

離散値系テイラー級数展開を用いた動画像認識

HAYANO, Seiji / SAITO, Yoshifuru / 斎藤, 兆古 / 早野, 誠
治 / KOSUGIYAMA, Toru / 小杉山, 格

(出版者 / Publisher)

法政大学計算科学研究センター

(雑誌名 / Journal or Publication Title)

法政大学計算科学研究センター研究報告 / Bulletin of Computational Science
Research Center, Hosei University

(巻 / Volume)

18

(開始ページ / Start Page)

143

(終了ページ / End Page)

147

(発行年 / Year)

2005-03-22

(URL)

<https://doi.org/10.15002/00025046>

離散値系テイラー級数展開を用いた動画像認識

小杉山 格 早野 誠治 齋藤 兆古
法政大学大学院工学研究科

従来の画像認識法の大部分は画像の幾何学的特異点抽出に基づいている。より具体的に人間の顔画像認識を考えれば、眼球間と眼球と口間の距離の比などを個々の顔情報の普遍量として顔画像識別を行う。しかし、このような従来の幾何学的特異点抽出に基づく顔画像認識法は、特異点の定義やその抽出法、顔表情の変化による特異点の変化などの問題点がある。このような従来の画像認識法の問題点を克服するために、筆者らは画像の色情報を利用した全く新しい画像認識法を提案し、画像を構成する色情報に基づくこの新手法が画像の幾何学的な情報を含めるため、良好な精度で画像認識を可能とする事を報告した。さらに、この新しい方法を動画像認識へ拡張し、静止画像に比較してより高精度な画像認識が可能である事も既に報告した。本論文では、動画像のみが含有する対象物の移動速度や加速度情報を色情報として抽出する方法を提案し、その妥当性を検討する。

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、職人芸の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1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[1]。

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工眼球計画を発足し、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・青色成分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像認識法を提案する。動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画像に対して3組得られる。R成分分布を例にとると、

$$I_R \in R_k \quad (2)$$

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。

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

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[2]。

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

2.4 動画の固有パターン

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

そこで、動画を構成する全フレームを重ねないように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、動画中の対象物の移動や変形が起きている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。

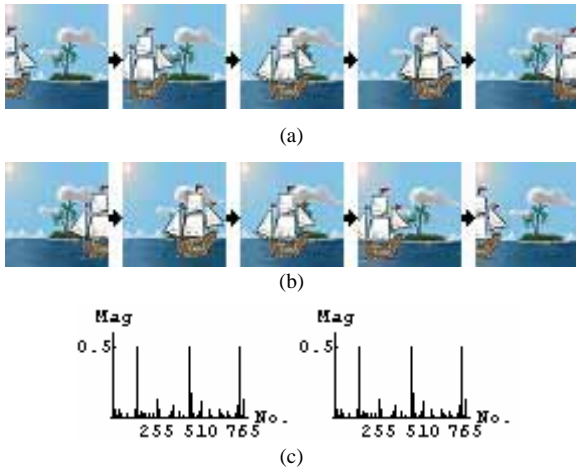


Fig.1 Eigen Pattern of Dynamic Image
(a)Sample Image No.1 , (b)Sample Image No.2
(c)Eigen Patterns(left:No.1 , right:No.2)

3. 固有パターンの一致性評価

本論文における動画認識とは、予め複数の動画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が

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、入力動画をデータベースのいずれかの動画像と同一とするものである。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、動画像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3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3倍である。この次数を m とすると、次数 m は、動画像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[3]。

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

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

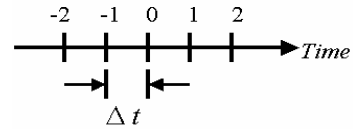


Fig.2 Time Domain Finite Difference

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 $t_2 = (-2 \Delta t)$ 、 $t_1 = (-\Delta t)$ 、 $t_0 = (0)$ 、 $t_1 = (\Delta t)$ 、 $t_2 = (2 \Delta t)$ はテイラー級数に展開可能とする。 $t_0 = (0)$ として、 $\frac{\partial^n \phi}{\partial t^n}$ 以上の項が無視できる場合、有限差分法は、 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、 ϕ_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、 $n \times n$ 正方形行列は、ヴァンデルモンド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]。そして、有限差分法が適用可能な条件は、テイラー級数展開可能であることだけである。

4.2 動画におけるテイラー級数展開

動画は時系列上に静止画像を並べたものであるから[5]、時間軸方向に対して、動画もテイラー級数展開可能である。式(9)で示した一般式を用いると時刻 a における関数は式(10)で表せる[6,7]。

$$\phi = \phi_a + (\Delta t) \frac{\partial \phi}{\partial t} + (\Delta t)^2 \frac{1}{2!} \frac{\partial^2 \phi}{\partial t^2} \quad (10)$$

式(10)における左辺は動画に、第一項目は時刻 a におけるフレーム画像、第二項目は時間に対して変化分を、そして第三項目はさらに時間に対しての変化分に対応する。また、式(10)の時間刻み幅 Δt は、現在のCCDカメラでは一時間に約30枚撮影可能であるが、ここでは $\Delta t=1$ として考える[8]。Fig.3に、式(10)に対応させて得られた画像を示す。ここで、式(10)における時刻 a は、50枚で構成される動画の28枚目のフレームとした。

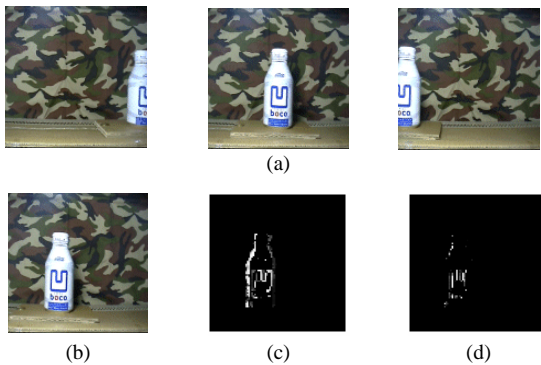


Fig.3 Taylor Series Expansion for Dynamic Image
(a)Left hand side on Equation (10)
(b)the First Term , (c)the Second Term , (d)the Third Term

4.3 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。4.2節で示したように、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、フレーム間の差分が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、差分を全フレームに通して行い、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、一階微分、二階微分を求めらることで、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.4 , 5 へと示す。



Fig.4 Database Dynamic Images



Fig.5 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

$$\frac{\partial \phi}{\partial t} = \frac{-\phi_1 + \phi_2}{\Delta t} \quad (11)$$

とつまり、式(11)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。ここで求められた一階差分画像と、その固有パターンを Fig.6 へと示す。

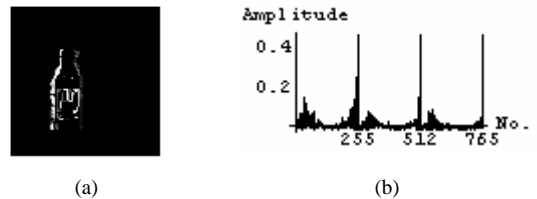


Fig.6 Difference image and Its Eigen Pattern
(a)1st order difference image
(b)Eigen Pattern extracted from 1st order difference

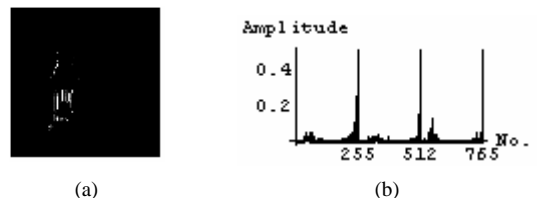


Fig.7 Difference image and Its Eigen Pattern
(a)2nd order difference image
(b)Eigen Pattern extracted from 2nd order difference

5.2 二階差分による固有パターン抽出

5.1 で示した一階差分と同様に、式(9)から二階差分を求めれば式(12)のようになる。

$$\frac{\partial^2 \phi}{\partial t^2} = \frac{-\phi_1 + 2\phi_2 - \phi_3}{(\Delta t)^2} \quad (12)$$

式(12)の ϕ_1, ϕ_2, ϕ_3 にフレーム画像を対応させることで二階差分画像を求めることができる。一階差分画像が時間的に変化した画像ならば、二階差分画像は、さらにその時間的変化分を抽出したことを意味する。このため、一階差分画像は速度、二階差分画像は加速度の意味を持つといえる。二階差分画像と、式(11)の演算を全てに施し、二階の差分動画から求められた固有パターンを Fig.7 へと示す。Fig.6 に示した一階差分による固有パターンと比べれば、二階差分で得られた固有パターンのヒストグラムは小さいことが分かる。

5.3 認識結果

5.1, 5.2 で示した差分画像からの固有パターンを用いて、Fig.3, 4 の背景の異なる動画像での認識を試みた結果を Fig.7 で示す。ここで示される認識結果は、一階、二階差分画像からの固有パターンを繋げた 1530 次の固有ベクトルを用いた。

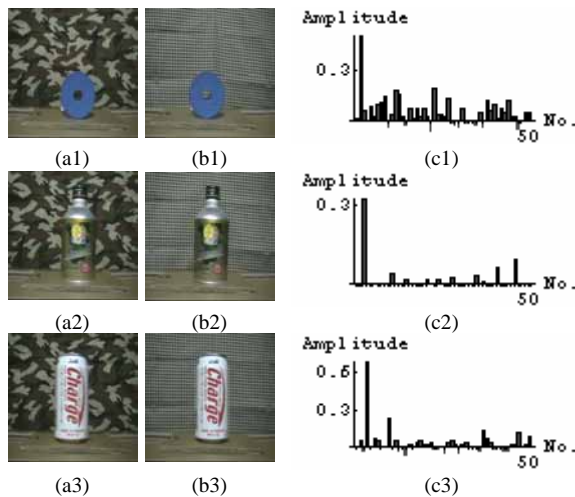


Fig.8 Cognized results by applying the 1st and 2nd time difference method to vectored images
(a1)-(a3) Cognized images, (b1)-(b3) Testing images
(c1)-(c3) Solution vectors

また、各固有パターンでの結果を比較するために、式(12)で示される認識率を定義する。認識率とは、全供試動画像数に対する認識に成功した動画像数を百分率で表したものである。Table.1 に各固有パターンにおける認識率を示す。Table.1 において、Order は固有パターンの次数を表している。

$$CognitionRate = \frac{Number\ of\ Successful\ images}{Number\ of\ Testing\ images} \times 100 \quad (13)$$

Table.1 Cognition Rates

Eigen Pattern	Order	Cognition Rate
Composite	765	26 %
1 st difference	765	78 %
2 nd difference	765	72 %
Composite and 1 st difference	1530	82 %
Composite and 2 nd difference	1530	82 %
1 st difference and 2 nd difference	1530	94 %
Composite, 1 st and 2 nd difference	2295	92 %

6. まとめ

本論文では、動画像に対しても離散系テイラー級数展開可能であることを明らかにした。そして、背景が異なる動画像に対して、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。その結果、背景画像が異なっていても差分動画像から得られた固有パターンを用いれば、良好な認識結果を得ることが可能であり、計算機の視覚情報処理能力が、より人間のそれへと近づいたということが確認できた。

今後の課題は固有パターンによる画像からの普遍量において同面積な色情報をもつ、異なる対象物は同一とみなされてしまう問題がある。幾何学的な情報を画像の周波数成分から普遍量として抽出し、認識を試みることである。

参考文献

- [1] 森文彦, 花岡かほる, "エキスパートシステム構築技法入門", オーム社, 1990 年.
- [2] 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 年.
- [3] G.Strang, "Linear Algebra and its Applications", Academic Press, 1976 年.
- [4] 齋藤兆古, "Mathematica による数値シミュレーション", 大学院参考教材, 2003 年.
- [5] 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 年.
- [6] 平野矢太郎, 田中尚夫, "工科系のための微分積分学", 裳華房 第 23 版 pp.28-32, 2000 年.
- [7] 山内二郎, 宇野利雄, 一松信, "電子計算機のための数値計算法", 培風館, 1972 年.
- [8] 小杉山格, 早野誠治, 齋藤兆古, 堀井清之, "有限差分を用いた動画像認識", 可視化情報学会誌, Vol.24, No.1 pp.247-250, 2004 年.

キーワード .

動画像認識、固有パターン、有限差分

Summary .

Dynamic Image Cognition by means of Discrete Taylor Series Expansion

Tohru Kosugiyama Seiji Hayano Yoshifuru Saito
Graduate School of Engineering , Hosei University

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 previous 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 been 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