固有空間法はなぜうまく働くか

出口光一郎　岡谷貴之
東北大学大学院情報科学研究科
{kodeg,okatani}@fractal.is.tohoku.ac.jp

概要：最近、画像解析に広く固有空間法が応用され、実用的な手法としての立場を確立している。しかし、画像そのものをデータとして見たときには、サンプル数に比べて次元数が遥かに大きいという、本来の固有空間法を適用するデータとは違った様相を持つ。ここでは、このようなデータに固有空間法を適用することで、副次的な効果としての次元数を下げる、ということが画像解析におけるモデルあてはめに本質的な役割をしているのではないかと論じる。

ここでは、画像からの姿勢推定というタスクを例に、画像データそのものに、直接的に姿勢パラメータを結びつける手法をまず述べる。続いて、同じことを、固有空間法を用いて実現し、その対比をすることでこのようなデータの次元数がサンプル数よりはるかに大きいときに何が起きているのかを見てみる。そして、いわゆる「過学習の理論」の立場からこの対比を解釈してみる。

キーワード　固有空間法、画像からの姿勢推定、汎化能力、過学習の理論

Why does the eigen space method work well?

Koichiro Deguchi　and　Takayuki Okatani
Graduate School of Information Sciences, Tohoku University
{kodeg,okatani}@fractal.is.tohoku.ac.jp

Abstract: The eigen space method has established its position among computer vision techniques, recently, as a tool for robust image analysis as well as efficient image data compression method. However, image data itself has a significant difference from statistical data to which the conventional eigen space method is conceptually applicable, because its dimensionality is much larger than the number of samples which are used to construct the eigen space. Here, we claim that the reduction of the dimensionality which is carried out in the process of the eigen space method plays an important role for model fitting to the image data.

In this article, we introduce the image based pose estimation task as an example to apply the eigen space method. We make comparison between the eigen space method and the direct interpretation method where a linear model combining poses and images. Then, we discuss the comparison from the viewpoint of the theory of over-learning.

Key words: Eigen space method, Image based pose estimation, Generalization, Over-learning.

1　はじめに

最近、画像解析に広く固有空間法が応用され、実用的な手法としての立場を確立している [1]。とくに、パラメトリック固有空間法と呼ばれる手法は、村瀬ら [2] によって、コンピュータビジョンの問題での簡便かつ強力な手法としての有効性が示され、広く用いられている。

固有空間法では、M 次元データのサンプルが、図 1のように、実質的にはずっと次元の低い部分空間に分布している、あるいは分布しているとみなせる場合、第一にはデータ圧縮として、第二には各データに共通する大域的な性質を抽出して引き去り、サンプル内での個体差を強調することで、より有効なデータ解析を実現するものである。統計的にデータを見たときに主成分分析をすることと同等である。

しかしその応用中には、本来の原理としての固有空間法の説明には合致していない様な画像群に対する応用も多くある。すなわち、画像そのものをデータとして見たときに、サンプル数に比べて次元数が遥かに大きい。一般には、次元数 M は画像の画素数にあたり、数万以上であるのに対して、サンプル数はせいぜい数百である。さらに、画像データの空間ではそれぞれのサンプルがほとんど独立であるとみなせるような場合が多い。図 1で言えば、M 次元空間に、極端にまばらにデータが分布していて、この図に示すような状況にない。これは、本来の固有空間法を適用するデータとは違った様相を持つ場合である。それらでも、パラメトリック固有空間法は機能し、何らかの有効な結果を与えている。

ここでは、画像からの姿勢推定というタスクを例に、このようなデータにパラメトリック固有空間法を適用することで、副次的な効果としての次元数を下げる、ということが画像解析におけるモデルあてはめに本質的な役割をしているのではないかと論じる。

そのため、画像データそのものに、直接的に姿勢パラメータを結びつける手法をまず述べる。続いて、同じことを、固有空間法を用いて実現し、その対比をすることでこのようなデータの次元数がサンプル数よりはるかに大きいときに何が起きているのかを見てみる。

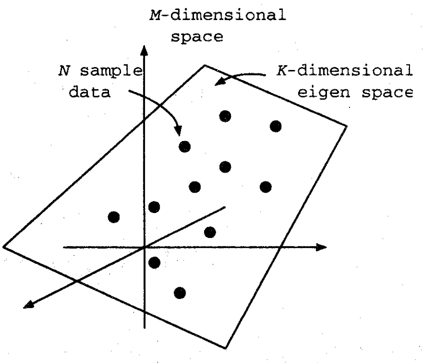

図 1: サンプルデータの分布とその固有空間

そして、いわゆる「過学習の理論」の立場からこの対比を解釈してみる。

2　画像からの姿勢推定問題の線形解法

2.1　画像と姿勢とその推定

ここでの例としての問題は、一枚の与えられた画像から物体（あるいは、カメラ）の姿勢を言い当てるというものである。一般的には、連続的に変化するパラメータがあって、そのパラメータに応じて画像が変化するとき、画像からパラメータを復元する。

具体的には、ターンテーブルに乗せられて対象が回転しているときの画像 N 枚をサンプルとして、まず、与える。ターンテーブルの回転角 ϕ が各画像に対する姿勢パラメータである。ただし、ここでは、$(\theta_1 = \cos\phi,\ \theta_2 = \sin\phi)$ を姿勢パラメータとする。

画像のサイズを、縦、横がそれぞれ I, J ピクセルであるとし、$M = I \times J$ とする。そして、画像を、各ピクセルの濃淡値を並べた M 次元のベクトル $\boldsymbol{x} = (x_1, x_2, \cdots, x_M)^\top$ で表わす。（以下、簡単のため $\boldsymbol{x}$ の平均は $\mathbf{0}$ とする。）したがって、問題は、M 次元のデータ（画像）$\boldsymbol{x}$ から、2 次元の姿勢パラメータ $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top$ を推定することがここでの問題となる。

これに対して、まず、姿勢パラメータが、$\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_N$ であるときの N 枚の画像、$\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N$ をサンプルとして与える。この前提の元

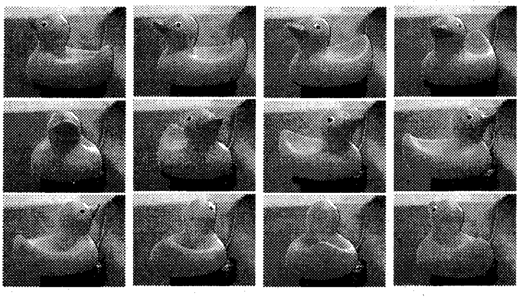

図 2: 姿勢推定問題におけるサンプル画像の例。対象を回転させながら撮影した画像の340枚のうちの１２枚を示す。このような画像の組をその姿勢パラメータとともにサンプルとして与える。その後、任意の姿勢での画像から、その姿勢パラメータを推定する。

で、新たに与えられた画像 $\boldsymbol{x}$ （上記のサンプル画像には含まれない）に対して、この画像が得られるであろう姿勢 $\boldsymbol{\theta}$ を求める。

図2に示すのは、後述の実験で用いたサンプル画像の一例である。このような画像の組をその姿勢パラメータとともにサンプルとして与える。その後、任意の姿勢での画像から、その姿勢パラメータを推定する。

2.2 見え方に基づく解法

画像から求めようとしている姿勢パラメータは、画像中の様々な特徴とその幾何学的な配置に反映されているはずである。本来は、画像生成の過程を考察して、それらの現れ方を物理モデル化し、一方で、そのような画像特徴を与えられた画像から抽出して、その物理モデルに基づいて、姿勢パラメータを導く。

しかし、そのようなことをせずに、姿勢パラメータと画像を直接に対応付け、画像生成過程の中味を考えない — すなわち、この姿勢パラメータ $\boldsymbol{\theta}$ では、こういう画像 $\boldsymbol{x}$ が見えるということのみに基づく — 手法を、一般に「見え方に基づく手法 (appearance based method)」と言う。

ここでは、最も直接的に、姿勢パラメータ $\boldsymbol{\theta}$ と、画像 $\boldsymbol{x}$ とは、線形に結び付けられるとしよう。すなわち、

$$\boldsymbol{\theta} = W^\top \boldsymbol{x} \tag{1}$$

という関係があるとする。$\boldsymbol{x}$は、M 次元ベクトルであり、$\boldsymbol{\theta}$ が P 次元ベクト（ここでは、$P = 2$）であれば、W は、$M \times P$ の行列で、$W = [\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_P]$ である。

ただし、$\boldsymbol{\theta}$ の各要素は独立に $\boldsymbol{x}$ から求まり、何も従属関係を仮定できないので、W の各列ベクトル $\boldsymbol{w}_j$ の決め方は独立である。つまり、$\boldsymbol{\theta}$ の第 j 要素 θ_j に対して、

$$\theta_j = \boldsymbol{w}_j^\top \boldsymbol{x} \tag{2}$$

となる M 次元ベクトル $\boldsymbol{w}_j$ の決め方が、ここでの姿勢決定のための問題となる。すなわち、各サンプル $\{\boldsymbol{\theta}_i, \boldsymbol{x}_i\}(i = 1, \cdots, N)$ に対して、$X = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N]^\top$ とおくと、

$$[\theta_{1j}, \theta_{2j}, \cdots, \theta_{Pj}]^\top = X\boldsymbol{w}_j \tag{3}$$

が連立して成り立つ解を求める。

ここで重要なことは、この連立方程式では、サンプル数 N （せいぜい、数百）が、決めなくてはならない未知数の個数である画像サイズ M （数万以上）よりはるかに小さいので、すべてのサンプルに対して (3) が成り立つ $\boldsymbol{w}$ は、いつも存在する。しかも、いくつでも存在する。したがって、何らかの評価基準を導入して、その意味での最適な $\boldsymbol{w}_j$ を決めることになる。

2.3 拘束付き最小化問題としての直接的対応付けの解法

一般には、(3) は不定である。そこで、解に何らかの基準を導入する。その基準を、$\boldsymbol{w}_j^\top Q \boldsymbol{w}_j \rightarrow \min.$ とする [3]。

したがって、$\boldsymbol{w}_j$ は、(3) という拘束付きの最小化問題、

$$\min_{\boldsymbol{w}_j, \boldsymbol{\lambda}} \ \{ \boldsymbol{w}_j^\top Q\, \boldsymbol{w}_j + \boldsymbol{\lambda}^\top (X\boldsymbol{w}_j - [\theta_{1j}, \theta_{2j}, \cdots, \theta_{Pj}]^\top) \}$$

の解となる。

$Q = I$ すなわち $|\ \boldsymbol{w}_j\ | \rightarrow \min.$ とする解は、

$$\hat{\boldsymbol{w}}_j = X(X^\top X)^{-1}\ [\theta_{1j}, \theta_{2j}, \cdots, \theta_{Pj}]^\top \tag{4}$$

と与えられることが知られている。

この解は次のように解釈される。

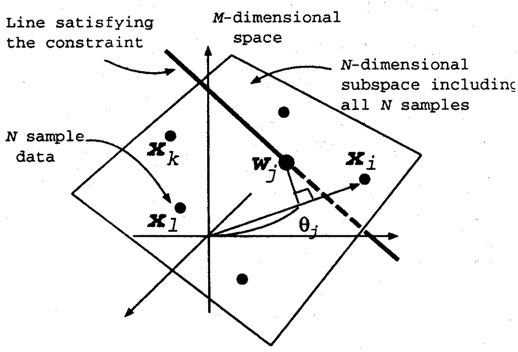

図 3: 直接的対応付けによる姿勢推定問題の線形解の幾何学的な意味。

(2) は、$\boldsymbol{w}_j$ が、$\boldsymbol{x}_i$ の分布する M 次元空間内のある (超) 平面を表す。各 i についてのこの式を一まとめにした (3) は、$\boldsymbol{w}_j$ がある直線上にあることを示す。そのような $\boldsymbol{w}_j$ の中で、$|\boldsymbol{w}_j| \to \min.$ となるものは、$\boldsymbol{x}_i$ の分布する M 次元空間内で、この直線が、すべての $\boldsymbol{x}_i$ $(i = 1, \cdots, N)$ を通る平面（$N \leq M$ なら必ず存在する）と交差する点を $\hat{\boldsymbol{w}}_j$ とすることを表す（図 3）。

3 パラメトリック固有空間法による姿勢推定

3.1 固有空間法

さて、この問題に対する固有空間法を用いた解法は、以下のように意味付けられる。

N 個のサンプル画像を K 個 $(K < N)$ の単位ベクトルの線型和で表すものとし、そのとき、れがこの N 個それぞれからの誤差の平均として最も近くなるように表すものとする。すなわち、

$$\hat{\boldsymbol{x}}_i = a_{i1}\boldsymbol{e}_1 + a_{i2}\boldsymbol{e}_2 + \cdots + a_{ik}\boldsymbol{e}_K \tag{5}$$

と置くときに、

$$J = \sum_{i=1}^{N} |\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i|^2 \longrightarrow \min. \tag{6}$$

となるように、$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_K\}$ を選ぶ。

このような単位ベクトルの組は、N 枚の画像集合 $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N\}$ の共分散行列

$$Q \equiv \frac{1}{n} X X^\top \tag{7}$$

の固有値問題

$$Q\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i \tag{8}$$

を解き，固有ベクトルを対応する固有値の大きい順に $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_K$ $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K)$ と K 個取ることで得られる。但し $|\boldsymbol{e}_i| = 1$ とする.

これらの固有ベクトルは互いに直交し，このサンプル画像集合に属する画像 $\boldsymbol{x}$ は，固有ベクトルとの内積 $a_j = \boldsymbol{e}_j^\top \boldsymbol{x}$ を用いて次のように展開できる (*Karhunen-Loève* 展開).

$$\hat{\boldsymbol{x}} = \sum_{i=j}^{K} a_j \boldsymbol{e}_j \tag{9}$$

固有空間とは $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_K\}$ の張る空間を言う．この空間は，画像データベクトルの元の M 次元空間の中で，与えられた画像集合を最小二乗誤差の意味で最もよくその内部に含む K 次元の空間となっている．もとの M 次元の画像 $\boldsymbol{x}$ は、$\hat{\boldsymbol{x}}$ の係数を用いて $K(\ll M)$ 次元のベクトル $\boldsymbol{a} = [a_1, a_2, \cdots, a_K]^\top$ で表現できる.

固有空間の必然性は、このサンプル集合の情報を保つ効率のよい空間であることにある。

3.2 パラメトリック固有空間法

さて、パラメトリック固有空間法は、サンプル画像に対するこの係数ベクトル $\boldsymbol{a} = [a_1, a_2, \cdots, a_K]^\top$ を、K 次元固有空間にプロットしていくと、パラメータの変化に応じて、その軌跡が「経験的に」滑らかになるということに基づく（図 4）。すなわち、固有空間内でのパラメータ変化に対する係数ベクトルの軌跡は、パラメータの次元（先の姿勢推定の例では、2 次元）の多様体とみなせる。

この根拠を理論的に与えることは困難であるが、*Karhunen-Loève* 展開とフーリエ展開の類似性、そして、その結果、係数 a_j の、j が小さいものは、画像の低空間周波数成分のスペクトルに対応し、さらに、画像の低空間周波数成分はパラメータの変化に対して、一般的に、あまり変化しないと、説明される [3]。

したがって、このことから、固有空間内でのパラメータ変化に対する係数ベクトルの軌跡は、「補間の能力」が期待できることになる。

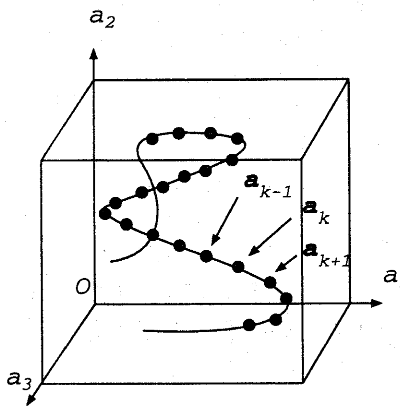

図 4: 固有空間内での係数ベクトルのプロット。パラメータの変化に応じて、その軌跡が「経験的に」滑らかになる。

すなわち、パラメータ $\boldsymbol{\theta}_k$ で得られる画像 $\boldsymbol{x}_k$ の係数ベクトルを $\boldsymbol{a}_k$ とし、$\boldsymbol{\theta}_{k+1}$ でのそれらを $\boldsymbol{x}_{k+1}$、$\boldsymbol{a}_{k+1}$ とすると、$\boldsymbol{\theta}_k$ と $\boldsymbol{\theta}_{k+1}$ の中間で得た画像の係数ベクトルも $\boldsymbol{a}_k$ と $\boldsymbol{a}_{k+1}$ の中間になる。また、その結果、ある与えられた画像 $\boldsymbol{x}$ から、その係数ベクトルを計算することで、固有空間内でその近傍にあるサンプル画像の係数ベクトルから、与えれた画像のパラメータ $\boldsymbol{\theta}$ が推定できることになる。

以上が、パラメトリック固有空間法のあらましである。

3.3 固有空間での姿勢パラメータとの直接的対応付け

姿勢のわかっているサンプル画像 $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N\}$ の、固有空間での表現

$$\boldsymbol{a}_i = [\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_K]^\top \boldsymbol{x}_i \tag{10}$$

について、前章と同様な姿勢パラメータとの線形対応付け

$$\theta_{ij} \sim \boldsymbol{v}_j^\top \boldsymbol{a}_i \tag{11}$$

$(i = 1, \cdots, N)$ を考えよう。

係数ベクトル $\boldsymbol{v}_j$ は、$\boldsymbol{a}$ と同じく K 次元であり、$K < N$ であるので、今度はすべての i に対して (11) をの両辺を等しくする $\boldsymbol{v}_j$ は存在しない。そこで、何らかの評価基準、例えば、

$$\sum_{i=1}^{N} |\theta_{ij} - \boldsymbol{v}_j^\top \boldsymbol{a}_i|^2 \rightarrow \min. \tag{12}$$

を考えて $\hat{\boldsymbol{v}}_j$ が選ばれる。

さて、以上のプロセスを元の画像データの空間で考えてみよう。(11) の右辺は、

$$\begin{aligned} \theta_{ij} \sim \boldsymbol{v}_j^\top \boldsymbol{a}_i &= \boldsymbol{v}_j [\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_K]^\top \boldsymbol{x}_i \\ &= \bar{\boldsymbol{w}}_j^\top \boldsymbol{x}_i \end{aligned} \tag{13}$$

であるので、(3) と同じ形になる。ただし、上式から、$\bar{\boldsymbol{w}}_j$ は、$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_K\}$ で張られた K 次元部分空間（固有空間）に閉じ込められている。

この部分空間は、$\boldsymbol{x}_i$ からこの部分空間内の点 $\hat{\boldsymbol{x}}_i$ におろした距離の 2 乗和に対して

$$\sum_{i=1}^{N} |\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i|^2 \longrightarrow \min. \tag{14}$$

となる空間であり、さらに、この $\hat{\boldsymbol{x}}_i$ が固有空間内の点としての $\boldsymbol{a}_i$ に他ならないので、(12) より、結局、$\bar{\boldsymbol{w}}_j$ は、

$$\sum_{i=1}^{N} |\theta_{ij} - \bar{\boldsymbol{w}}_j{}^\top \boldsymbol{x}_i|^2 \rightarrow \min. \tag{15}$$

と選ばれる（図 5）。

以上をまとめると、

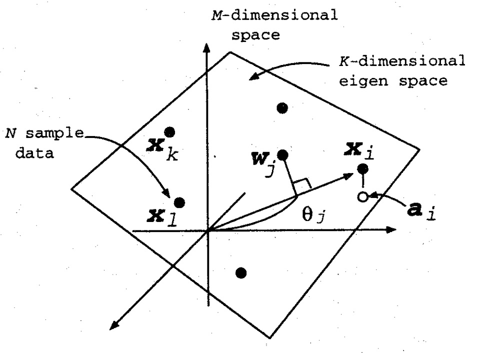

図 5: 固有空間での画像とパラメータの直接的対応付けのあらまし

- 前章で示した、サンプル画像データと姿勢パラメータとの直接的対応付けでは、M 次元空間で、N 個のサンプル点を含む N 次元部分空間内に、最もよく両者を対応付ける $\hat{\boldsymbol{w}}$ をとる。（対応付けの残差は 0 となる。）
- 本章で示した、固有空間でのサンプル画像と姿勢パラメータの直接的対応付けでは、M 次元空間内の、K 次元固有空間内に、最もよく N 組の対応を付ける $\bar{\boldsymbol{w}}$ をとる。（対応付けの残差は 0 とならない。）

最も重要な点は、$K < N$ であり、このために後者の残差は 0 にはならないが、その代りに、将来的に得られるかもしれないサンプルも許容するような空間内で解を探す。ここに、固有空間法での汎化しかけがある。

4 学習として見たときのパラメトリック固有空間法

4.1 パラメータ推定の学習

前章で、モデルあてはめと汎化能力の間の一つの関係を見た。

この関係を、例えとして先に挙げた画像からの姿勢推定問題、それを一般化した画像からのパラメータ推定問題を、さらに一般化して、学習の問題として解釈してみる。

学習データとして、データの組、$D = \{(\boldsymbol{x}_1, \boldsymbol{\theta}_1), (\boldsymbol{x}_2, \boldsymbol{\theta}_2), \cdots, (\boldsymbol{x}_N, \boldsymbol{\theta}_N)\}$ が与えられている。このとき、$\boldsymbol{\theta}$ と $\boldsymbol{x}$ との間には、

$$\boldsymbol{\theta} = f(\boldsymbol{x}; \boldsymbol{w}) \tag{16}$$

という関係がある（とする）。ここで、$f(\cdot; \boldsymbol{w})$ はある関数形で、これは固定されている。$\boldsymbol{w}$ は、その関数の具体的な形を決めるパラメータであり、このパラメータがデータの組に応じて変るとする。（ここで、$\boldsymbol{\theta}$ を指す「パラメータ」という語とは使い分けることに注意。）関数形を決めるパラメータ $\boldsymbol{w}$ は、L 次元を持つとする。前記の2.2節での例では、$f(\boldsymbol{x}; \boldsymbol{w})$ は、線形式（内積）$\boldsymbol{w}^\top \boldsymbol{x}$ であった。さらに、このときは、$L = M$ であった。

さて、サンプルを与えられた上での、パラメータ推定の学習とは、次々に D のサンプルデータ $(\boldsymbol{x}_i, \boldsymbol{\theta}_i)$ を参照しながら、$\boldsymbol{w}$ をチューンナップして、正しい $f(\boldsymbol{x}; \boldsymbol{w})$ を得、$\boldsymbol{x}$ から $\boldsymbol{\theta}$ を求める関数を得ようというものである。

さらに、データの組 $(\boldsymbol{x}_i, \boldsymbol{\theta}_i) \in D$ については、

$$\boldsymbol{\theta}_i = f(\boldsymbol{x}_i; \boldsymbol{w}) + \boldsymbol{n}_i \tag{17}$$

（ただし、$\boldsymbol{n}_i$ は観測ノイズ）であるとする。

ノイズ $\boldsymbol{n}_i$ は、重要である。結果的に、モデルの不完全さを補う役割も持つからである。

一般には、D 内のデータを用いて、

$$\sum_i \| \boldsymbol{n}_i \|^2 \longrightarrow \min. \tag{18}$$

となるような $\boldsymbol{w}$ を求め、$f(\boldsymbol{x}; \hat{\boldsymbol{w}})$ を得る。

続いての問題は、このようにして得た $f(\boldsymbol{x}; \hat{\boldsymbol{w}})$ が、$\boldsymbol{\theta}$ の推定において、どれくらい汎化能力を持つかである。D 内の $\boldsymbol{x}$ に対して、どれくらいの精度で $\boldsymbol{\theta}$ を言い当てることができるかは、(18) の評価で与えられる。D にはなかった新たな $\boldsymbol{x}$ に対して、十分よい $\boldsymbol{\theta}$ をこの $f(\boldsymbol{x}; \hat{\boldsymbol{w}})$ で導けるかが汎化能力である。

これは、前節でのパラメトリック固有空間法の場合にあてはめれば、「補間能力」と称したものにあたる。すなわち、固有空間内での係数ベクトルの軌跡が、パラメータ $\boldsymbol{\theta}$ の変化に対して十分に滑らかどうかが、ここでの汎化能力と同等の評価である。

4.2 学習の汎化能力

一般に、学習の汎化能力については、「過学習の理論」と呼ばれるものがある。すなわち、「学習し過ぎは良くない」ことを理論化されている。そこでは、学習誤差と汎化誤差は相容れないことが示されている。ここで、学習誤差とは、学習サンプルに対する誤差であり、汎化誤差とは、あらたに遭遇するであろう相手に対する誤差の期待値である。学習誤差を最小化すると、汎化誤差が一般には大きくなる。これを過学習という。この過学習の度合いは、チューンナップするパラメータ1個あたりのサンプル数で決まる。

おおよその概要を、図6に示す [5]。

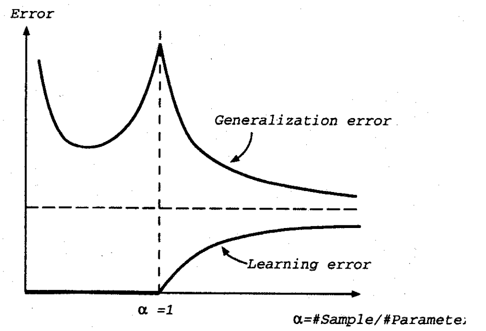

図 6: 学習誤差と汎化誤差。チューンナップするパラメータ1個あたりのサンプル数で決まる。

4.3 学習の汎化能力から見た姿勢推定問題の解法

前節での、汎化能力の一般的な傾向を、第2章での我々の姿勢推定問題へと類推してみる。

第2章でのサンプル画像そのものとの直接的な対応付けでは、チューンナップするパラメータの数は M、サンプル数は N で、チューンナップするパラメータ1個あたりのサンプル数 $\alpha = N/M$ は非常に小さい値となり、図6においては、左はしのほうにある場合となる。

すなわち、2.2節での解法では、汎化能力は一般には期待できない。

一方、固有空間法においては、その次元 K は、サンプル数 N に対して必然的に、$K \leq N$ であるので、一旦、固有空間に落したあとであれば、線形モデルをあてはめても、図6においての臨界点 α_0 を越えた右側になる。すなわち、汎化能力が期待できる領域に入る。

このように解釈すると、姿勢パラメータのあてはめに対する、2つのアプローチの決定的な差が分かるとともに、汎化能力を得るために一つの手法としての固有空間法の性質が浮かぶ。

5 実験結果

画像と姿勢パラメータの直接的対応付けとして2、3章で与えた手法の、適用結果を示す。

図2に示したように、回転テーブルの上に物体をのせ、カメラを固定して画像を撮影した。画像から推定する対象は、カメラに対する姿勢（回転テーブルの角度）とする。ただし、角度をそのまま対象とすると、物体を1周させたとき画像は元に戻って同一となるが角度は ±360 度となり、明らかに不都合である。これを回避するため、回転角を θ_i として、冗長な表現 $(\cos\theta_i, \sin\theta_i)$ を学習・推定した。

回転テーブルの回転角を1度ずつ変化させながら物体をほぼ1周させ、340枚の画像を取得した。このうち18枚のサンプルを回転角0度から等間隔に取りだし、学習サンプルとした。得られた学習結果を残りの $340 - 18 = 322$ 枚の画像に適用して姿勢角（の余弦と正弦）を計算したのが図7である。図のそれぞれのプロットについて、横軸は $\cos\theta$、縦軸は $\sin\theta$ を表す。各プロットは、固有空間の次元を18(次元を落とさない), 9, 5と順に落としたものである。

18次元(次元を落とさない)での結果から、18個のサンプルによる対応付けでそのものが、この場合は十分な汎化の能力を持っていることが示されている。さらに、次元を9、さらに5と落した、固有空間でも、かなりよい推定ができており、固有空間法のこの面での有効性を示している。

6 おわりに

画像からの対象の姿勢パラメータを推定する問題への、見掛けに基づくアプローチにおいて、パラメトリック固有空間法による解法が、どのような幾何学的な意味を持つのかを論じた。ここでは、元のサンプル画像により直接的に姿勢パラメータを対応づける手法を対比して、固有空間法がサンプルデータ以外の新たなデータにも対応しうる汎化能力を持つ仕掛けを探った。

また、固有空間法を適用することで、副次的な効果としての次元数を下げる、ということが画像解析におけるモデルあてはめに本質的な役割をしているのではないかと論じた。

これらの議論は、対象とする画像の性質そのものにも大きく依存する微妙なものである。ここでの、議論が、どこまでの画像の性質をカバーするものであるか、また、どのような五臓の性質に依存する議論であるかは、今後の課題である。

参考文献

[1] 村瀬、固有空間法による画像認識、松山ほか編「コンピュータビジョン－技術評論と将来展望」(新技術コミュニケーションズ)、pp.206-218 (1998)

[2] Murase, Nayar, Visual learning and recognition of 3-D objects from appearance, IJCV, vol.14, pp.5-25 (1995)

[3] Okatani, Deguchi, Yet another Appearance-Based Method for Pose Estimation Based on a Linear Model, MVA2000, pp.258-261 (2000)

[4] Deguchi, A Direct Interpretation of Dynamic Images with Camera and Object Motions for Vision Guided Robot Control, IJCV, vol.37, pp.7-20 (2000)

[5] 甘利、村田、Muller、学習の数理モデル－汎化能力と過学習－、外山ほか編「脳と計算論」(朝倉書店)、pp.37-53 (1997)

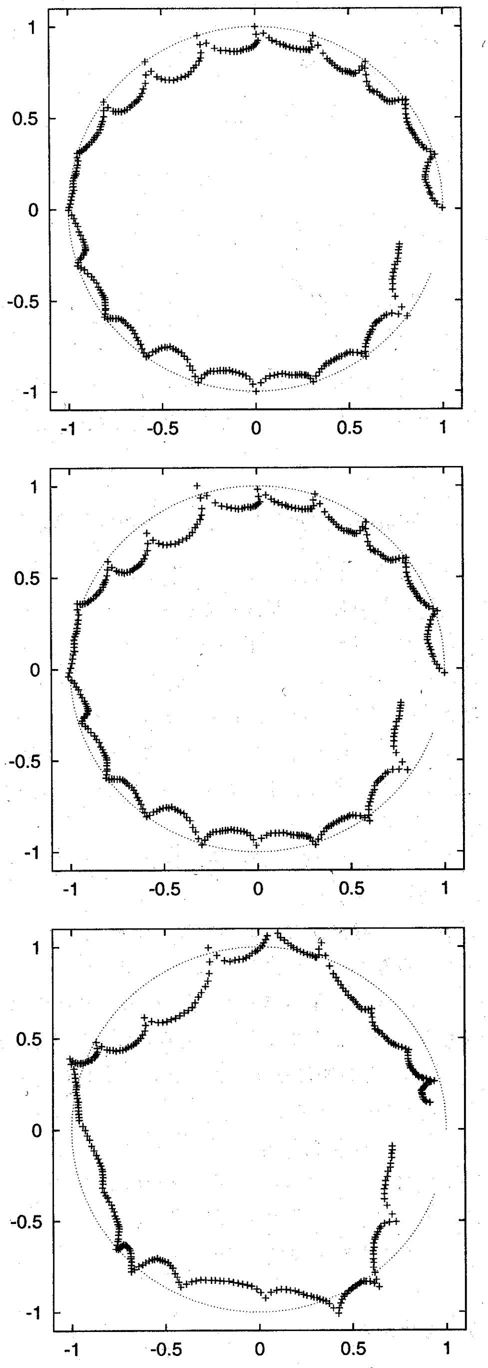

図 7: 固有空間の次元を 18, 9, 5 として推定された姿勢。横軸、縦軸はそれぞれ姿勢角の余弦と正弦。円周状に広がるドットは真の値を、菱形は推定された値を表す。