

有限差分を用いた動画像認識

小杉山 格 , 早野 誠治 , 齋藤 兆古(法政大学 大学院)

堀井 清之(白百合女子大学)

Dynamic Image Cognition by means of Finite Difference
Toru KOSUGIYAMA, Seiji HAYANO, Yoshifuru SAITO and Kiyoshi HORII

ABSTRACT

Most of the conventional image cognition methodologies are based on the geometrical singular points extraction of the target images. In human face image cognition, it is proposed that the method of discriminating the ratio of the distance between eyeballs and a mouth etc. as characteristics of each face information. This recognizing method has problems such as a definition or extracting method of the singular points, and change of the singular point distribution by change of facial expression. In order to overcome these difficulties, we have previously proposed a new image cognition methodology based on the color information extracting strategy. In the present paper, we have reported that the image cognition method based on the color information which composing each of the target images could recognize in fairly good accuracy. Furthermore, it has reported already that this method is generalized to the dynamic images cognition. As a result, it has been reported that the dynamic images was possible to recognize more precisely than static image.

In this paper, we propose the method of extracting the time variations of the color information by means of the finite differences. It is shown that this new methodology makes it possible to cognize the target objects having any background images.

Keywords: Dynamic image cognition, Eigen pattern, Finite Difference

1. はじめに

マイクロテクノロジーの広汎な普及と製品は全地球を席卷する電子機器を普及させ、ソフトウェア産業と呼ばれる、従来の人類では想像出来ない、電子形態の製品を製造する産業が生まれた。

初期の計算機は、いわゆる、職人芸的知識と操作技術を必要としたが、ハードとソフトウェアの進展が従来不可能と考えられていたヒューマンインターフェイスを実装可能とし、家庭の主婦から子供まで老若男女を問わず計算機を計算機と意識せずに使う環境整備は円熟期に入ったと言えるであろう。これは、人間が高度な知識を獲得せずに計算機を駆使できることを意味し、人間は学ばなくても計算機が賢く学んだことに他ならない。当然であるが、機械である計算機が自己学習して人間に近づいたのでは無い。計算機を創造した人間の英知がより高度な計算機を創造した結果である。

この意味で計算機を計算機と意識させないで使う時期を計算機の第一世代とすれば、計算機は人間にフレンドリーな機械から人間の友人そのものへ進展するのは自然であろう。すなわち、計算機の第 2 世代は人間の友人に

なり得る機械で作られた擬似人間氾濫時代と言えよう。

この第 2 世代の計算機を実現すべく誕生した学術分野の代表として、機械に学習機能を持たせんとするニューラルネットが上げられるであろう。ニューラルネットの基本的着想は人間の脳が持つ学習機能を計算機へ移植せんとする大きな試みである。

他方、筆者らは、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口眼球計画を発足し、研究・開発結果をこれまでに報告してきた。

筆者等の提案する画像認識手法は、画像認識の古典論では常套手段である幾何学的特異点間の比などへ着目せず、個々の画像を構成する画素の濃淡分布や赤・緑・青成分分布情報に基づいている。

この方法では画像を構成する背景画像も固有情報として勘案されるため、背景が同一でなければ使えない。また、幾何学的情報は、画像を構成する赤・緑・青成分の割合として維持されるため、間接的な情報となる。換言すれば、我々が提案する画像認識法は、ジグソーパズルを組む前に何の絵で有るか認識可能である利点を有する反面、濃淡情報や赤・緑・青色情報分布のみ一致すれば

同一画像と識別する欠点を有する。

本稿では、動画画像が時間軸方向へテイラー級数展開可能と仮定して、背景画像を無関係とする動画画像認識法を提案する。動画画像は必然的に時間軸方向へ離散化された複数のフレーム画像から構成される。このため、動画画像を時間軸方向へ解析的にテイラー級数展開可能であるから、時間軸方向へ有限差分法を適用し、離散値系テイラー級数展開を行い、テイラー級数展開の各項へ従来から筆者等が提案している合成画像法を適用し、背景が異なる動画画像認識を行う。

本稿は、此処で提案する方法の初期実験結果に関して報告するものである。

2. 画像の固有パターン

2.1 画像の構成要素

計算機のスクリーン上に描かれる画像は、色成分により赤(R)、緑(G)、青(B)の濃淡で表現されている。また、画像はスクリーン上で2次元平面状の画素の集合であり、

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

として表すことができる。

2.2 色成分固有ベクトル

静止画像の固有パターンは、画像中のR、G、B各成分の0から255までの濃度分布で与えられ、1つの画像に対して3組得られる。R成分分布をとすると、

$$I_R \in R_k \quad k = 1,2,3,\dots,m \times n \quad (2)$$

で与えられる。次にR成分分布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)で導く。 $Round[*]$ は括弧内の値を整数化する演算を表し、 $Max[*]$ は括弧内の集合における最大値の要素を求める演算を表す。

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

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。

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

式(4)で求められたE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色成分固有ベクトルを導出する。

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は色成分固有ベクトルを構成するR成分、G成分、B成分の3成分の集合であり、式(5)で表される。

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

2.4 動画画像の固有パターン

固有パターンEを動画画像へと拡張する。動画画像は複数のフレーム画像によって構成されるので、全フレームを通した固有パターンを抽出する必要がある。

そこで、動画画像を構成する全フレームを重ねるように1枚の静止画像へ合成した合成静止画像から固有パターンを抽出する。合成画像法により抽出された固有パターンは、動画画像中の対象物の移動や変形が起これいても同一の固有パターンを抽出することが可能である。Fig.1に同一対象物でフレームの順番が異なる動画画像から固有パターンを抽出した結果を示す。同図(c)より、固有パターンはいずれも等しいことが確認できる。

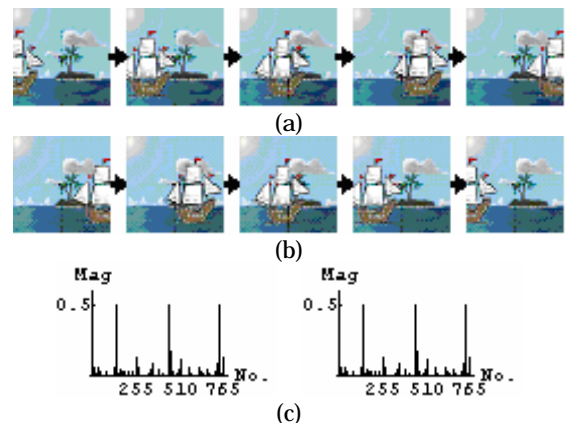


Fig.1 Eigen pattern of dynamic image
(a)Sample image No.1, (b)Sample image No.2,
(c)Eigen patterns(left:No.1, right:No.2)

3. 固有パターンの一致性評価法

本論文における動画画像認識とは、予め複数の動画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、認識対象として与えられる未知の入力動画画像から固有パターンを抽出し、データベースと入力間の固有パターンの一致性を線形システム方程式で評価し、入力動画画像をデータベースのいずれかの動画画像と同定するものである。R、G、B各成分が収められた固有パターンをEとして、n個の動画画像データからシステム行列Cを構築

すると、システム行列 C は、

$$C = [E_1, E_2, E_3, \dots, E_n] \quad (6)$$

で与えられる。従って、入力動画の固有パターンを E_x とすると、次の線形システム方程式が得られる。

$$E_x = C \cdot X \quad (7)$$

式(7)における X は、動画データベース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。また、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、本論文において最大 2 要素の固有ベクトルで構成されており、且つ、各固有ベクトルがダイナミックレンジ D で正規化した分布をヒストグラム化したものであるから、固有パターンの次数はダイナミックレンジ D の最大 2 倍である。この次数を m とすると、次数 m は、動画データベースに格納された固有パターンの個数 n よりも大きいと仮定すれば、システム行列 C が m 行 n 列の長方形行列となり、解ベクトル X の導出に最小自乗法を適用することができる。

$$X = [C^T C]^{-1} C^T E_x \quad (8)$$

式(8)で最大値をとる要素から、動画は認識される。

4. 有限差分法

4.1 離散値系テイラー級数展開の原理

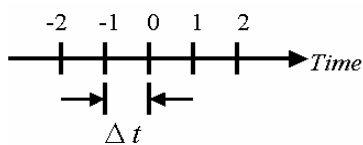


Fig.2 Time domain finite difference

Fig.2で、時間軸方向の偏微分を原点0で有限差分に置き換えることを考える。ここで、各点における $\phi_{-2} = (\phi(-2t))$, $\phi_{-1} = (\phi(-t))$, $\phi_0 = (\phi(0))$, $\phi_1 = (\phi(t))$, $\phi_2 = (\phi(2t))$ はテイラー級数に展開可能とする。 $\phi_0 = (\phi(0))$ として、 $\frac{\partial^n \phi}{\partial t^n}$ 以上の項が無視できる場合、 n 個の接点に関する次の一般式、

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{bmatrix} = \begin{bmatrix} 1 & \Delta t & \dots & (\Delta t)^{n-1} \\ 1 & 2\Delta t & \dots & (2\Delta t)^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & n\Delta t & \dots & (n\Delta t)^{n-1} \end{bmatrix} \begin{bmatrix} \phi_0 \\ \frac{\partial \phi}{\partial t} \\ \vdots \\ \frac{1}{(n-1)!} \frac{\partial^{n-1} \phi}{\partial t^{n-1}} \end{bmatrix} \quad (9)$$

を解き、 $\phi_0, \frac{\partial \phi}{\partial t}, \frac{1}{(n-1)!} \frac{\partial^{n-1} \phi}{\partial t^{n-1}}$ を求めることに帰する。式(9)右辺、 $n \times n$ 正方行列は、ヴァンデルモンデ型の行列であ

るため、解析的に逆行列を作ることが可能である。このように、有限差分法は任意の微係数を近似的に求めることができる。

4.2 動画認識への応用

背景画像が異なる動画を計算機上で認識しようとする場合、全フレームを重ねることによる合成画像法では、異なる背景情報のため固有パターンも等価とならない。動画は時系列上に静止画像を並べたものであるから、時間軸方向に対して、動画もテイラー級数展開可能であり、動画を構成するフレーム間の有限差分で、微係数を近似的に求めることが可能である。さらに、これは、フレーム間の差分は動画中の背景画像情報の削除を可能とし、移動している対象物の輪郭情報抽出につながる。また、差分を全フレームを通して行い、その結果を1枚の静止画像にまとめる合成画像法を実行することは、一種の積分と考えられる。

ここで、式(9)に対して、 $n=2$ 、すなわち、一階微分を求めることで、動画中から対象物の情報を抽出することを試みる。また、認識対象とする背景画像の異なる動画の一例をFigs.3, 4へと示す。

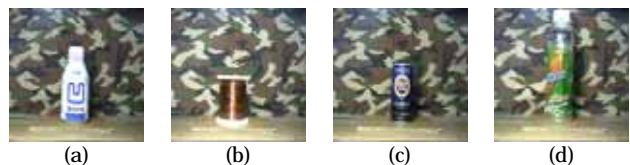


Fig.3 Database dynamic images

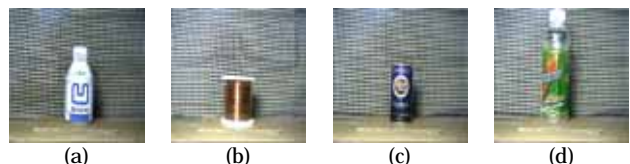


Fig.4 Testing dynamic images

5. 動画認識

5.1 一階差分による固有パターン抽出

式(9)より1階差分は

$$\frac{\partial \phi}{\partial t} = \frac{-\phi_1 + \phi_2}{\Delta t} \quad (10)$$

と求まり、式(10)の ϕ_1, ϕ_2 に動画を構成するフレーム画像を対応させて、差分画像を算出する。

一階差分では、動画を構成するフレーム数(=n)より1枚少ない $n-1$ 枚の差分画像が求まる。差分画像では変化のない背景画像は黒となって現れる。黒は画像を表現する数値データでゼロに対応する。したがって、異なる背景画像を持った動画でも、時間的に変化した対象物の情報が得られる。最後に、得られた差分動画に対して、2.4節で示した合成画像法を適用することで、時間差分画像の普遍量が抽出される。

5.2 カラー画像特性ベクトルによる固有パターン抽出

カラー画像は常に波長の依存した3成分で構成されている。可視光は波長の最も長い成分は赤 (R), 中間の波長は緑 (G), そして, 最も波長が短い青 (B) からなる。R,G,B各成分をそれぞれ, デカルト座標系の x, y, z の各成分に対応させれば画像から反射光ベクトル分布が得られる。すなわち, カラー画像は3次元のベクトルとして考えることができる。

3次元反射光ベクトルが時間と共に変化する画像が動画画像と言える。この反射光ベクトルを利用して動画画像認識を行う手法を, 此处で提案する。Fig.5は元画像, 反射光ベクトル分布画像, そして一次時間変化分の反射光ベクトル分布の例を示す。

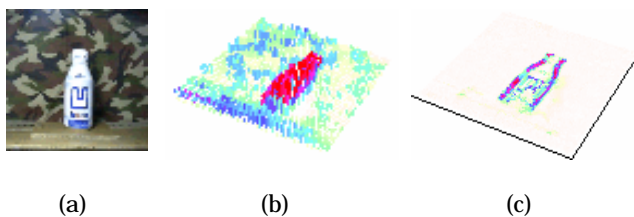


Fig.5 Color characteristic vectors
 (a) Original image
 (b) Color vector of original image
 (c) 1st difference image of color vector

反射光ベクトル分布画像を構成する各ベクトルは x, y, z 成分からなる。各成分に対して前節で述べた有限差分法で一階差分を行う。ベクトル間で差分された3次元のベクトルから, ベクトルの大きさの時間変化率を算出する。各画素におけるベクトルの大きさの時間変化率をモノクロ画素値とみなし, モノクロ動画画像を得る。このモノクロ動画画像から, ベクトルの大きさの時間変化率を0から255の256段階として, モノクロ固有パターンを抽出することができる。また, 各ベクトルの大きさの時間変化率が $x-y$ 平面となす角度に対しても, 角度変化固有パターンを抽出することが可能である。ここでは, Fig.5の動画画像から, 抽出されたベクトルの大きさの時間変化率と角度変化固有パターンの例をFig.6に示す。

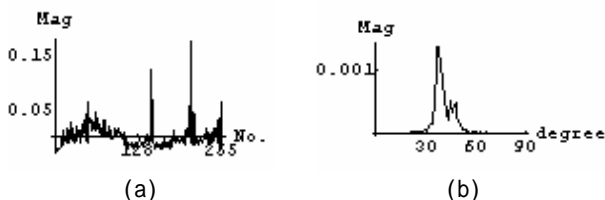


Fig.6 Eigen Pattern of Vectored Dynamic Images
 (a) Vector Magnitude Eigen pattern,
 (b) Vector Angle Eigen Pattern

5.3 ベクトル化動画画像を用いた動画画像認識

Fig.7に一階差分, 二階差分によって得られたベクトル大きさの一階と二階時間変化率固有パターンを同時に用

いた場合の認識結果を示す。

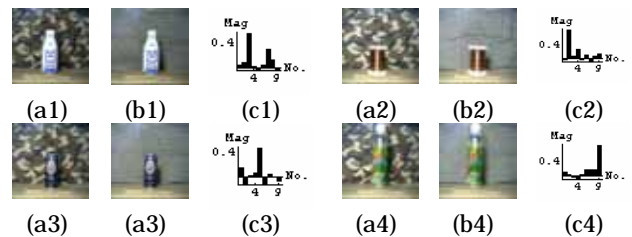


Fig.7 Cognized results by applying the 1st and 2nd time difference method to vectored images.
 (a1)-(a4)Cognized images, (b1)-(b4)Testing images, (c1)-(c4)Solution vectors

6 . まとめ

本論文では, 背景が異なる動画画像に対して, 有限差分法を適用することによって対象物の色情報時間変化量の固有パターンを抽出する手法を提案した。光を構成するR,G,B成分をそれぞれ x, y, z 直交座標系の値に対応させて得られるベクトル化動画画像を用いた全く新しい動画画像認識法を提案し, その妥当性を検証した。その結果, 背景画像が異なっても良好な認識結果を得ることが可能となり, 計算機の視覚情報処理能力が, より人間のへと近づいたということが確認できた。また, ベクトルの大きさを固有パターンとして構築することにより, 計算機内で保たれる情報量の削減も可能となった。

今後の課題はベクトル化画像で, 反射光ベクトルと $x-y$ 平面に対する角度で固有パターンを構成し, 動画画像認識を行うことである。

参考文献

- 1) 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画像認識, 可視化情報学会誌, Vol.22, No.1 (2002) pp.243-246.
- 2) 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画像認識, 可視化情報学会誌, Vol.23, No.1 (2003) pp.95-98.
- 3) 齋藤兆古: Mathematica によるウェーブレット変換, 朝倉書店 (1996)
- 4) 齋藤兆古: Mathematica による画像処理入門, 朝倉書店 (1998)
- 5) G.Strang: Linear Algebra and its Applications, Academic Press, Inc(1976)