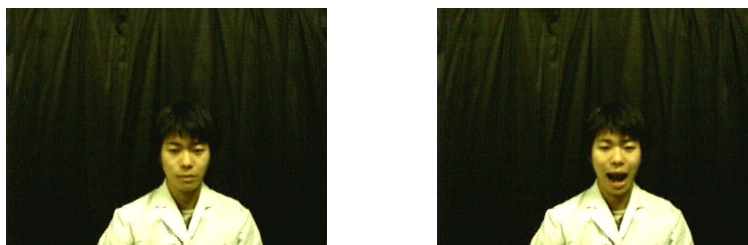


図7 人物動画画像1

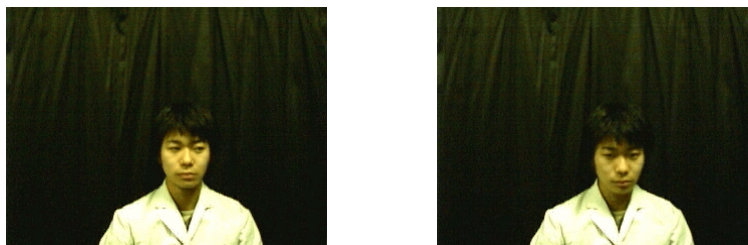


図8 人物動画画像2

図7，図8の動画画像から固有パターンを抽出した．それぞれの固有パターンを図9，図10に示す．図9，図10の固有パターンは，ほぼ同値であり，固有パターンは表情の変化(動き)に影響されない性質を持つ事が確認できる．

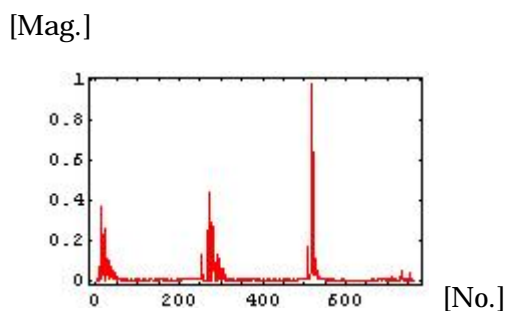


図9 人物動画画像1の固有パターン

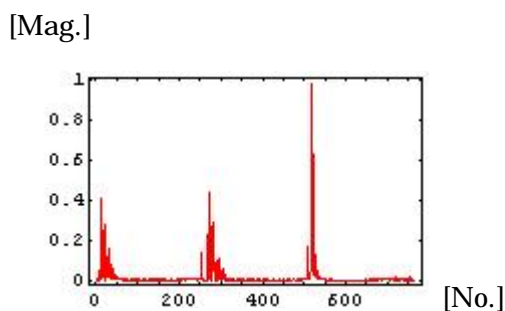


図10 人物動画画像2の固有パターン

固有パターンは対象物の動き，位置，角度に影響を受けないため，人物画像から固有の特徴量を抽出することができた．従来の幾何学的特異点抽出に基づいた画像認識では困難とされていた人物画像の認識に有用であると考えられる．

2.3 時間軸情報

次に動画の時間軸情報から特徴量を抽出する方法を述べる。本論文では画像の時間軸情報として RGB 成分の時系列変化をゆらぎとして捕らえ、そのゆらぎから動画の特徴量を抽出することを試みる。まず、従来のゆらぎ画像解析法について述べる。次に固有パターンを従来のゆらぎ解析へ導入した、新しいゆらぎ画像解析法について述べる。

2.3.1 従来のゆらぎ画像解析

従来のゆらぎ画像解析法とは、RGB 成分の強度値を 1 フレームに対する平均値から算出し、それを全フレームに対して求め時系列に並べることによってゆらぎを 1 次元数値配列として取得し、その数値配列に対してフーリエ解析を用いてゆらぎの特徴を求める。この方法は一般的な「 $1/f$ ゆらぎ」の解析を行なう場合に用いられる。

従来のゆらぎ解析の有用性の検証として「 $1/f$ ゆらぎ」を解析する。「 $1/f$ ゆらぎ」とは自然界に多く存在し、例えば小川のせせらぎ、ロウソクの炎、爽やかなそよ風などの快いリズムが相当する。従来のゆらぎ解析は、ある信号を離散的フーリエ変換により、各周波数に対するフーリエ・パワースペクトラムを計算し、フーリエ・パワースペクトラムの概形でゆらぎの解析を行なう。すなわち、周波数の低下とともにフーリエ・パワースペクトラムが増加するような信号の中で、フーリエ・パワースペクトラムの振幅が周波数に対して反比例する信号を「 $1/f$ ゆらぎ」と呼ぶ。

従来のゆらぎ画像解析で視覚的に判りやすくするために、よく行われる方法はフーリエ・パワースペクトラム対周波数の両対数図を描き、図の傾きによってゆらぎの種類を大まかに分類し、傾きが -1 のときを「 $1/f$ ゆらぎ」、傾きが -2 のときを「 $1/f^2$ ゆらぎ」というように分類する方法である。図 1 1 にフーリエ・パワースペクトラムの一例を示す。

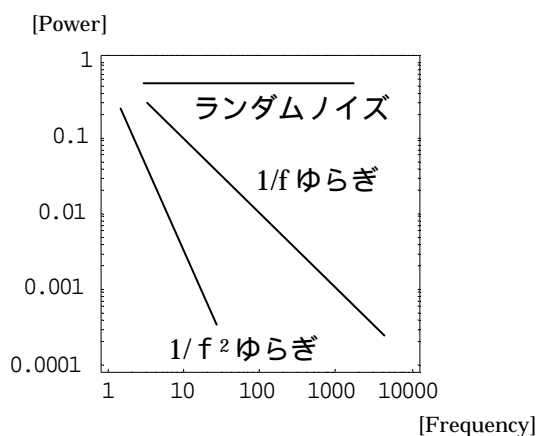


図 1 1 フーリエ・パワースペクトラムの例

実際に CCD カメラにより撮影したロウソク炎の動画を計算機に取り込み、RGB 成分のゆらぎを抽出する。

計算機に取り込まれた動画画像から RGB 成分強度値それぞれの平均値を求め、時系列に並べる。その結果、時系列（フレーム系列）方向の RGB 成分のゆらぎを得ることができる。図 1 2 はろうそく炎とその RGB 成分のゆらぎを時系列に表した波形である。

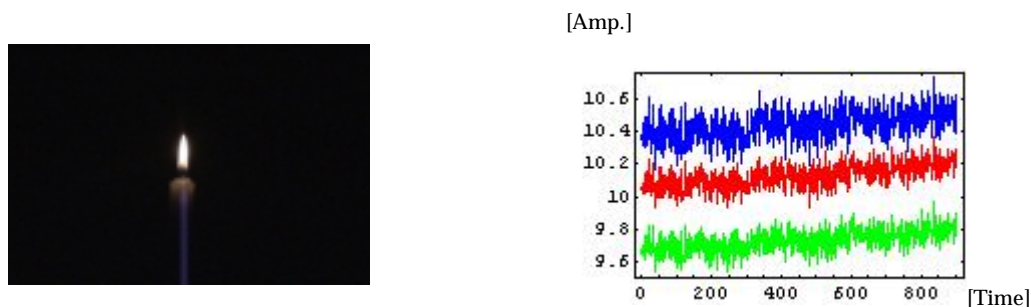
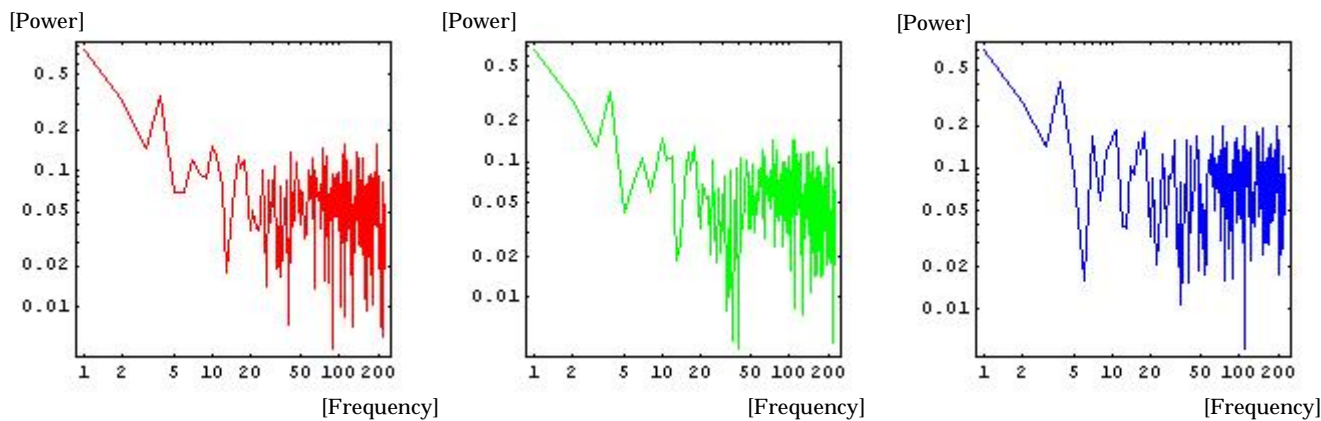


図 1 2 ろうそく炎と RGB 成分のゆらぎ

図 1 2 に示した波形に対して、ゆらぎの種類を従来のゆらぎ解析により吟味する。画像から得られた図 1 2 の信号に離散的フーリエ変換（DFT）を適用し、各周波数に対するフーリエ・パワースペクトラムを計算する。フーリエスペクトラム・パワーと周波数をそれぞれ、縦軸と横軸に対応させ、両対数図を描く。

図 1 3 に示すような両対数軸上に描かれる曲線へ累乗近似を適用し、回帰直線の傾きを求める。ただし、DFT スペクトラムは、全サンプル数に等しい実部と虚部を与えるから、独立なパワースペクトラムは全サンプル数の半分までで有り、さらに DFT の精度を勘案して、全サンプル数の 1/4 項までの低周波領域から傾きを求める。図 1 3 は典型的なパワースペクトラム対周波数特性である。周波数が高いほど DFT の精度が悪く、振動的にパワースペクトラムが変化することが判る。



(a)R 成分

(b)G 成分

(c)B 成分

図 1 3 1/f ゆらぎの RGB 成分フーリエ・パワースペクトラム

図 1 3 に示したフーリエ・パワースペクトラムから傾きを求める方法を述べる。フーリエ・パワースペクトラムが n 次の数値配列で構成されているとすると図 1 3 に示したフーリエ・パワースペクトラムは式 (6) で表される。

$$f = [y_1, y_2, y_3, \dots, y_n] \quad (6)$$

この数値配列を両対数グラフで表した場合の傾きを求める方法として本研究では累乗近似を採用し、直線に回帰させる。すなわち、式(7)により回帰直線の傾きを求めることができる。

$$X_i = \log i \quad Y_i = \log y_i$$

$$a = \frac{n \sum_{i=1}^n X_i \cdot Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n Y_i \cdot Y_i - \sum_{i=1}^n Y_i \sum_{i=1}^n Y_i} \quad (7)$$

式(7)を用いて、ろうそく炎のゆらぎのフーリエ・パワースペクトラムの傾きを算出したところ RGB 成分のそれぞれの傾きは R = -0.24 G = -0.21 B = -0.12 となった。ろうそくの炎は「1/f ゆらぎ」が存在するので、フーリエ・パワースペクトラムの傾きは -1 であると予想していたが、ゼロに近い値を取っている。これは動画画像からゆらぎを抽出する際に RGB 成分の平均化を行なっているためろうそく炎以外の背景部分のゆらぎ(ホワイトノイズ)に影響を受け「1/f ゆらぎ」特性が平均化してしまったことに起因する。

2.3.2 時系列周波数特性によるゆらぎ画像解析

従来のゆらぎ画像解析では RGB 成分の平均値を用いて解析を行なっていたため正確な解析ができなかった。ここでは、RGB 成分強度値の 1 から 255 において、より詳細に解析をおこなうため、ゆらぎ画像解析へ固有パターンを導入し、新手法のゆらぎ画像解析を提案する。

図 1 4 (a)に示す動画画像の 1 フレームから固有パターンを抽出すると図 1 2 (b)のようになる。さらに、動画画像の全フレームから固有パターンを抽出し、式(11)に示すように固有パターンを時系列に並べ、固有パターンのゆらぎを図 1 4 (c)に示す。

$$F = [E_1, E_2, \dots, E_n] \quad n: \text{フレーム数} \quad (11)$$

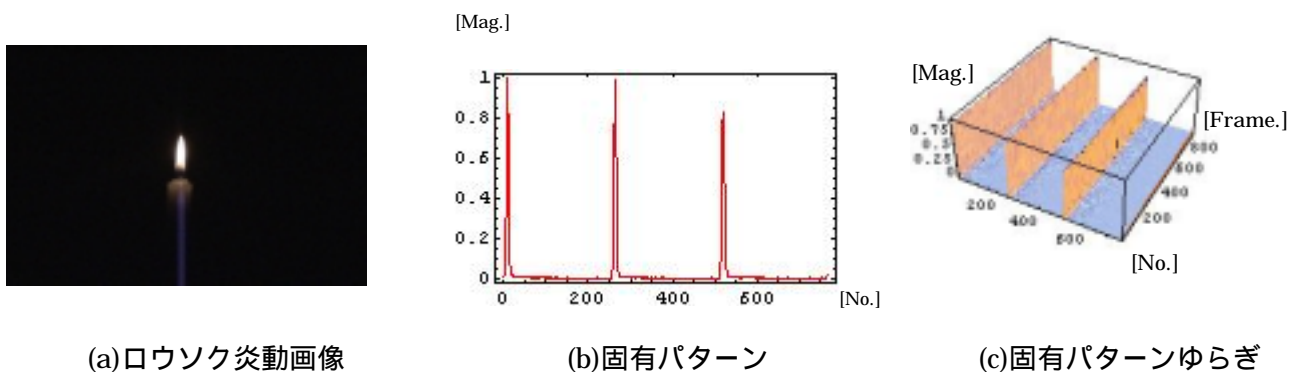


図 1 4 固有パターンと固有パターンのゆらぎ

ロウソク炎動画画像より得られた固有パターンのゆらぎを時系列(フレーム)方向に DFT 解析し、各周波数に対するフーリエ・パワースペクトルを計算する。そして、累乗近似を用いて両対数図で表した場合の傾きを求める。ただし、全サンプル数の 1/4 項までの低周波領域を計算に採用した。その結果、RGB 成分の 1 から 255 までの強度分布のゆらぎの特性を表した数値配列を得ることができ、横軸に RGB 成分強度値、縦軸にフーリエ・パワースペクトラムの傾きをとリグラフ化した。-1 に近い程「1/f ゆらぎ」に近い信号が存在することを意味し、0 に近いほどランダムノイズであることを意味する。

図 1 5 は、図 1 4 (c) の RGB 固有パターンのゆらぎにおけるフーリエ・パワースペクトラムに累乗近似を適用して得られた傾き分布である。ロウソク燃焼画像は RGB 成分それぞれの強度値の 255 で「1/f ゆらぎ」をもっていることを表している。この傾き分布は RGB 成分の 1 から 255 の周波数特性を表しており、従来のゆらぎ解析では不可能であった動画画像から正確なゆらぎ解析を可能としている[6]。これを時系列周波数特性と呼ぶ。

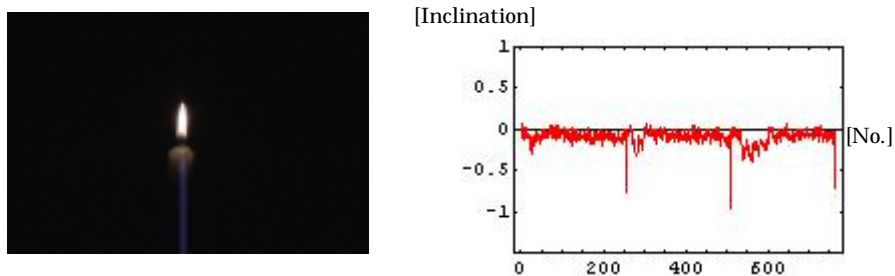


図 1 5 時系列周波数特性

2.4 時間軸情報の特徴

2.4.1 光源のゆらぎ検証

ここでは RGB 成分それぞれの強度値の 1 から 255 における解析が可能である時系列周波数特性の検証を行なう。正確な解析ができていないか否かを検証するため、人工的に白熱電球と LED に「1/f ゆらぎ」や「1/f^{0.5} ゆらぎ」を与え、ゆらぎ解析を行ない正しい解析ができていないかを検証する。

白熱電球や LED などの電気照明に 1/f ゆらぎを導入するためには、その基となる 1/f ゆらぎ電気信号の生成が必要である。この生成には、抵抗の熱雑音を利用する方法、自然界における 1/f ゆらぎをデータとして計算機に取り込む方法、カオスによる出力が 1/f ゆらぎとなるような現象を計算機もしくは電子回路で直接発生させる方法などが考えられる。本論文では計算機で発生させたランダムノイズから 1/f ゆらぎ信号の生成を試みる。

計算機の乱数によって発生させたランダムノイズをフーリエ変換し、実数部と虚数部のデータ配列に式(12)で示した 1/f フィルタを畳み込み演算し、逆フーリエ変換することで 1/f ゆらぎ信号を生成する。ただし式(12)の n はランダムノイズの全サンプル数を表す。図 1 4 にランダムノイズを示す。図 1 5 にフィルタリング後の 1/f ゆらぎ信号を示す。

$$filter = \left[0, 1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{\frac{n}{2}-2}, \frac{1}{\frac{n}{2}-1}, \frac{1}{\frac{n}{2}}, \frac{1}{\frac{n}{2}-1}, \frac{1}{\frac{n}{2}-2}, \dots, \frac{1}{3}, \frac{1}{2}, 1 \right] \quad (12)$$

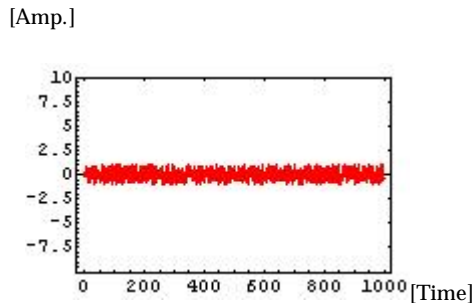


図 1 4 ランダムノイズ

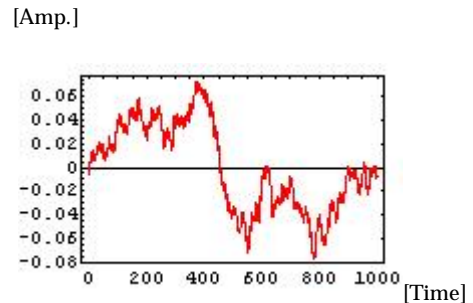


図 1 5 1/f ゆらぎ信号

また、ランダムノイズのフーリエ・パワースペクトラムとフィルタリング後の 1/f ゆらぎ信号のフーリエ・パワースペクトラムを図 1 6 と図 1 7 に示す。ランダムノイズのフーリエ・パワースペクトラムの傾きは約 0 であるのに対し、フィルタリング後の 1/f ゆらぎ信号のフーリエ・パワースペクトラムの傾きは約 -1 であることが確認できる。

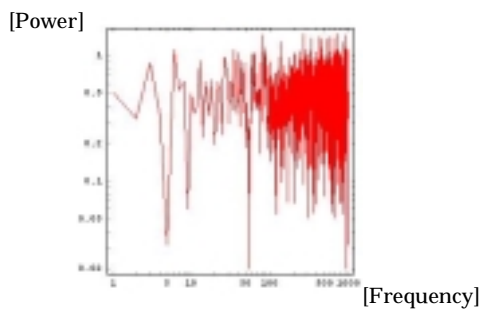


図 1 6 ランダムノイズのフーリエ・パワースペクトラム

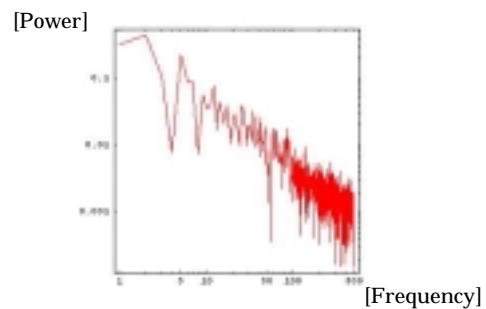


図 1 7 1/f ゆらぎ信号のフーリエ・パワースペクトラム

得られた 1/f ゆらぎ信号を DA 変換ボードにより電圧に変換し白熱電球や LED に様々なゆらぎを持たせる。ただし、光源のゆらぎを CCD カメラで撮影するため、信号のサンプリング周波数を CCD カメラのフレームレート(29.97Hz)以下に設定した。

時系列周波数解析を用いて、通常白熱電球、1/f^{0.5} ゆらぎ白熱電球、1/f ゆらぎ白熱電球、通常の LED、1/f^{0.5} ゆらぎ LED、1/f ゆらぎ LED の 6 種類の光源の動画解析を行なった。解析の結果を図 1 8 に示す。

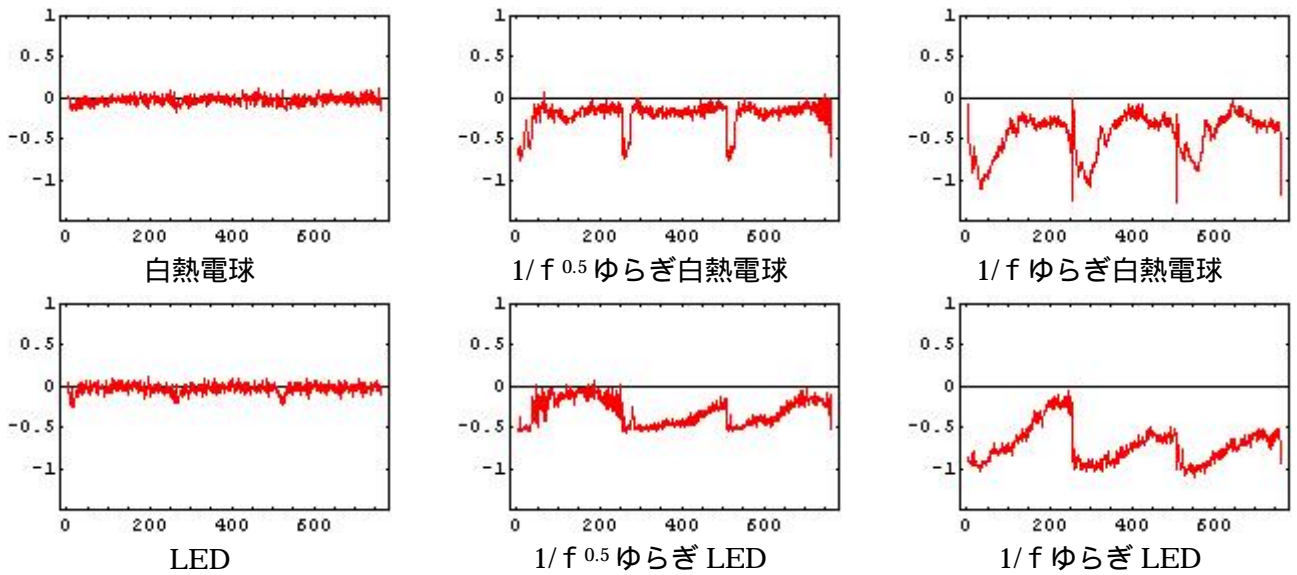


図 18 様々なゆらぎの時系列周波数特性

制御していない白熱電球と LED は全ての RGB 成分強度値において約 0 の数値を示している。また、 $1/f^{0.5}$ ゆらぎに制御した白熱電球と LED は約 -0.5 の数値を示している。 $1/f$ ゆらぎに制御した白熱電球と LED は -1 の数値を示している。よって、画像の RGB 成分ヒストグラムの時系列変化にもとづく時系列周波数特性による分布定数系ゆらぎ解析は、RGB 成分の 1 から 255 までの強度分布における、正確なゆらぎの解析を可能とし動画の時間軸特徴を明確に表している。

2.4.2 自然界のゆらぎ画像解析

動光源画像から時系列 RGB 固有パターンの DFT 解析により動画の時系列周波数特性解析の方法を述べた。ここでは、時系列周波数特性を用いてロウソクやガスの燃焼炎、更には自然界の風景などを解析した結果を述べる。動画の 1 例とその時系列周波数特性を図 19 に示す。

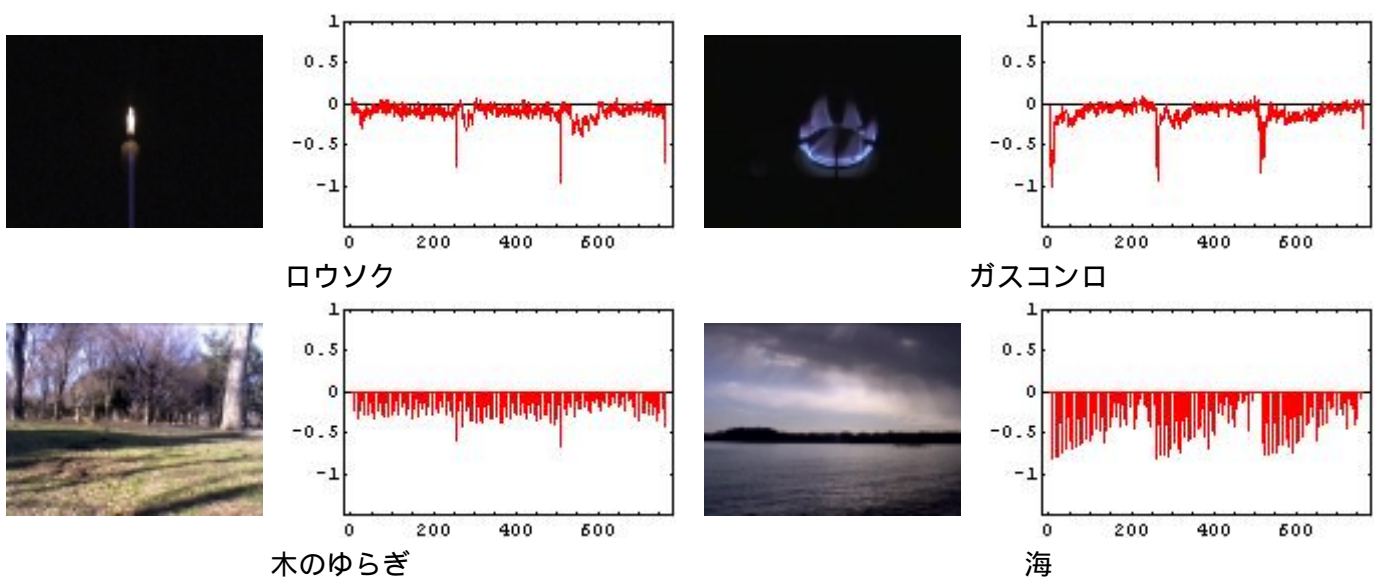


図 19 ゆらぎ解析結果

図19に示すようにロウソクはRGB成分の高い強度値で「 $1/f$ ゆらぎ」をもっている。ガスの炎はRGB成分の低い強度値で「 $1/f$ ゆらぎ」を持っている。また、木のゆらぎは「 $1/f^{0.5}$ ゆらぎ」、海は「 $1/f$ ゆらぎ」をもっている。これにより、可視化情報は対象物固有の時系列情報をもっていることがわかる。

また、図20、図21に異なる場所で撮影した木のゆらぎ動画像を示す。それぞれの時系列周波数特性を図22、図23に示す。



図20 木のゆらぎ1



図21 木のゆらぎ2

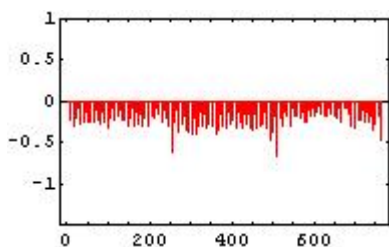


図22 木のゆらぎ1の時系列周波数特性

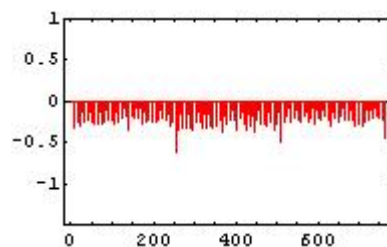


図23 木のゆらぎ2の時系列周波数特性

異なる場所で撮影した木のゆらぎから、ほぼ等しい時系列周波数特性を得た。この結果からわかるように時系列周波数特性は対象物固有の時間軸特徴を抽出しており、動画像の時間軸情報から物体の固有周波数特性を抽出できていることがわかる。

第 3 章 可視化情報の特徴量を利用した画像認識

2 章で述べた動画画像の特徴量である固有パターンと時系列周波数特性を用いて動画画像認識を行なう。最初に、1 次元数値配列として格納された動画画像の特徴量における一致性評価方法について述べる。次に固有パターンによる動画画像認識について述べ、更に 3 台の CCD カメラを用いることによって認識結果の妥当性判断を計算機に行なわせることが可能であることを述べる。最後に固有パターンと時系列周波数特性による動画画像認識について述べる。

3.1 最小自乗法による一致性評価

本論文における画像認識とは予め複数の画像から特徴量を抽出し、データベースを構築しておく。次に入力画像特徴量を抽出し、最小自乗法を用いてデータベースと入力画像固有パターンにおける一致性を評価し、入力画像の同定を行うものである。ただし、特徴量とは固有パターンや時系列周波数特性の 1 次元数値配列のことを言う。

例として n 個の画像データから特徴量を抽出しデータベースを構築する。

各画像データは 1 個の特徴量ベクトルを持つ。よって n 個のデータベース特徴量ベクトルは式(13)のシステム行列を構築可能とする。

$$C = [E_1, E_2, E_3, \dots, E_n] \quad (13)$$

入力画像データの特徴量ベクトルを Y とすると、線形システム方程式は式(14)で表すことができる。 X は解ベクトルであり、式(13)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(13)を構成する特定の特徴量ベクトルが等しいことを意味する。

$$Y = C \cdot X \quad (14)$$

特徴量ベクトルの要素数は特徴量の選択によって変わってくるが本論文では固有パターンや時系列周波数特性を特徴量とするので要素数は最低でも 765 個あり、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ことができる。すなわち、システム行列 C は正方行列ではなく縦長長方形となり、その結果として逆行列が存在しない。したがって、式(14)のシステム方程式から解ベクトル X 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本論文では最小自乗法を採用する [7]。

ここで最小自乗法について説明する。未知数の数よりも式の数が多い場合、全ての式を満足する解は存在しない。従って、近似解ベクトルを X とすれば、誤差ベクトル e は式(15)で与えられる。

$$e = Y - CX \tag{15}$$

ベクトル e の 2 乗ノルムを最小にする方法が最小自乗法である。式(16)に誤差ベクトルの 2 乗ノルムを示す。

$$\begin{aligned} \|e\|^2 &= e^T e \\ &= [Y - CX]^T [Y - CX] \\ &= Y^T Y - 2Y^T CX + X^T C^T CX \end{aligned} \tag{16}$$

この 2 乗ノルムの極値が、誤差 e の最小値となるので、そのときの X が、求める解ベクトルとなる。極値を求める為に X に関して微分すると式(17)になる。

$$\begin{aligned} \frac{\partial}{\partial X} \|e\|^2 &= -2Y^T C + 2X^T C^T C = 0 \\ X^T &= Y^T C [C^T C]^{-1} \end{aligned} \tag{17}$$

したがって、解ベクトル X の導出は式(18)で与えられる。

$$X = [C^T C]^{-1} C^T Y \tag{18}$$

解ベクトル X の最大値を取る要素から画像は認識される。

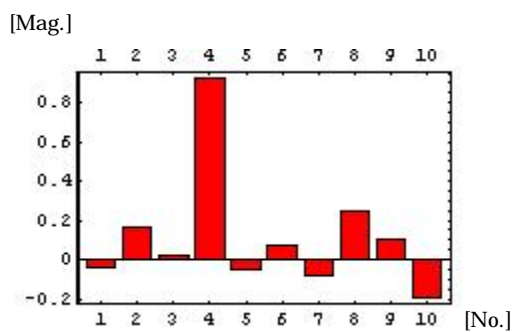


図 2 4 解ベクトルの例

また、システム行列 C が横長長方形列となった場合の不適切な線形システム方程式の解法としては、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[8]。

3.2 固有パターンを利用した動画像認識

3.2.1 1フレーム画像認識

動画像の1フレームのみを使った画像認識を行なう。計算機のスクリーン上の可視化情報は、上述したハードウェア依存性だけでなく、映像生成条件やR,G,Bの色成分のダイナミックレンジへ依存する性質を持つ。固有パターンはRGB情報の変化に影響を受けるため、RGB成分を一定基準に従うダイナミックレンジで抽出する。即ち、映像の生成条件、光源・背景を固定し映像化する。図25にデータベースとなる画像の一例を示す。画像データベースは9個の人物顔画像と1個の人物が映っていない背景画像である。撮影は、背景及び光源が一定の条件で行うため、暗幕で囲った暗室下において光源は単一の白色電球を用いた。

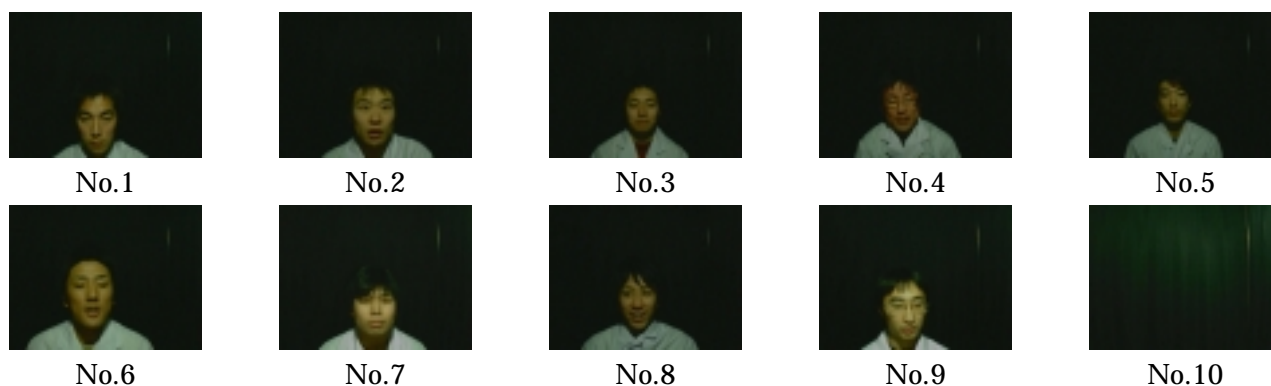


図25 データベース

図26に認識対象となる画像の一例を示す。対象は人物9人の顔画像と1個の人物が映っていない背景画像である。

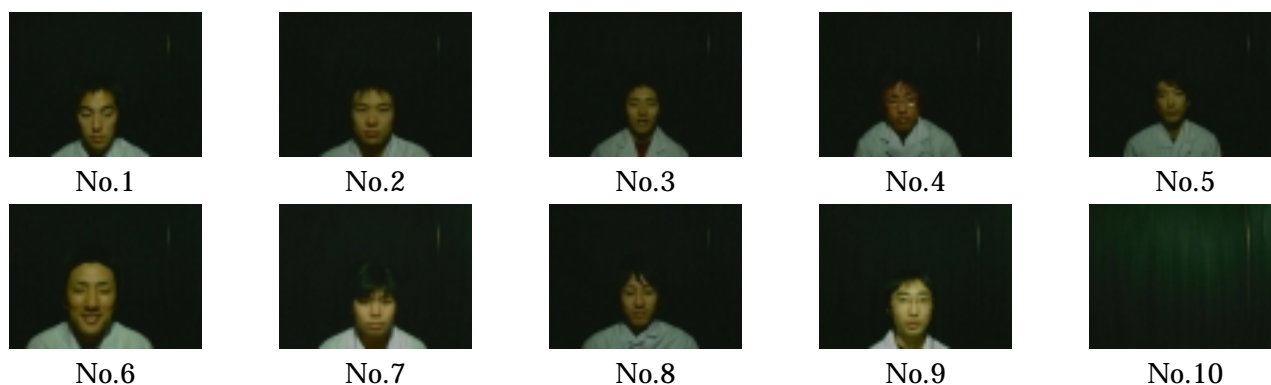


図26 認識対象

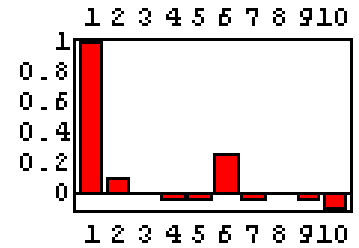
図25の画像データベース、及び図26の認識対象画像からそれぞれ固有パターンを抽出し、最小自乗法を用いて解ベクトル X を求め、解ベクトル中の最大値をとる要素で同定を行なった。認識結果と解ベクトルを図27に示す。解ベクトルの横軸は画像データベースの番号であり、図25データベースの番号と対応している。



認識対象(No.1)



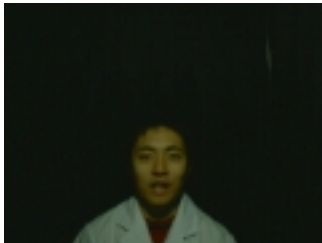
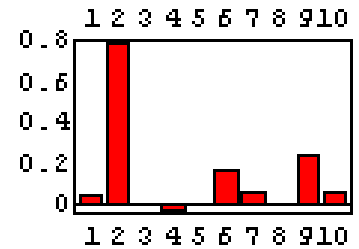
出力画像(正解)



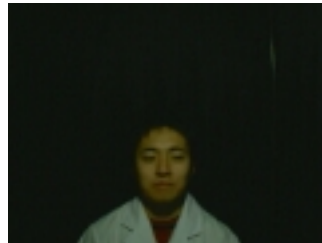
認識対象(No.2)



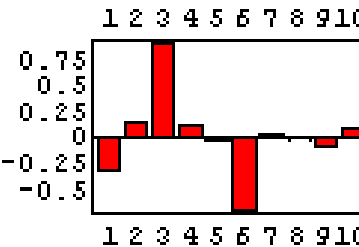
出力画像(正解)



認識対象(No.3)



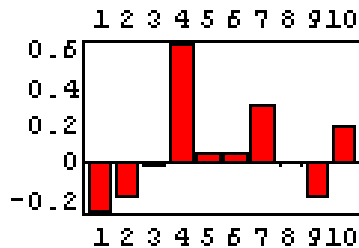
出力画像(正解)



認識対象(No.4)



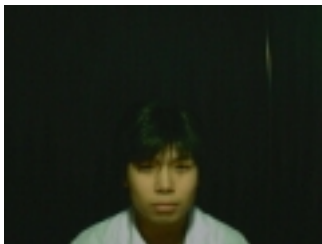
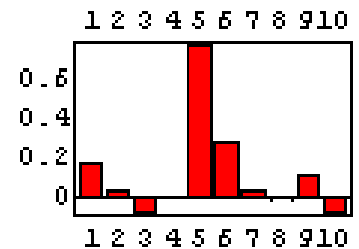
出力画像(正解)



認識対象(No.5)



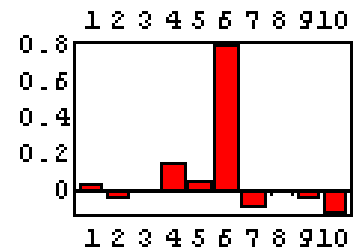
出力画像(正解)



認識対象(No.6)



出力画像(正解)



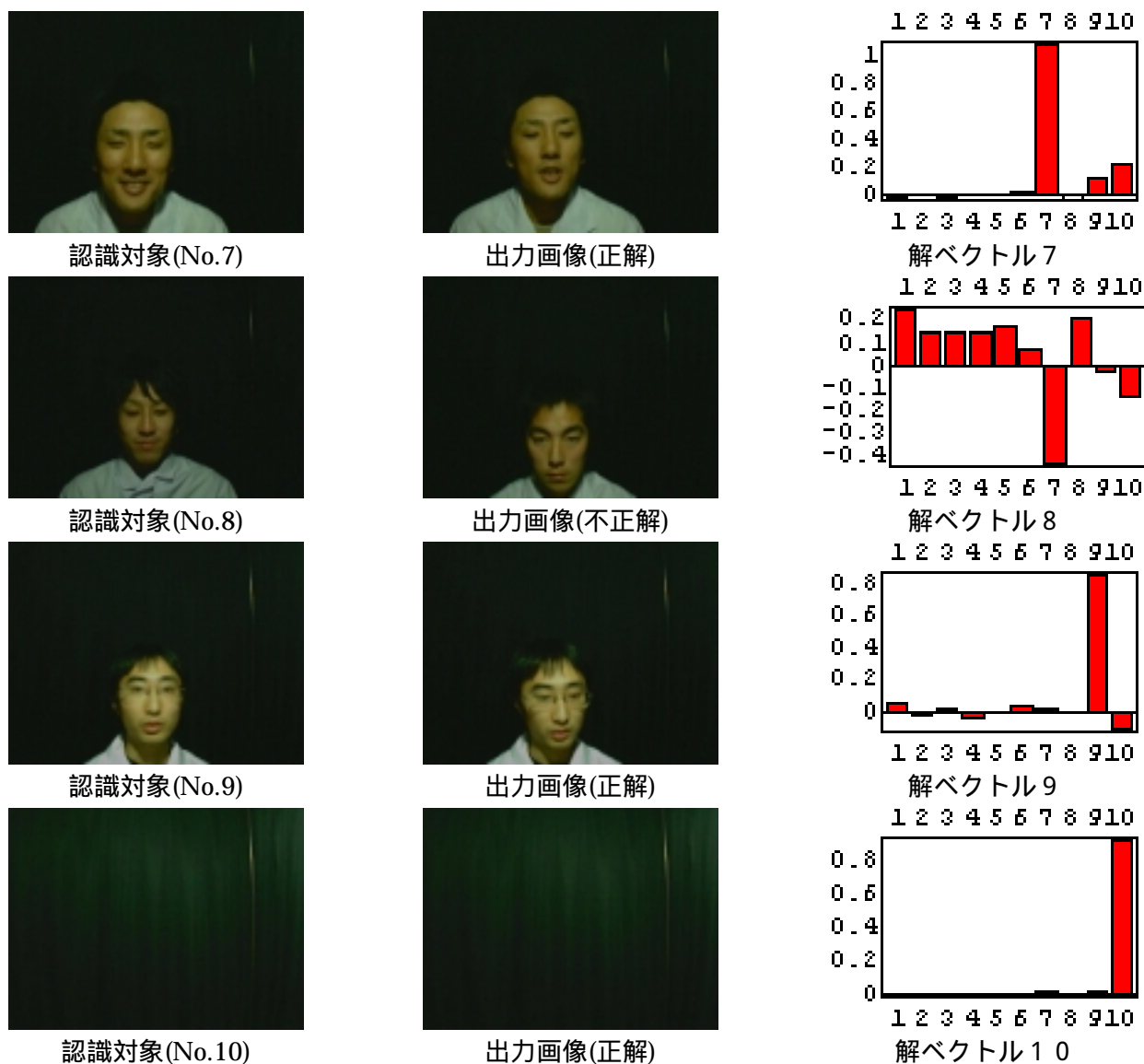
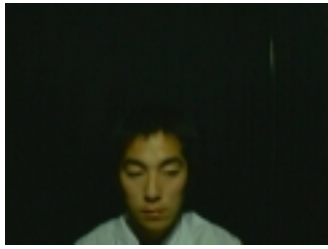


図 2.7 1 フレーム画像認識結果

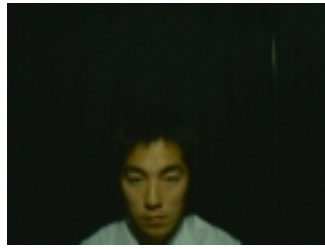
図 2.7 の結果から，No.8 は誤認識であるが，その他は正解である．従来の幾何学的特徴を利用した画像認識では顔の表情に変化がある場合において対応不可能であったが，固有パターンによる画像認識は表情の変化に対応可能である事がわかる．

3.2.2 動画像認識

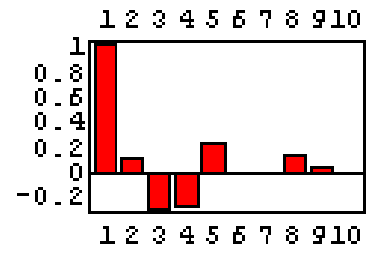
図 2.5 のデータベースと図 2.6 の認識対象動画像の全フレームを用いて動画像認識を試みた．認識結果と解ベクトルを図 2.8 に示す．解ベクトルの横軸は画像データベースの番号であり，図 2.5 データベースの番号と対応している．



認識対象(No.1)



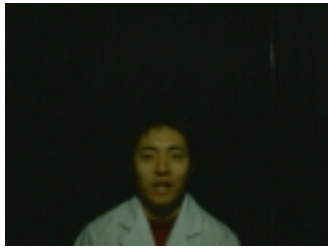
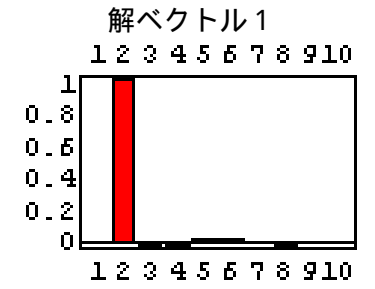
出力画像(正解)



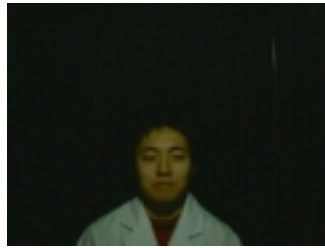
認識対象(No.2)



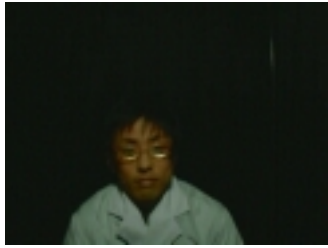
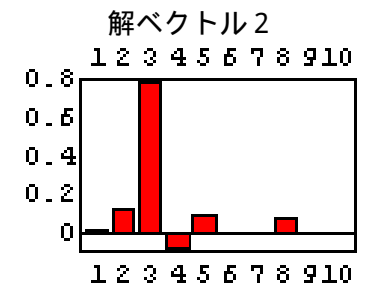
出力画像(正解)



認識対象(No.3)



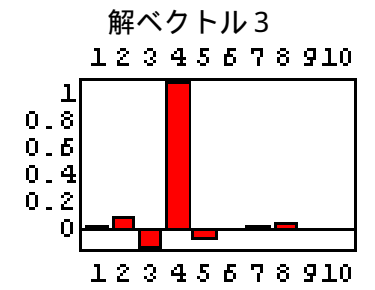
出力画像(正解)



認識対象(No.4)



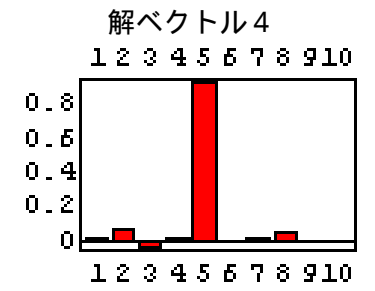
出力画像(正解)



認識対象(No.5)



出力画像(正解)



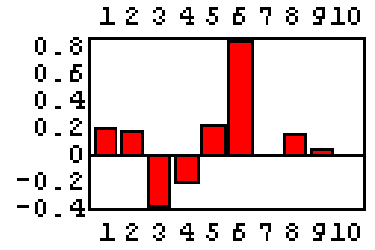
解ベクトル 5



認識対象(No.6)



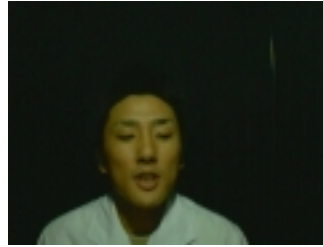
出力画像(正解)



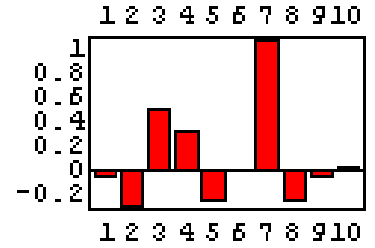
解ベクトル6



認識対象(No.7)



出力画像(正解)



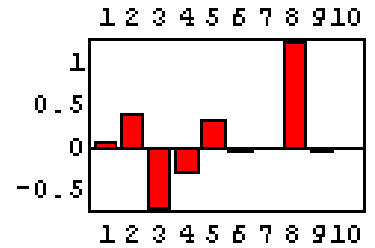
解ベクトル7



認識対象(No.8)



出力画像(正解)



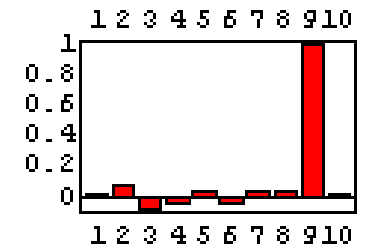
解ベクトル8



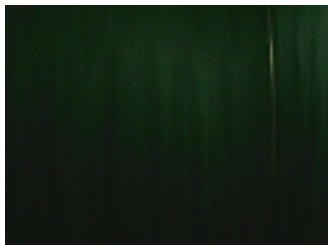
認識対象(No.9)



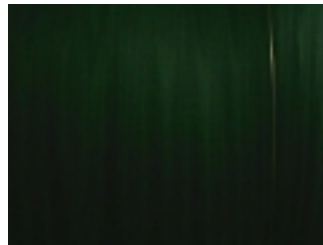
出力画像(正解)



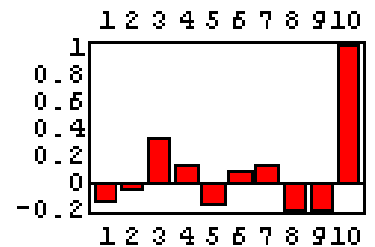
解ベクトル9



認識対象(No.10)



出力画像(正解)



解ベクトル10

図 2 8 動画像認識結果

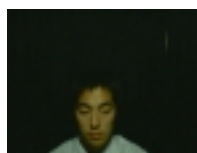
図 2 8 の結果から 図 2 7 に示した 1 フレームのみの認識で誤認識した No.8 においても正しく認識することができた . これは , 動画像の全フレームから固有パターンを抽出することが 1 フレーム画像固有パタ

ーンを加算平均することと等価であり，結果として固有パターンの情報量が豊富になったと考えられる．

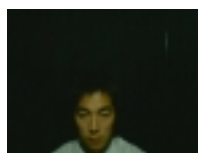
3.2.3 3 CCD 動画像認識

動画像認識のさらなる認識率の向上と認識結果の妥当性を計算機で判定可能とするために固有パターンと線形システム方程式を利用した動画像認識を3台の CCD カメラを併用して行う．

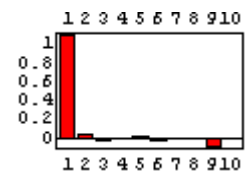
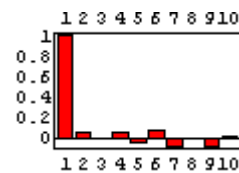
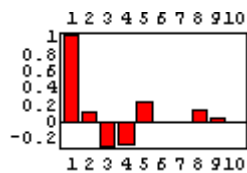
CCD カメラ1，CCD カメラ2，CCD カメラ3を，それぞれ認識対象物に対して異なる角度から撮影できるように配置し，3台の CCD カメラでそれぞれ動画像認識をおこない，独立した3個の解ベクトル X を求める．3個の解ベクトルを比較し2個或いは3個等しい解が得られたならば，その解は信頼性があると判断され，最終的な解として決定する．また3個の解ベクトルを比較した結果，すべて異なる解が得られた場合，その解は信頼性がなく該当なしと判断する．図28は最終的な認識結果と3台の CCD カメラの個々の解ベクトルを示す．データベースと認識対象は図25，図26に示した人物を用いる．更に妥当性判定が可能であることを示すためにデータベースに格納されていない人物(No.11)を新たに認識対象に加えた．



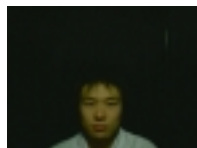
認識対象(No.1)



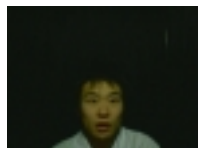
出力画像(正解)



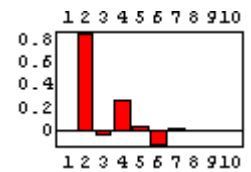
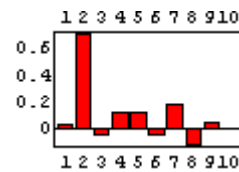
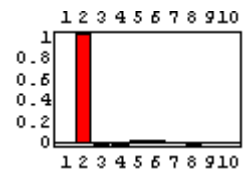
解ベクトル1



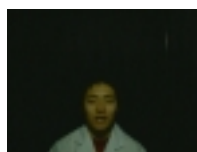
認識対象(No.2)



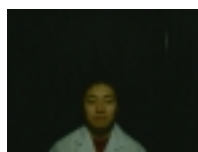
出力画像(正解)



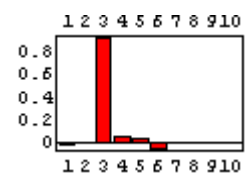
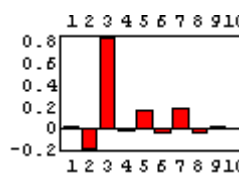
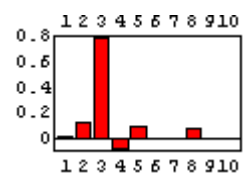
解ベクトル2



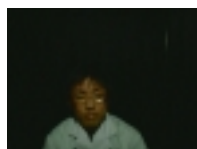
認識対象(No.3)



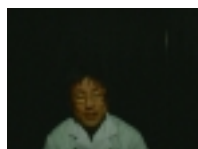
出力画像(正解)



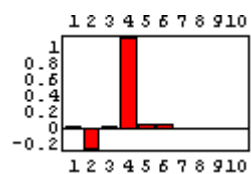
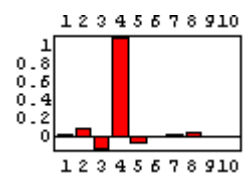
解ベクトル3



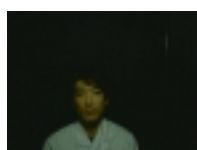
認識対象(No.4)



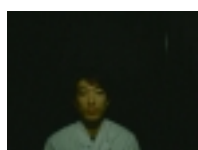
出力画像(正解)



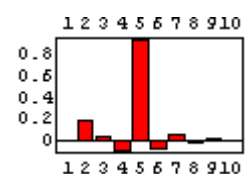
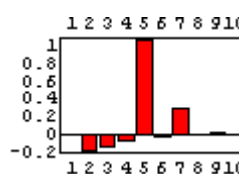
解ベクトル4



認識対象(No.11)



出力画像(正解)



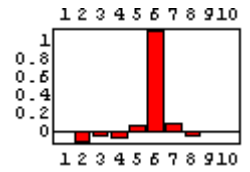
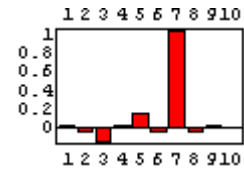
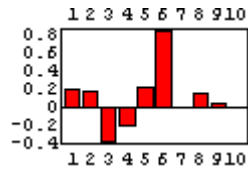
認識対象(No.5)



出力画像(正解)



解ベクトル5



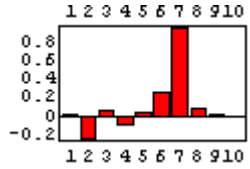
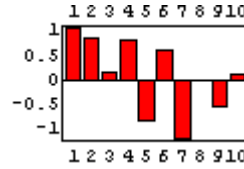
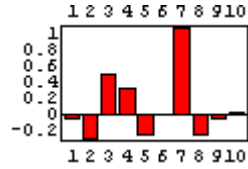
認識対象(No.6)



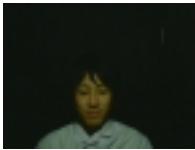
出力画像(正解)



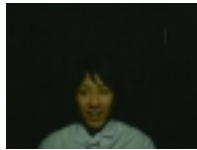
解ベクトル6



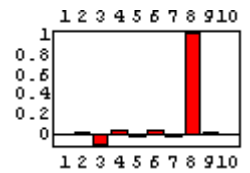
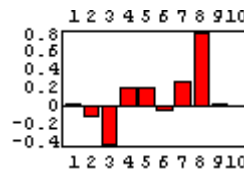
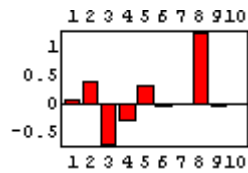
認識対象(No.7)



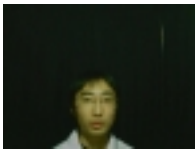
出力画像(正解)



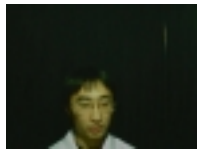
解ベクトル7



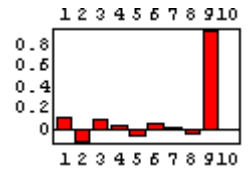
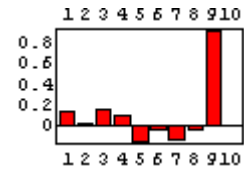
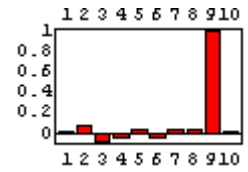
認識対象(No.8)



出力画像(正解)



解ベクトル8



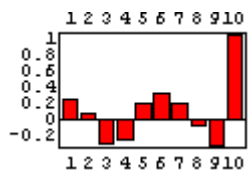
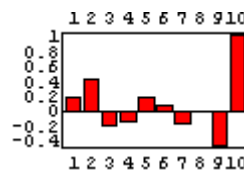
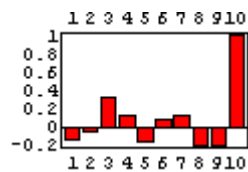
認識対象(No.9)



出力画像(正解)



解ベクトル9



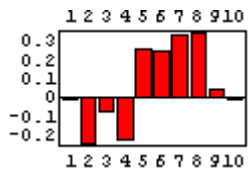
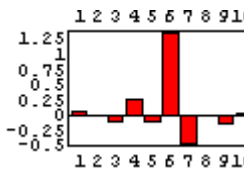
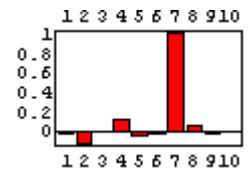
認識対象(No.10)



出力画像(正解)



解ベクトル10



認識対象(No.11)

出力画像(該当なし)

解ベクトル11

図 2 8 3CCD 人物認識結果

3 台の CCD カメラを併用することにより信頼性のない結果の識別が可能となり，図 2 8 (No.11)に示すようにデータベースにない認識対象を「該当なし」として判別することができ，認識率の向上と計算機による妥当性判定が可能になった。

3.3 固有パターンと時系列周波数特性を用いた動画像認識

人間は無限に存在する風景画像において、種類別に認識をおこなっている。例えば図27に示した画像は「波打ち際と雲の画像」として認識できる。「波打ち際と雲の画像」というものは無限に存在するが、人間は図27を「波打ち際と雲の画像」の1種として分類できる。本章では、このように撮影場所によって無限に存在する風景画像を本論文で提案する動画像認識法により識別する。

2章で述べたように、物体は固有の時系列周波数特性をもっており、自然界の風景においても固有の時系列周波数特性がある。このことから、時系列周波数特性を用いることにより風景の識別が可能であると考えられる。

本章では固有パターンと時系列周波数特性を用いて、風景画像(木、波など)の認識をおこない、さらに本動画像認識が人物と風景の両方の動画像認識に対応できることを報告する。

3.3.1 固有パターンと時系列周波数特性による動画像認識

可視化情報の特徴量は2章で述べたように固有パターンと時系列周波数特性で表すことができる。固有パターンは対象物の位置、角度に影響を受けない空間情報である。時系列周波数特性は物体固有の時間軸情報である。これらの特徴量を合わせ風景画像の認識をおこなう。固有パターンと時系列周波数特性を合わせた動画像の特徴量を図27に示す。

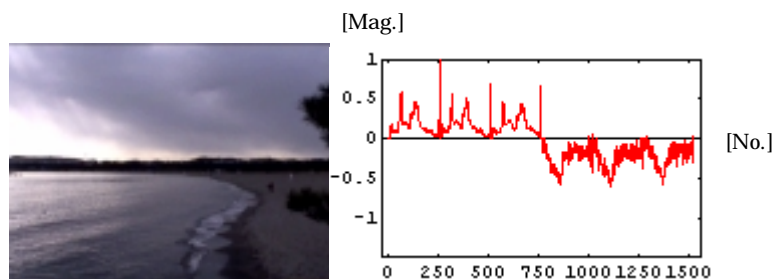


図27 動画像の特徴量

動画像の特徴量を用いて風景認識をおこなう。データベースの1例を図28、認識対象を図29に示す。動画像認識の結果と解ベクトルを図30に示す。解ベクトルの横軸はデータベースの番号であり、図28データベースの番号と対応している。認識結果は全て正解であり本論文で述べた固有パターンと時系列周波数特性を合わせた特徴量は、風景動画像であっても認識可能であることが示された。

また、比較として固有パターンのみを用いた動画像認識結果を図31に示す。認識対象のNo.1, No.2, No.6で誤認識をしており、固有パターンのみでは対応できないことがわかる。したがって、固有パターンと時系列周波数特性を合わせた特徴量での動画像認識は風景認識において有用であることがわかる。



No.1



No.2



No.3



No.4



No.5



No.6

図 2 8 風景データベース



No.1



No.2



No.3



No.4



No.5



No.6

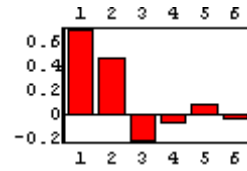
図 2 9 風景データベース



認識対象(No.1)



出力画像(正解)



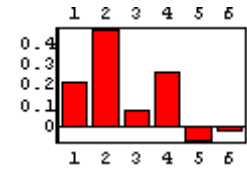
解ベクトル1



認識対象(No.2)



出力画像(正解)



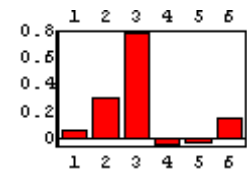
解ベクトル2



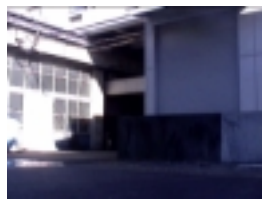
認識対象(No.3)



出力画像(正解)



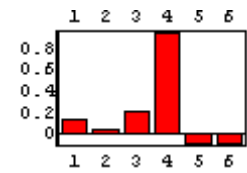
解ベクトル3



認識対象(No.4)



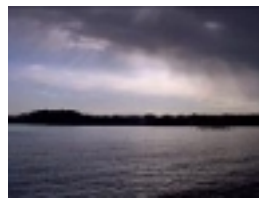
出力画像(正解)



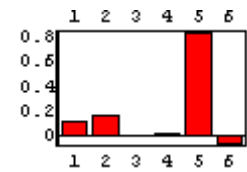
解ベクトル4



認識対象(No.5)



出力画像(正解)



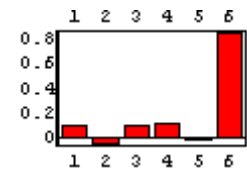
解ベクトル5



認識対象(No.6)



出力画像(正解)



解ベクトル6

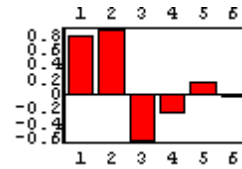
図 3 0 固有パターンと時系列周波数特性による動画像認識結果



認識対象(No.1)



出力画像(不正解)



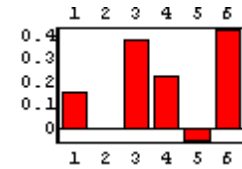
解ベクトル1



認識対象(No.2)



出力画像(不正解)



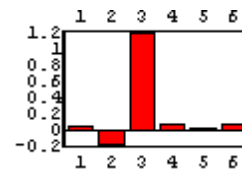
解ベクトル2



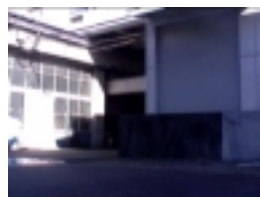
認識対象(No.3)



出力画像(正解)



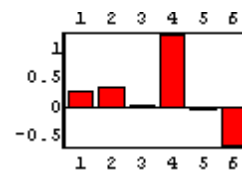
解ベクトル3



認識対象(No.4)



出力画像(正解)



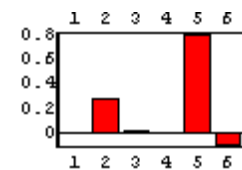
解ベクトル4



認識対象(No.5)



出力画像(正解)



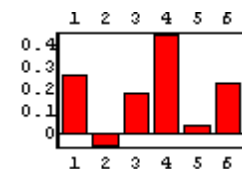
解ベクトル5



認識対象(No.6)



出力画像(不正解)

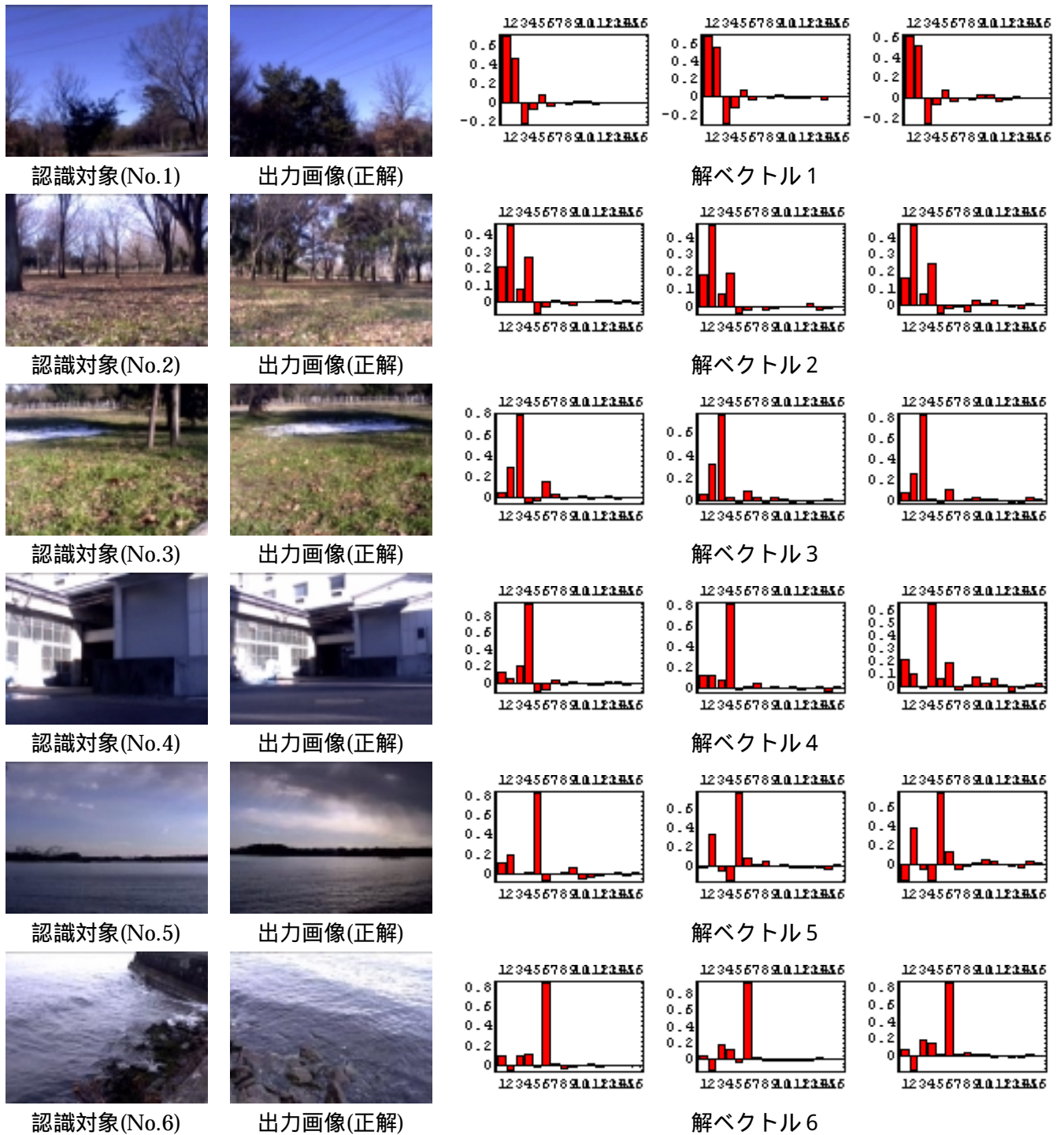


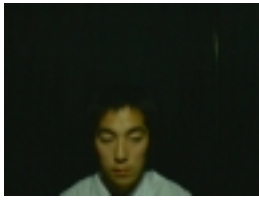
解ベクトル6

図3 1 固有パターンによる動画像認識結果

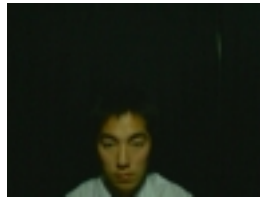
3.3.2 固有パターンと時系列周波数特性による 3CCD 動画像認識

以上述べた動画像の特徴量（固有パターンと時系列周波数特性），3CCD 動画像認識法を用いて，人物と風景の混合データベースにおける動画像認識をおこなう．データベースに図 2 5 の人物と図 2 8 の風景を用い，認識対象に図 2 6 の人物と図 2 9 の風景を用いた．その結果を図 3 2 に示す．

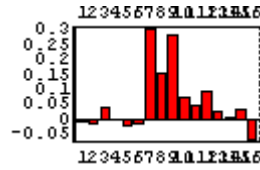




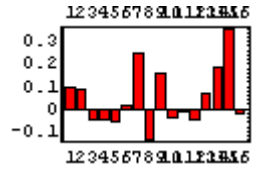
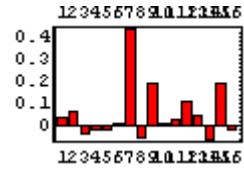
認識対象(No.7)



出力画像(正解)



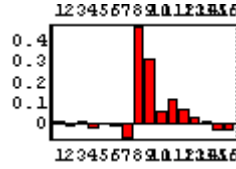
解ベクトル7



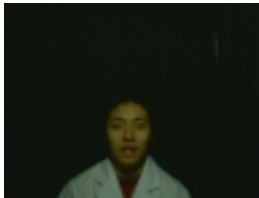
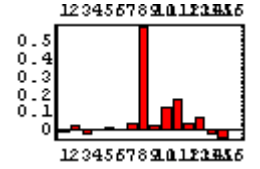
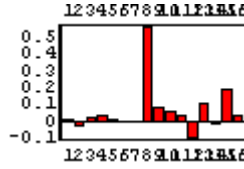
認識対象(No.8)



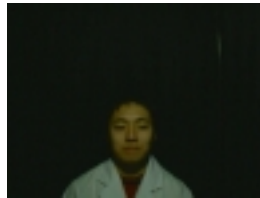
出力画像(正解)



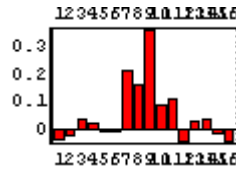
解ベクトル8



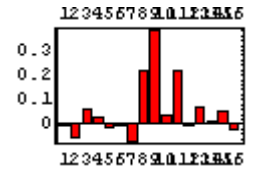
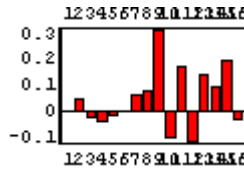
認識対象(No.9)



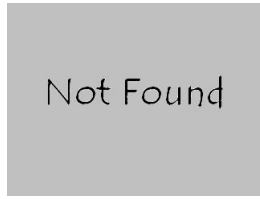
出力画像(正解)



解ベクトル9

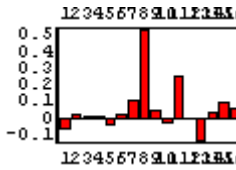


認識対象(No.10)

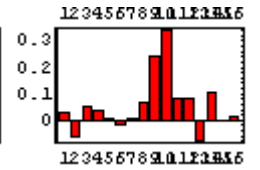
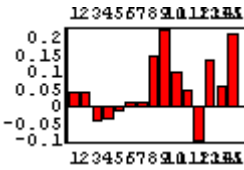


Not Found

出力画像(該当なし)



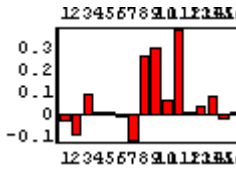
解ベクトル10



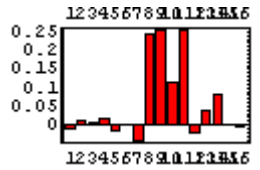
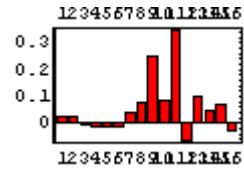
認識対象(No.11)



出力画像(正解)



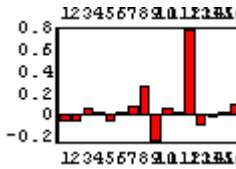
解ベクトル11



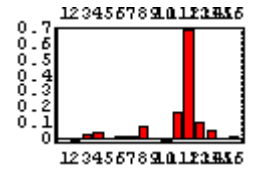
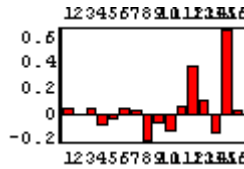
認識対象(No.12)



出力画像(正解)



解ベクトル12



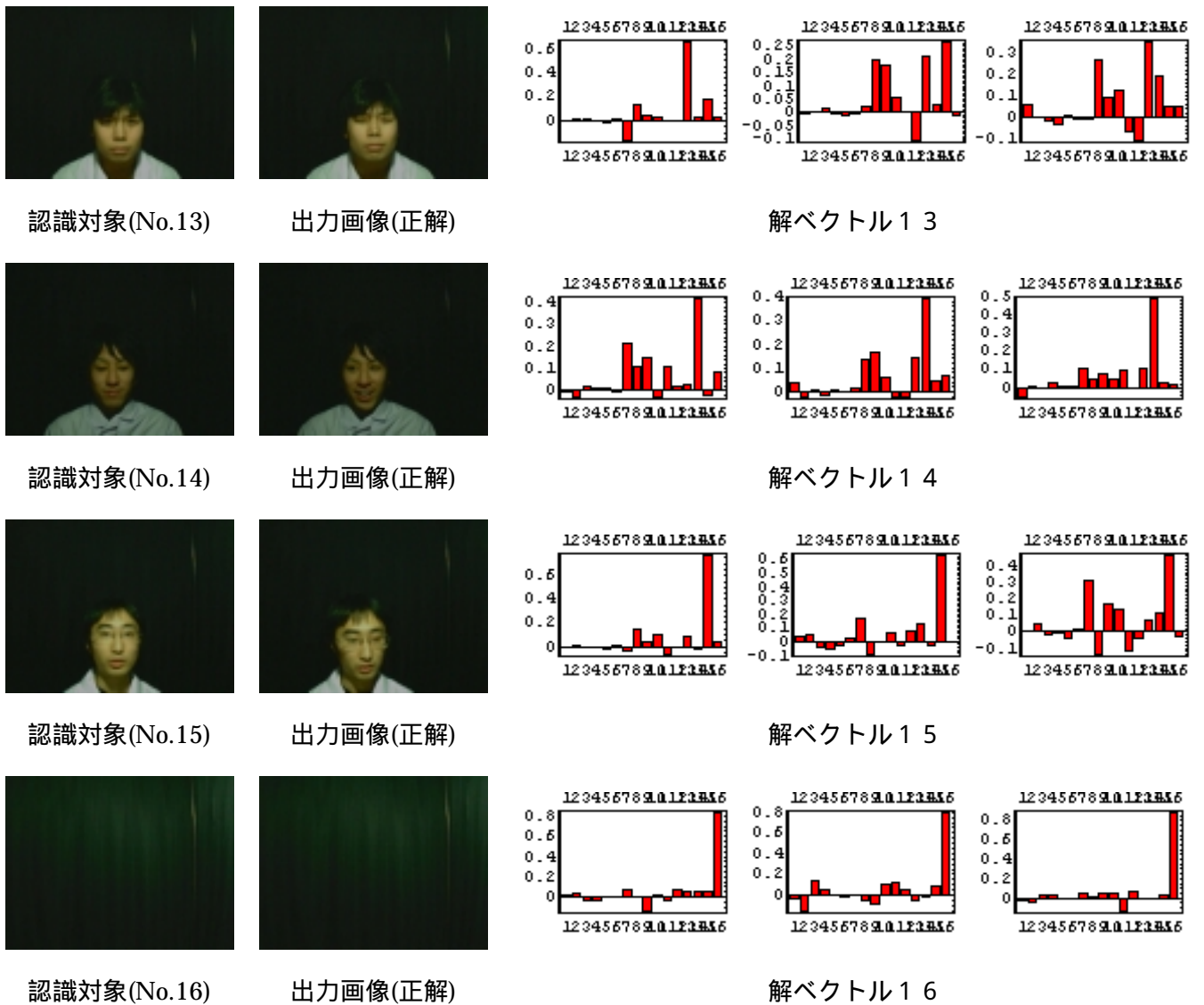


図 3 2 動画画像認識結果

動画画像の特徴量として固有パターンと時系列周波数特性を用い 3 台の CCD カメラで独立に認識を行ない、認識結果の多数決を取るという 3CCD 動画画像認識法で解を求めた。その結果、図 3 2 に示すように、認識対象 16 個中 15 個正解、1 個該当なしという認識結果になった。したがって、固有パターンと時系列周波数特性を合わせた特徴量は、従来の幾何学的的特異点抽出に基づいた画像認識では不可能であった人物の表情変化に対応し、更に撮影場所により異なる風景画像の認識を可能とした。

第4章 結 論

本論文では、人間の視覚情報処理能力を計算機によって実現する第一段階として画像認識ソフトの基幹部分の開発をおこなった。可視化情報は空間情報と時間軸情報によって構成されており、これらの特徴量を RGB 情報から抽出する方法を提案し、さらに、その特徴量を利用した動画画像認識手法について述べた。

第2章では可視化情報の空間情報と時間軸情報の特徴量である固有パターンと時系列周波数特性の抽出方法について述べた。空間情報の特徴量である固有パターンは、対象物の座標や角度情報、更に動き情報を削除した特徴量であることを示した。動画画像の時間軸情報は、RGB 成分のゆらぎをフーリエ変換し、得られたフーリエ・パワースペクトラムへゆらぎ解析手法を適用して時系列周波数特性を抽出した。この時系列周波数特性は物体固有の時間軸情報であり、異なる場所で撮影した風景画像でも同じ分類（例えば、化学反応、自然界の雲の動き、海の波の動きなど）のものであれば同様な特性を示す特徴量であることを示した。

第3章では可視化情報の空間情報と時間軸情報の特徴量を用いた動画画像認識について述べた。特徴量データの一致性評価は、最小自乗法による線形システム方程式の解ベクトルからおこなった。また3台の CCD カメラを用いて独立した認識をおこない、多数決を取ることでより認識率が向上し、妥当性判断が可能であることを示した。

本論文で提案した動画画像認識手法は、従来の幾何学的特徴を利用した画像認識では不可能であった人物の表情に変化がある場合においても認識可能であり、また風景など撮影場所により無限に存在する自然界の動画画像に対しても識別が可能であることを示した。

よって、本論文は、人間の視覚情報処理機能に近い動画画像認識エンジンを提案し、その妥当性を検証した。その結果、より人間の視覚情報処理能力へ近い動画画像認識が計算機によって可能であることが判明した。

参考文献

- [1] 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, pp.95-98, 2003
- [2] 若林健一, 早野誠治, 齋藤兆古, 國井利泰: 画像の固有パターンと画像認識への応用, 可視化情報学会誌, Vol.19, No.1, pp.91 ~ 94, 1999
- [3] 武者利光: ゆらぎの世界 自然界の $1/f$ ゆらぎの不思議
- [4] 武者利光, ゆらぎ現象研究会: ゆらぎの科学, 森北出版株式会社
- [5] 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, pp.243-246, 2002
- [6] 丸山和夫, 早野誠治, 齋藤兆古, 加藤和夫, 望月菜穂子, 山崎慶太, 堀井清之: 色情報による光源特徴の一解析法, 可視化情報学会誌, Vol.24, No.1, pp.223-226, 2004
- [7] G. Strang, "Linear Algebra and Its Applications", Academic Press, Inc., 1976.
- [8] Hisashi Endo, Seiji Hayano, Yoshifuru Saito and Kenzo Miya, "Generalized Vector Sampled Pattern Matching Method -Theory and Applications-", Studies in Applied Electromagnetics and Mechanics, Vol.23 Electromagnetic Nondestructive Evaluation (VI) (Eds. F. Kojima), IOS Press, The Netherland, pp.285-292, 2002.
- [9] 中島龍興, 近田 玲子, 面出 薫: 照明デザイン入門, 彰国社, 1995
- [10] 小杉幸夫, 高橋孝夫, 鈴木晴夫, 高倉公明, 鈴木真人, 池辺潤, 武者利光: $1/f$ ゆらぎのリズムを持つ電気刺激による除痛について, 電子通信学会, 1978