環境情報の構造化と獲得情報量の最大化を規範とした学習制御の研究

A Learning Algorithm for Structuring and Maximizing of Envoronmental Information

1071024 (登録番号)

研究代表者 (横浜国立大学 准教授) 藤 本 康 孝

研究の目的

ヒトは外界の情報の約 7 割〜9 割を視覚から得ているといわれている。ロボットの可能性を広げるためには、視覚情報の学習と認識技術の向上が必要不可欠となってくる。画像認識技術は、コンピュータや CCD カメラの進化とともにめざましく発展しており、産業界では不良品検出、指紋認証、顔認識システム、交通情報の調査などの分野で活用されている。しかし、コンピュータによる認識技術は特定の限られた環境下でのみ有効であり、未だ人間のように柔軟に環境を認識し理解するシステムには至っていないのが現状である。

本研究では、より柔軟な自律ロボットの実現を目指し、外界から得られる情報 (環境情報) を再構築し、その情報に基づいて行動を決定するロボットの開発を目的とする。環境情報の学習に主成分分析 (Principal Components Analysis; PCA) を適用し、低次元化された確率分布モデルとして情報を蓄積する。逐次的に入力される情報に基づいて、PCA における固有空間を効率的に更新する手法を検討する。さらに、その固有空間においてクラスタリングを行い、状況認識を行う方法について検討する。

PCA は、Kirby らによって顔画像の圧縮に用いられ、Turk らによって顔画像の向きや大きさが一定であれば、高い認識率を得ることが示され、それ以来画像認識に多く用いられている。また、顔画像だけでなく、3 次元での物体の認識にも適用する方法も提案されている。PCA では、画素値の共分散行列から固有値を求める必要があるが、この計算には多くの時間を要する。そこで本研究では、計算時間の短縮を目的としてニュートン法を用いた方法を提案する。この方法では、それまでに得られた固有値と固有ベクトルの誤差分のみを修正することで、効率良く計算することができる。これにより、得られた画像をリアルタイムに認識し、またその情報を逐次学習することを目指す。

研究の内容、成果

1. 主成分分析 (PCA)

画素数 N の静止画像の各画素の濃度値を $y_k^{(i)}$, $(i=1,\cdots,n,\ k=1,\cdots,N)$ で表す。n は画像枚数を表す。画素ごとに正規化を行うことで認識率が向上することが期待されるため、平均 0、分散 1 に正規化を行ってから PCA を行う。

全サンプル画像の画素ごとの平均 c_k と標準偏差 σ_k は、$c_k = \frac{1}{n}\sum_{i=1}^{n} y_k^{(i)}$ $\sigma_k^2 = \frac{1}{n}\sum_{i=1}^{n}(y_k^{(i)} - c_k)^2$ として求まる。これを用いて、正規化を行った画像ベクトル $\boldsymbol{y}_i$ は、次のように表せる。

$$\boldsymbol{y}_i \equiv \left[\frac{y_1^{(i)}-c_1}{\sigma_1}, \frac{y_2^{(i)}-c_2}{\sigma_2}, \cdots, \frac{y_N^{(i)}-c_N}{\sigma_N}\right]^T \quad (1)$$

このとき、画像集合の共分散行列 $\boldsymbol{P}$ は $\boldsymbol{P} = \frac{1}{n}\sum_i^n(\boldsymbol{y}_i\boldsymbol{y}_i{}^T)$ と求まる。この共分散行列の固有

値を求め、対応する固有ベクトル $\boldsymbol{v}$ を基底ベクトルとすることで固有空間に投影することができる。このとき、固有ベクトルを主軸とし、対応する固有値をその主軸方向の標準偏差とする多次元分布を考えることができ、主要な固有値と主軸で分布を近似することができる。共分散行列は画素の 2 乗という大きな行列であり、普通に計算することは難しいが、十分小さい固有値は無視することができることから、画像の次元を数次元まで落すことができる。

固有空間上では、2 枚の画像相関 $\boldsymbol{y}_i^T\boldsymbol{y}_j$ が大きければ 2 点間距離が小さくなるため、画像判別に利用できる。

2. オンライン学習での固有値計算

逐次的に画像を取得する場合、画像を 1 枚得るごとに固有値と固有ベクトルを求める必要がある。しかし、毎回計算するには全ての画像データを所持していなければならず、また計算時間もかかるため効率が悪い。そこで、効率の良い共分散行列の更新法、またニュートン法を用いて入力画像に対する固有値、固有ベクトルの変化分のみを計算し修正する方法とその逆行列計算を近似的に行う方法を提案する。

共分散行列の更新　共分散行列を求めるために、取得した画像データを蓄積していかなければならないが、これは膨大なデータ量となり、学習する画像枚数に限界ができてしまう。そこで、それまでの共分散行列と入力画像から、新しい共分散行列を更新する。i 枚目までの画素毎の平均を $c^{(i)}$、標準偏差を $\sigma^{(i)}$ と表すと、n 枚目の新規画像 $y^{(n)}$ を取得したとき、画素毎の平均と標準偏差は、それまでの平均と標準偏差を用いて次式で更新することができる。

$$c_k^{(n)} = \frac{n-1}{n}c_k^{(n-1)} + \frac{y_k^{(n)}}{n} \tag{2}$$

$${\sigma_k^{(n)}}^2 = \frac{n-1}{n}\left({\sigma_k^{(n-1)}}^2 + {c_k^{(n-1)}}^2\right) + \frac{{y_k^{(n)}}^2}{n} - {c_k^{(n)}}^2 \tag{3}$$

と表せる。また、i 枚目までの共分散行列を $P^{(i)}$ とすると、n 枚目までの共分散行列 $\boldsymbol{P}^{(n)}$ は、$n-1$ 枚目までの共分散行列 $\boldsymbol{P}^{(n-1)}$ を用いて表すことができる。その (a,b) 要素を $P_{a,b}^{(n)}$ としたとき、

$$\begin{aligned} P_{a,b}^{(n)} &= \frac{1}{\sigma_a^{(n)}\sigma_b^{(n)}}\Big(P_{a,b}^{(n-1)}\sigma_a^{(n-1)}\sigma_b^{(n-1)} + y_a^{(n)}y_b^{(n)} \\ &\quad -nc_a^{(n)}c_b^{(n)} + (n-1)c_a^{(n-1)}c_b^{(n-1)}\Big) \end{aligned} \tag{4}$$

と更新していくことができる。これにより、画像枚数が増えても共分散行列と平均、標準偏差を所持しておけば更新し続けることができる。

ニュートン法　逐次的に画像を取得する場合、画像を 1 枚得るごとに固有値と固有ベクトルを求めなければならない。そこで、ニュートン法を用いて入力画像による変化分のみを計算し修正する方法を提案する。

$\boldsymbol{P}$ の固有値、固有ベクトルをそれぞれ λ_i、$\boldsymbol{v}_i (i = 1, \cdots, m : m$ は必要な固有値の数$)$ とする。このとき、$\lambda_i\boldsymbol{v}_i = \boldsymbol{P}\boldsymbol{v}_i$、および、正規化した固有ベクトルについて $\boldsymbol{v}_i^T\boldsymbol{v}_i = 1$ が成り立つ。$n-1$ 枚目までの共分散行列 $P^{(n-1)}$ の i 番目の固有値を $\lambda_i^{(n-1)}$、固有ベクトルを $v_i^{(n-1)}$ とし、$P^{(n)}$ の i 番目の固有値、固有ベクトルの近似解として $\lambda_i = \lambda_i^{(n-1)}, \boldsymbol{v}_i = \boldsymbol{v}_i^{(n-1)}$ とおくと、式誤差

$$\boldsymbol{e}_i = \begin{bmatrix} (\lambda_i I - \boldsymbol{P}^{(n)})\boldsymbol{v}_i \\ \boldsymbol{v}_i^T\boldsymbol{v}_i - 1 \end{bmatrix} \tag{5}$$

が存在することになる。固有値、固有ベクトルを求めるために方程式 $\boldsymbol{e}_i = 0$ を解けば良い。ヘッセ行列 $\boldsymbol{H}_i$ は、

$$\boldsymbol{H}_i = \begin{bmatrix} \dfrac{\partial \boldsymbol{e}_i}{\partial \boldsymbol{v}_i^T} & \dfrac{\partial \boldsymbol{e}_i}{\partial \lambda_i} \end{bmatrix} = \begin{bmatrix} \lambda_i I - \boldsymbol{P}_i^{(n)} & \boldsymbol{v}_i \\ \boldsymbol{v}_i^T & 0 \end{bmatrix} \tag{6}$$

と求められるので、ニュートン法を用いて、

$$\begin{bmatrix} \boldsymbol{v}_i \\ \lambda_i \end{bmatrix} := \begin{bmatrix} \boldsymbol{v}_i \\ \lambda_i \end{bmatrix} - \alpha \boldsymbol{H}_i^{-1}\boldsymbol{e}_i \tag{7}$$

のように、誤差ベクトルが $\boldsymbol{e}_i = 0$ となるまで更新することで、固有ベクトルを求めることができる。α はステップ幅を表わす。実際は逆行列を直接計算することはせずに、線形方程式を共役勾配法により解くことで (7) 式を計算する。

近似逆行列による方法　ニュートン法を用いると、誤差分のみの計算となり早い収束が望めるが、(7) 式の第 2 項 $\boldsymbol{H}^{-1}\boldsymbol{e}$ の線形方程式の計算に時間を要する。そこで、この部分の計算時間の効率化を考える。i 番目の固有値に対応するヘッセ行列 $\boldsymbol{H}_i$ の固有値、固有ベクトルは、共分散行列 $\boldsymbol{P}$ の固有値 λ と固有ベクトル $\boldsymbol{v}$ を用いて表すことができる。$\boldsymbol{H}_i$ の固有ベクトルからなる行列 $\boldsymbol{V}_i$ および固有値を対角に並べた行列 $\boldsymbol{\Lambda}_i$

$$\boldsymbol{V}_i = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \frac{1}{\sqrt{2}}\boldsymbol{v}_i & \cdots & \boldsymbol{v}_m & \frac{1}{\sqrt{2}}\boldsymbol{v}_i \\ 0 & 0 & \cdots & \frac{1}{\sqrt{2}} & \cdots & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \tag{8}$$

$$\boldsymbol{\Lambda}_i = \mathrm{diag}\,[\lambda_i-\lambda_1,\cdots,\lambda_i-\lambda_{i-1}, 1,\lambda_i-\lambda_{i+1},\cdots,\lambda_i-\lambda_m,-1] \tag{9}$$

を用いて、ヘッセ行列は $\boldsymbol{H}_i \simeq \boldsymbol{V}_i\boldsymbol{\Lambda}_i\boldsymbol{V}_i^T$ と近似できる。したがって、ヘッセ行列の近似逆行列は (8) 式、(9) 式より、

$$\begin{aligned} \boldsymbol{H}_i^{-1} &\simeq \boldsymbol{V}_i\boldsymbol{\Lambda}_i^{-1}\boldsymbol{V}_i^T \\ &= \begin{bmatrix} \displaystyle\sum_{j\neq i}^{m} \frac{1}{\lambda_i-\lambda_j}\boldsymbol{v}_j\boldsymbol{v}_j^T & \boldsymbol{v}_i \\ \boldsymbol{v}_i^T & 0 \end{bmatrix} \end{aligned} \tag{10}$$

と表わせる。ただし、m はヘッセ行列 $\boldsymbol{H}_i$ を十分に再現するために必要な固有値の数である。

(10) 式を (7) 式のニュートン法の更新式に代入し、誤差ベクトル $\boldsymbol{e}_i = 0$ となるとき、固有値、固有ベクトルが求まる。

3. クラスタリング

固有空間の投影した後、その対象物がどのような画像かを判別する必要がある。事前学習を行わない教師なしデータを自動的に分類するクラスタリングの方法として、本研究では階層的手法を用いた。階層的手法は、全ての点が 1 個の対象だけを含む別のクラスタである状態を初期状態として、それぞれのクラスタ間の類似度を計算し、次々と併合していくことで任意のクラスタ数に分類する方法である。クラスタ間の類似度の計算法には最短距離法を用いた。

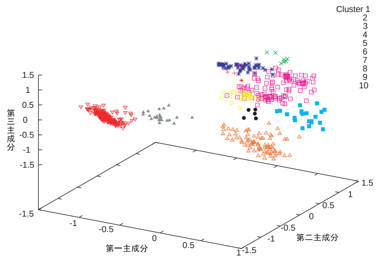

図 1: 560 枚までのクラスタリング結果

図 2: 主成分画像

4. 実験

実験条件　日常環境での PCA による学習実験を行った。実験では、日常環境として研究室内画像を 560 枚を用い、PCA により固有空間と主成分画像を求めた。PCA を行う際には画像を (30×40)pix に縮小して用いた。

結果　図 1 に、第 1〜第 3 成分での固有空間におけるクラスタリング結果と、またそのときの固有値と主成分画像を示す。クラスタリングでは画像枚数の 3 以下のクラスタを孤立点とし、残りのデータ数が全体 90%を満たすまで最短距離法によるクラスタの融合を繰り返した。最短距離法の計算には、第 5 主成分までの固有空間を用いた。 図 2 の主成分画像からは、第 1 主成分と第 2 主成分では奥に人影が、また第 3 主成分では手前に人影が見える。このときの各クラス

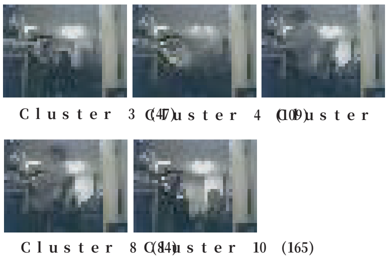

図 3: クラスタの平均画像 (括弧内は要素数)

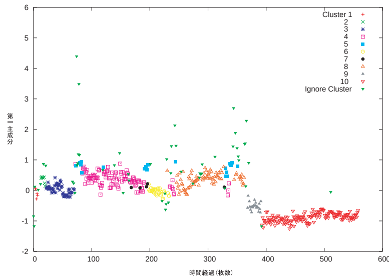

図 4: 入力画像が属するクラスタの時間変化

タの平均画像は図 3 のようになった。Cluster3,4 では奥の方の人が変化し、Cluster6,8 では手前の人の姿勢が変化している様子が分かる。このクラスタの時間変化を表したものが図 4 であり、これによって手前の人物は 200 枚付近で現れて 350 枚付近で退席している、など対象の行動パターンを知ることができる。

計算時間の比較　650 枚の画像に 1 枚を追加し、固有空間を再構成したときの各手法の計算時間を表 1 に示す。 ベキ乗法は求める固有値と次の固有値の差が小さい場合に収束に時間がかかるという性質があるため、主成分の数が多くなるほど時間がかかる傾向がある。ニュートン法を用いた場合、計算時間にばらつきが無く安定していることが分かる。さらに、逆行列計算に共役勾配法を用いたニュートン法と近似逆行列を用いた提案手法を比較すると、平均で約 1/6 に計算時間を短縮することができた。この点から大量の画像を扱う場合に、提案手法は非常に有効であると言える。

表 1: 固有値と画像の学習にかかる時間 (秒)

成分	固有値	べき乗法	ニュートン法 (共役勾配法)	ニュートン法 (近似逆行列)
1st	1922.7	1.58	2.21	0.551
2nd	328.0	3.08	2.52	0.318
3rd	155.8	3.20	2.74	0.400
4st	93.2	5.28	2.85	0.518
5nd	57.8	57.40	3.96	0.656

今後の研究の方向、課題

本研究では、PCA を用いて画像情報のクラスタリングを行い、また未知画像に対して PCA の固有空間を効率よく更新する方法について検討した。この構造化した環境モデルを確率分布として捉え、情報理論的意味での情報量を最大化するようにロボットの行動を生成する方法についても検討を行っている。これまでのところ比較的単純なアクティブビジョンシステムを対象としているが、これを一般化し、汎用性を実現することが課題である。

成果の発表、論文等

1. 藤本康孝, 加藤英夫: ロボットビジョンによる環境情報の構造化の研究, 電気学会産業計測制御研究会資料, IIC-08-24, pp. 21–25, (2008.3)
2. 李 雪, 藤本康孝: アクティブビジョンによるオンライン学習に関する研究, 電気学会産業計測制御研究会資料, IIC-08-89, pp. 17–22, (2008.3)
3. 七田和典, 藤本康孝: ステレオビジョンとオプティカルフローを併用した 3 次元環境認識の高速化に関する研究, 日本ロボット学会学術講演会, 2N12, (2007.9)
4. 李 雪, 藤本康孝: アクティブビジョンシステムによる環境情報の構造化に関する研究, 日本ロボット学会学術講演会, 2N16, (2007.9)