

## 解説

## 動的計画法によるパターンマッチング†



大田 友一† 山田 博三††

## 1. ま え が き

二つのパターンを照合し、一致の度合を評価したり、一致させるために一方のパターンに加えるべき変形を求めたり、一方のパターンを他方の部分パターンとして採したりする操作は、コンピュータビジョンにおいてももっとも基本的な処理の一つであり、パターンマッチングと呼ぶ。

通常、一方のパターンはデータとして与えられる画像であり、もう一方は、モデルとして与えられる標準パターンである。標準パターンは、求めたい形状をテンプレートとして陽に表現する場合が多いが、望ましい形状を評価関数に埋め込んだかたちで陰に表現する場合もある。

求めたい対象の特徴を標準パターンに十分に反映させるためには、標準パターンを大きくとる必要があるが、逆に、対象の変形への対応が困難となる。したがって、標準パターンを複数の部分テンプレートに分割し、各部分要素のもつべき特徴と、部分要素間の関係として対象を表現する。これにより、局所的な変形は、各部分要素で吸収し、大局的な変形は、部分要素間の関係で吸収することが可能になる<sup>1)</sup>。

しかし、このような標準パターンにもっともよく適合するパターンを画像中で採し出す操作は、結局、部分要素間の関係により変形可能なすべてのテンプレートを生成し、おのおのについて照合処理を行うのと等価である。コンピュータビジョンにおける問題では、部分要素数が数十になる場合も珍しくなく、現実にはパターンマッチングを実行するためには、計算量を削減する手法が不可欠な技術となっている。

動的計画法 (dynamic programming) は、線形計画法、非線形計画法とならぶ数値計画法の一手法であ

り、 $n$  段階の決定を必要とする問題を、 $n$  個の 1 段階決定問題に分解して解くことにより、計算量を対数のオーダーに低減させることができる。

本稿では、コンピュータビジョンにおけるいくつかの事例を、パターンマッチングの観点から眺め、信頼性の向上には大局的評価が必要であること、その結果として解の探索空間が膨大になること、その解決法として動的計画法が有効であることを、事例に即して解説する。また、動的計画法を適用するために、本来の問題に加えられている条件や制限についても言及する。

## 2. 動的計画法とは

動的計画法 (以下、DP と呼ぶ) は、多段階最適決定過程の問題を電子計算機で高速に解くための数学的手法であり、1950 年代に Bellman が発展させた<sup>2), 3)</sup>。

いま、 $x_i$  ( $i=1, \dots, 3$ ) を、離散値をとる状態変数とし、 $x_i$  上で定義される目標関数  $h(x_1, x_2, x_3)$  を最小にする  $x_i$  の組合せを探す問題を考える<sup>4)</sup>。もし、関数  $h(\cdot)$  についてまったく予備知識がない場合には、 $(x_1, x_2, x_3)$  のすべての可能な組合せについて  $h(\cdot)$  を評価し、最小値をもつものを探す必要がある。これは、図-1 (a)において、 $x_{start}$  から  $x_{end}$  に至るすべての経路を調べることに相当する。

しかし、 $h(\cdot)$  が、

$$h(x_1, x_2, x_3) = f_1(x_1) + d_1(x_1, x_2) + d_2(x_2, x_3) \quad (1)$$

のように分解できる場合には、 $x_1$  に依存するのは  $f_1(\cdot) + d_1(\cdot)$  の部分のみであり、 $x_2$  のおのおのの値について  $x_1$  で最小化することにより、

$$f_2(x_2) = \min_{[x_1]} [f_1(x_1) + d_1(x_1, x_2)] \quad (2)$$

を得て、

$$\min_{[x_1]} h(\cdot) = \min_{[x_1]} [f_2(x_2) + d_2(x_2, x_3)] \quad (3)$$

のように  $x_1$  を消去できる。図-1 で、 $f_1(x_1)$  は、 $x_{start}$  から  $x_1$  の各状態へ移動する場合のコスト、すなわち、 $x_1$  の各状態に至るのに必要な最小コスト、

† Pattern Matching Using Dynamic Programming by Yuichi OHTA (Inst. of Information Sciences and Electronics, University of Tsukuba) and Hiromitsu YAMADA (Machine Understanding Division, Electro Technical Laboratory).

†† 筑波大学電子情報工学系

††† 電子技術総合研究所情報情報部

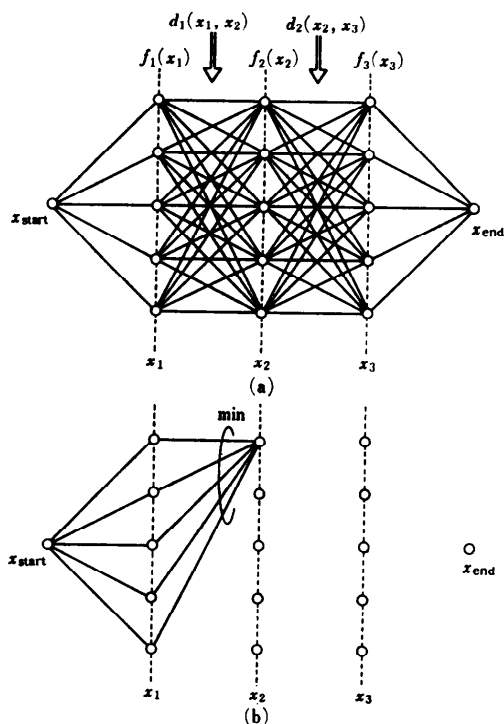


図-1 動的計画法

$d_1(x_1, x_2)$  は、 $x_1$  の各状態から  $x_2$  の各状態へ移動するコストと考えれば、(2)式の操作は、同図(b)のように、 $x_2$  の各状態において、 $x_1$  からそこに至るすべての経路を調べ、 $x_2$  の各状態に至る最小コストを求めることに相当する。

同様の操作を、 $f_2(\cdot) + d_2(\cdot)$  について行い  $f_3(x_3)$  を得て、

$$\min_{\{x_i\}} h(\cdot) = \min_{\{x_i\}} [f_3(x_3)] \quad (4)$$

とすることができ、最後に、 $x_3$  について最小化することにより、 $h(\cdot)$  の最小値を求められる。

(2)式は、一般には、

$$f_n(x_n) = \min_{\{x_{n-1}\}} [f_{n-1}(x_{n-1}) + d_{n-1}(x_{n-1}, x_n)] \quad (5)$$

の漸化式で書け、1段階決定過程を表す。これを  $n$  回、再帰的に順次評価することにより、 $n$  段階決定過程  $h(\cdot)$  の最小値を求めることになる。

状態変数の数を  $n$ 、各状態変数がとり得る離散値の数を  $k$  とすれば、すべての状態の組合せは  $k^n$  とおりあるのに対して、上記の DP により調べるべき部分状態の数は、 $(n-1) \times k^2 + k$  個にすぎず、ほぼ対数のオー

ダに計算量を低減できる。

数理計画法の用語に従って、多段階決定過程の各段階における決定の選択の全系列、すなわち、 $x_i (i=1 \cdots n)$  の値の組合せ、を政策 (policy) と呼び、 $h(\cdot)$  を最小にする政策を最適政策 (optimal policy) と呼ぶことにすると、DP は、次に示す最適性原理 (principle of optimality) を利用して、多段階決定過程を1段階決定過程の系列に分解する。

**最適性原理** 最適政策とは、最初の状態や決定がどうであっても、この最初の決定の結果として起こった状態について、以後の決定は最適政策となるように構成しなければならない、という性質をもつ。

いかなる多段階決定過程も、形式的に1段階決定過程に分解し、解を計算することは可能である。しかし、求められた解が、最適政策となっていることが保証されるのは、その決定過程が、次に示す Markov 性を満たす場合に限られる。

**Markov 性** 各段階での決定が、その時点での系の状態のみに依存し、過去の決定の履歴に無関係に行えなければならないという性質。

この意味するところは、上に示した例題から明らかであろう。

複数の部分要素で構成された標準パターンと画像とのパターンマッチング問題を、DP を利用して解く場合には、個々の部分要素の照合操作が、各決定段階に相当し、図-1 の  $x_{start}$  から  $x_{end}$  に至る経路が標準パターンの変形を表す。したがって、多段階逐次決定問題として取り扱うためには、各部分要素を順序付けする必要がある、しかも、Markov 性を満たす問題として定式化しなければならない。

### 3. 画像からのパターン抽出

画像中から、物体の輪郭などの連結したエッジを抽出する処理は、画像解析における重要な前処理技術として、古くから研究されてきた。画像には、抽出したい輪郭の信号以外にも、多くの種類の信号や雑音が重畳して記録されており、局所的なエッジ抽出演算子を施すだけでは、満足できる輪郭の抽出が困難なことはよく知られている。

抽出したい輪郭の形状を先験知識として設定し、パターンマッチングにより、トップダウン的に抽出を行う手法が有効であり、形状が、直線や円弧など、少数のパラメータで記述できる場合には、Hough 変換を用いることができる<sup>6)</sup>。

求めたい輪郭の形状自身は明確でなく、なめらかな曲線であること、微分値が大きな部分を通ること、のような性質が規定されている場合には、これらの性質を評価する関数を設定し、輪郭上での評価値の積分を最大（最小）にするような形状を求める。これは、一般には、正則化 (regularization) の枠組みにおいて、最適化問題として変分法により扱うことができ、山登り的に、目標とする形状を抽出できる<sup>9)</sup>。

一方、輪郭に沿った道のりをパラメータとする座標系をとれば、輪郭は順序付けされた1次元の系列と見なすことができ、上記の最適化問題を DP の逐次決定過程にあてはめて解くことも可能である。

輪郭抽出への DP の利用は、Montanari<sup>7)</sup> によって最初に試みられ、Ballard<sup>9)</sup> は、コントラストの悪い X 線写真における腫瘍の輪郭抽出に適用した。

最近、正則化の枠組みが注目されるのにもない、変分法の代わりに DP を利用する試みも報告されている。Amini<sup>9)</sup> は、Kass<sup>10)</sup> によって提案された、エネルギー最小化による輪郭 (active contour) を、DP を利用して求めた。いま、輪郭のパラメータ表現を  $v(s) = (x(s), y(s))$  としたとき、輪郭のもつエネルギーを、

$$E_{\text{snake}} = \int [E_{\text{int}}(v(s)) + E_{\text{image}}(v(s)) + E_{\text{con}}(v(s))] ds \quad (6)$$

とする。  $E_{\text{int}}$  は、輪郭の内部エネルギー、  $E_{\text{image}}$  は、画像データの性質に応じて輪郭に加わる力、  $E_{\text{con}}$  は、輪郭の形状について対話的指示などで特別に加えた拘束力を表す。  $E_{\text{image}}$  は、輪郭の通過する部分の画像の明度値とその微分値により定義され、  $E_{\text{con}}$  とともに、輪郭上の各画素ごとに評価できる。一方、  $E_{\text{int}}$  は、

$$E_{\text{int}}(v(s)) = \alpha(s) |v_x(s)|^2 + \beta(s) |v_y(s)|^2 \quad (7)$$

として、  $v(s)$  の  $s$  に関する一次、二次微分値を用いて定義され、なめらかさの拘束を与える。  $v_i$  を、輪郭上の各画素とし、上式を差分式として表すと、

$$E_{\text{int}}(i) = \alpha_i |v_i - v_{i-1}|^2 + \beta_i |v_{i+1} - 2v_i + v_{i-1}|^2 \quad (8)$$

となり、輪郭上のエネルギーの総和は、  $\sum d_i(v_{i+1}, v_i, v_{i-1})$  の形をもつ。したがって、2. の (5) 式を単純に拡張し、  $v_{i+1}$  画素において、  $v_{i-1}$  と  $v_i$  に関して最小化を図ることにより、DP の漸化式を構成できる。各  $v_{i+1}$  画素において選択すべき決定は、輪郭をその画素から隣接画素のいずれかへ移動させることである。

上記の DP の過程を輪郭に沿って一順行うことにより、輪郭は以前の輪郭の隣接部分に移動する。この過程を、エネルギーの減少がなくなるまで繰り返す。図-2 (a) に、初期値として与えた輪郭を、(b) に、最終的に求めた輪郭を示す。輪郭は、30 節点をもつ折れ線を与えている。

この例では、一順の過程内では最適解が求められることが保証されるが、微小移動の繰り返しを行うため、大局的に最適解に到達する保証はない。しかし、変分法により微小移動させる場合に比べて、オイラーの方程式の表現に必要な高次の導関数を用いなくてよく、離散的な画像上で評価するのに適しており、収束も保証されて安定に動作するなどの利点がある<sup>11)</sup>。

#### 4. 2次元形状の照合

陽に形状を表現したモデルとデータの照合に DP が最初に使われたのは、音声認識においてである。ここでは、標本点（時間軸）を非線形に伸縮させて、モデル要素と入力要素の最適な対応付けを行うために DP

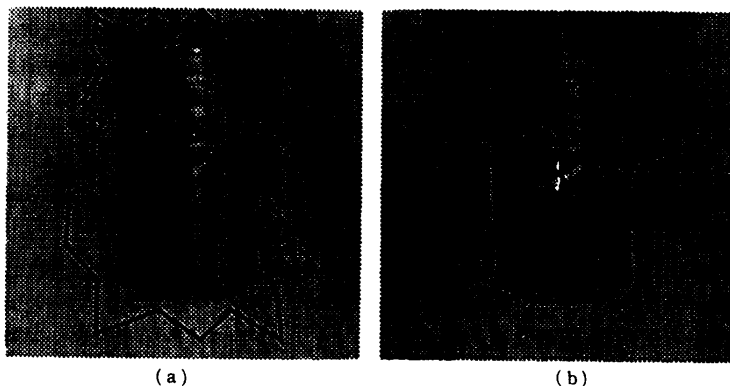


図-2 輪郭の抽出例<sup>9)</sup>

が用いられている。この手法は非常に大きなインパクトを与え、音声認識に不可欠な技術となっている<sup>(12),(13)</sup>。

一方、2次元の形状を定義した標準パターンと画像データの照合にDPを用いる場合、問題を逐次決定過程として構成する方法は、必ずしも明らかではない場合が多い。

#### 4.1 輪郭形状の照合

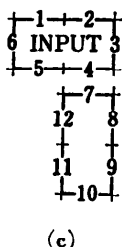
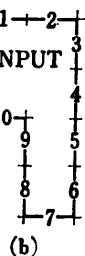
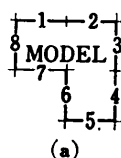
2次元の形状を、その輪郭で表現することにより、1次元系列として扱うことができる。したがって、輪郭の巡回性を考慮する必要を除けば、音声認識の場合とはほぼ同様の取り扱いが可能である。

音声の場合、入力要素、モデル要素とも、時間軸に沿って並んでおり、対応関係が時間軸方向に反転することはない。また、単語や音素の抜けを考えないかぎり対応における時間軸方向の大きな飛び越しを考えなくてもよい、という単調性の条件を用いて、解探索の大幅な削減を行っている。

輪郭照合の具体例として、図-3 (a)と(b)の対応問題を考える。ここでの対応要素は、右廻りに順序づけされた輪郭の多角形近似線分である。輪郭線分番号に関する演算は、輪郭の巡回性を考慮して (たとえば  $m=1$  のとき  $m-1$  は  $M$ ) 行うものとする。

この問題を解くためのDPの漸化式は、

$$f(m, i) = \min \begin{cases} f(m-1, i-1) + d_1(m, i) \\ f(m-1, i-2) + d_2(m, i-1, i) \\ f(m-2, i-1) + d_3(m-1, m, i) \end{cases} \quad (9)$$



と表される。この式の意味は、モデル線分  $m$  を入力線分  $i$  に固定したときの、モデル線分1から  $m$  までの対応の最小コスト  $f(m, i)$  は、

(1) モデル  $m$  と入力  $i$  の1対1対応のコスト (たとえば線の方角や長さの違いから計算)  $d_1$  と、 $m-1$  までの最小コスト  $f(m-1, i-1)$  との和、

(2) モデル  $m$  と入力  $i-1$  および  $i$  の1対2対応時のコスト  $d_2$  と、 $f(m-1, i-2)$  との和、

(3) モデル  $m-1$  および  $m$  と入力  $i$  の2対1対応時のコスト  $d_3$  と、 $f(m-2, i-1)$  との和、の3項の最小値として求められる、である (図-3 (d) 参照)。

モデル側の順序と入力側の順序に乱れがなく、1対1、1対2、2対1の対応しか許されないことを仮定し、各決定段階で最小化を行う範囲を削減している。

輪郭の巡回性にもなう始点の移動を許すため、 $m=1$  における初期値は、入力の各要素  $i$  に対して

$$f(1, i) = \min \{d_1(1, i), d_2(1, i-1, i)\}$$

で設定する。

最適対応関係は、 $f(M, i)$  の最小値をとる  $i$  をもとに、漸化式評価時に選んだ項を記憶している  $M \times I$  行列でポインタをさかのぼることにより決定される (図-3 (d) の太線)。

現実の画像では、輪郭線に“切れ”が生じることを前提とする必要がある。この場合、図-3 (c) のように入力要素の順序に乱れが生じ、上記の単調性を仮定したアルゴリズムはそのままでは適用できない。

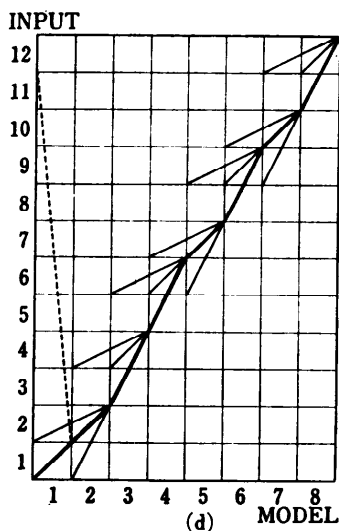


図-3 輪郭形状の照合

順序性が仮定できない場合、モデル要素  $m$  と入力要素  $i$  が対応するとき、モデル  $m-1$  に対応可能な要素 ( $i$  に先行可能な要素) としてすべての入力要素を考慮する必要がある。しかし、このように順序性の拘束を完全になくすと、DP の計算量が増大するだけでなく、不必要に広い範囲の解を探索することになり、期待しない結果を出力する危険性も増加する。

山田<sup>14)</sup>は、要素間の距離を手がかりとして、 $i$  に先行可能な要素を  $i$  の関数として制限する工夫を行っている。この先行可能要素の導入により、単調性が仮定できる場合と比べて計算量をあまり増大させずに、非単調性をもつ対象の照合ができる。この手法は手書き漢字に対して提案され、その変形は距離画像における多面体の認識にも用いられている<sup>15)</sup>。

また、モデル要素をグループ化し、グループ間でのみ線の切れを許すという方法もある<sup>16)</sup>。

#### 4.2 可変テンプレートによるパターン抽出

比較的少数のパラメータの調節により 2 次元テンプレートの形状を可変にし、パラメータの最適な設定問題を逐次決定過程として構成することにより、2 次元形状の照合を DP で行える。

輪郭形状を多角形近似線分で与え、テンプレートの各辺の方向は固定にし、長さのみを独立にある範囲で可変にすることにより、多角形の形を変えるようにする。このようなテンプレートを入力画像の上に置いたとき、辺上を入力画像の特徴量とモデルの特徴量の差の平均値が最小になるような長さの系列を決定する。

輪郭の始点が未知の場合には、すべての画素を始点としてテンプレートのあてはめを行い、良い評価を与えたものを抽出する。パラメータにより拘束された空間のみを探索するので、総当たりの 2 次元照合に比べて比較的少ない計算量でよい。

この方法は、1970 年前後に比較的粗い分解能の文字の認識手法として提案されたが<sup>17-19)</sup>、辺の方向の指定が制限され、計算量も多いことから、あまり用いられることはなかった。

最近、モデル線分方向の任意化、線分方向性評価、モデル輪郭の始点と終点の巡回性評価、粗密法 (coarse-to-fine) による計算量の低減などの改良により 512×512 程度の濃淡画像に対しても適用できるものとなり、医用画像<sup>20-22)</sup>、産業応用<sup>23)</sup>、文字認識<sup>24)</sup>など広い範囲で用いられ、汎用的かつ実用性の高い手法として認められつつある。

#### 4.3 ネットワーク状テンプレートの利用

図-4 に示すような、2 次元的に配置された部分要素が、相互に密接な関係をもつテンプレートの場合には、目標関数が、2. の (1) 式のように分解できない。このような目標関数をそのまま DP 漸化式に展開すると、部分要素相互の関係の深さに従って記憶量と計算量が指数関数的に増大し、DP による計算量削減の効果は期待できない。

Fischler<sup>25)</sup> は、 $M$  個の部分要素と、相互の位置関係がバネで拘束されたテンプレートと、入力画像を照合し、最適な配置を求める問題を、

$$h() = \sum_{m=1}^M \left\{ d(m, i_m, j_m) + \sum_{m'=1}^{m-1} e(m, m', i_m - i_{m'}, j_m - j_{m'}) \right\} \quad (11)$$

なる目標関数の最小化問題として設定した。ただし、 $d()$  は、 $m$  が  $(i_m, j_m)$  に置かれたときの部分テンプレートマッチングによるコスト、 $e()$  は、 $m$  と  $m'$  がそれぞれ  $(i_m, j_m)$  と  $(i_{m'}, j_{m'})$  にあるときのコストである。すなわち、 $m$  の配置のコストは、単に  $m$  の位置だけから決まるのではなく、それ以前に配置されたすべての要素に影響される。

これを、

$$f(m, i_m, j_m) = \min_{(i_{m-1}, j_{m-1})} \left\{ \{ f(m-1, i_{m-1}, j_{m-1}) + d(m, i_m, j_m) + \sum_{m'=1}^{m-1} e(m, m', i_m - i_{m'}, j_m - j_{m'}) \} \right\} \quad (12)$$

なる漸化式に展開し、DP を適用する。ここでは、要素  $m-1$  の配置に関して最小化するに当たって、要素 1 から  $m-2$  の配置は、該当する  $f(m-1, i_{m-1}, j_{m-1})$  を与えたものをそのまま用いて  $e()$  を評価する。

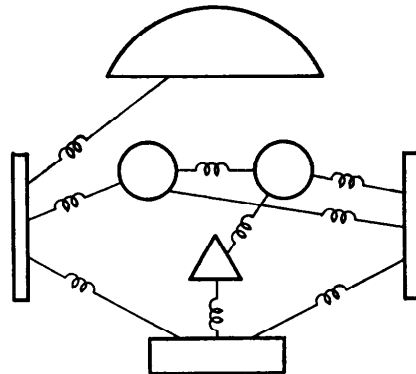


図-4 顔のネットワーク状テンプレート

上記の決定過程は、各決定段階において、過去の決定の履歴を参照することになり、Markov 性を満たさず、得られる結果が最適解である保証はないが、テンプレートの2次元性を考慮しながら計算量を抑える一つの現実的な方法である。

上に述べたような、目標関数としては部分要素相互の関係を正確に表現し、DP 適用時に、準最適解で妥協する方法に対して、あらかじめテンプレートを加工し、相互関係の枝を刈り取って最小木を構成し、根から遠い要素から順に決定することにより、最適解を保証する方法もある<sup>14)</sup>。

#### 4.4 2次元画像の照合

Moore<sup>26)</sup> は、2枚の画像  $r_{mn}$  と  $u_{ij}$  の画素単位の照合を

$$f(i, j, m, n) = \min_{\{(i', j', m', n')\}} \{f(i', j', m', n') + d(i', j', m', n', i, j, m, n)\} \quad (13)$$

なる漸化式評価により行う方法を提案している。ここで、 $(i', j', m', n')$  は、 $(i, j, m, n)$  で構成される4次元空間において、 $\{i, j, m, n\}$  と  $\{i-1, j-1, m-1, n-1\}$  で構成される4次元単位立方体の  $(i, j, m, n)$  を除く15頂点であり、漸化式は  $(i, j, m, n)$  の小さい点から順に、 $1 \times 1 \times M \times M$  回計算される。これは、DP の1次元から2次元への素直な拡張であるが、計算量と記憶量が問題であろう。対応の範囲を、 $1:1, 1:0$  に限定している点や、画像の位置ずれによる周辺部での不一致を、 $1:0$  対応の評価で吸収しようとしている点も問題であり、実際の対象に対する効果も不明である。

磯道<sup>27)</sup> は、2次元画像の照合において、 $x$  方向の1次元 DP を  $y$  方向に、次に  $y$  方向の1次元 DP を  $x$  方向に行う方法を提案している。

### 5. 両眼立体視における弾性マッチング

#### 5.1 立体視における対応問題

両眼立体視<sup>28)</sup>では、左右の画像上で対応を求める操作の自動化(対応問題)が非常に困難な課題である<sup>29)</sup>。

対応の決定を困難にする最大の要因は、左右画像中に類似した特徴が多数存在することにより生じるあいまいさであり、これに対処する方策の一つは、近接して存在する複数個の対応候補の照合を総合的に処理することにより個々の対応候補がもつあいまいさを解消することである<sup>30)</sup>。

左右画像中の複数個の候補同士の照合を総合的に評

価して、互いに整合性を保った最適な対応の集合を得る問題は、結局のところ、多対多の組合せにより生じる膨大な空間中で最適解を探索する問題である。したがって、それに必要な計算量を削減する手法の導入が、立体視における重要な研究課題となり、粗密法<sup>31)</sup>、弛緩法<sup>32)</sup>などとともに、以下に述べるように DP を用いる手法が研究されている。

立体視では、図-5 に示すように、左右のカメラのレンズ中心を結ぶ直線とシーン中の任意の一点で決定される平面(エピポーラ面)が、左右の画像面と作る交線はエピポーラ線(epipolar line)と呼ばれ、エピポーラ面上にのっているシーン点の左右画像上の像は、必ず、左右一対のエピポーラ線上に存在する性質がある。したがって、立体視における左右画像間の対応問題は、左右エピポーラ線間の対応関係を求める問題に分解できる。これは、音声認識の場合と同様、2本の1次元信号波形間の弾性マッチングの問題として扱え、DP の枠組みにあてはめることができる。

一般に、立体視における対応点が見つかるべきヒューリスティックな条件として、次の3個の制約があげられる<sup>33)</sup>。

○適合性(compatibility): 対応する点は類似の性質をもつ。

○一意性(uniqueness): 一方の画像上の点は、他方の画像上の2点以上と対応しない。

○連続性(continuity): 視差はなめらかに変化する。

DP を対応問題に適用する場合には、前述の Markov 性の条件を満たしつつ、上記3個の制約条件を多段階逐次決定過程に組み込む必要がある。

#### 5.2 DP による対応探索

図-6 に示すように、左右一対のエピポーラ線を2本の座標軸とする平面をとると、平面上の任意の点は、左右エピポーラ線上の2点間の対応を表す。したがって、左右のエピポーラ線の対応関係は、この平面

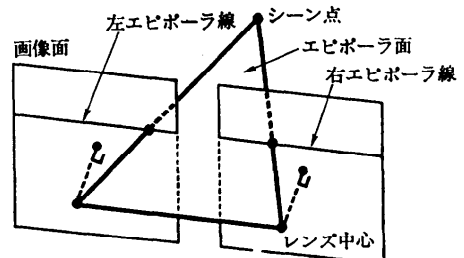


図-5 立体視

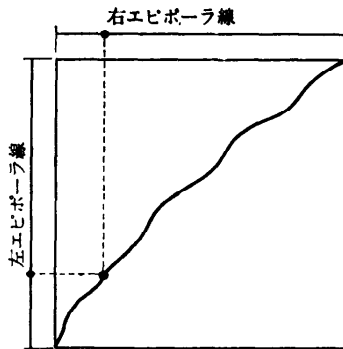


図-6 探索平面

上の点の集合により表現できる。この平面を探索平面と呼ぶ。

パターンマッチングの観点からは、モデルとデータに相当する左右のエピソード線をそれぞれ複数の部分に分割し、部分要素ごとの照合と、部分要素間の関係の評価する関数の値を最大(小)にする対応関係を求める問題である。

逐次決定過程を構成するための部分要素への順序付けは、モデルと見なすエピソード線上での要素の並びに従えばよいが、前述の一意性の制約を確保するためには、若干の注意が必要である。すなわち、一意性を保証するためには、各決定段階において、モデル要素と対応させようとするデータ要素が、それ以前に他のモデル要素と対応させられていないかどうかを調べる必要がある。しかし、この操作は、決定の過去の履歴を調べることであり、このような操作を組み込んだ決定過程は、DP で最適解が見つかることを保証する条件である Markov 性を満たさない。

これを回避するため、左右のエピソード線上で、互に対応する部分要素の順序は入れ替わらないという仮定を導入する。これによって、図-6 の探索平面上で、対応を示す点列は、左下から右上に単調上昇することになり、各決定段階で局所的に単調上昇性を守るだけで、対応の一意性が保証できる。しかし、この副作用として、図-7 に示したような、左右エピソード線上で対応する要素の順序が入れ替わる対象は、正しく処理することができないという制限が生じることに注意が必要である。

対応の適合性と連続性は評価関数により表現する。対応させる部分要素として、エピソード線上の個々の画素を用いる方法を画素対応法<sup>34)</sup>、エピソード線上で検出されたエッジを用いる方法をエッジ対応法<sup>35), 36)</sup>、

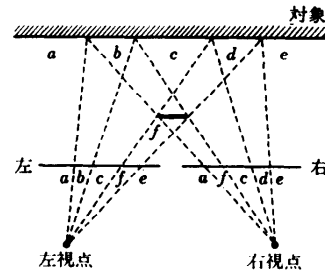


図-7 左右エピソード線上で要素の順序が入れ替わる例

エピソード線上で2個のエッジで区切られた区間を用いる方法を区間対応法<sup>37), 38)</sup>と呼ぶ。画素対応法とエッジ対応法では、適合性の評価関数は、 $d_1(x_i)$  の形である。ただし、 $x_i$  は決定段階となっている画素またはエッジを表す。一方、区間対応法では、現決定段階のエッジと、直前のエッジとの間の区間の性質を評価するため、 $d_2(x_{i-1}, x_i)$  の形の関数により適合性を評価する。連続性の評価は、直前の部分要素の対応状況との関係に基づくので、 $d_3(x_{i-1}, x_i)$  の形の関数になる。結局、エピソード線全体の対応の評価は、 $\sum d(x_{i-1}, x_i)$  の形となり、2. に示したように、再帰評価の漸化式により最適解を求めることができる。

エッジ対応法や区間対応法のように、雑音による影響を受けやすいエッジを決定段階とする場合には、雑音により生じたエッジを無視して対応を求める機構を組み込むことが実用上重要である。このため、区間対応法では、数本のエッジを跨いで区間を張る場合も対応の可能性として許している。これは、(5)式を、

$$f_n(x_n) = \min \begin{cases} \min[f_{n-3}(x_{n-3}) + d(x_{n-3}, x_n)] \\ \min[f_{n-2}(x_{n-2}) + d(x_{n-2}, x_n)] \\ \min[f_{n-1}(x_{n-1}) + d(x_{n-1}, x_n)] \end{cases} \quad (14)$$

のように変形した場合に相当するが、評価関数  $d(\quad)$  が2変数なので、計算量の組合せ的な増大はない。(14)式を高速に評価する演算装置を組み込んだ、区間対応法による立体視専用ハードウェアも試作されている<sup>39)</sup>。

### 5.3 階層的 DP の導入

階層的 DP は、レベルの異なる2種類の決定過程が組み合わせり、

- 全体の過程の評価は、上位の決定過程の評価関数で定義される、
- 上位の決定過程の各段階における評価値(すなわ

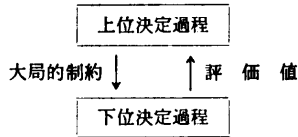


図-8 階層的 DP の構成

ち、(5)式の  $d()$  は、下位の決定過程の評価関数で定義される、

という問題に適用できる。すなわち、図-8 に示すように、上位の決定過程は、下位の決定過程に対して決定に関する大局的な制約を与え、下位の決定過程は、与えられた制約下で局所的に最適な政策を決定し、その評価値を上位の決定過程に提供する。階層的 DP は、2段 DP マッチングとして連続単語音声認識のための手法として提案されたのが最初であり、そこでは上位の決定過程として単語境界の探索を、下位の決定過程として上位から指示された境界の中での最適単語の探索を行っている。

一対のエピポーラ線間の最適対応問題は、上述のように、比較的素直に DP の枠組みで処理することができる。しかし、立体視において対応点を満たすべき三つの制約のうち、連続性の制約は、エピポーラ線内部だけではなく、隣接するエピポーラ線間でも考慮する必要がある。また、両眼立体視にもう一台カメラを追加して冗長性を増加させた3眼立体視では、考慮すべき制約も増加する。このような新たな制約を組み込んだ対応問題を DP の枠組みに載せることは一般には困難であるが、当該問題を、階層性のある2個の対応問題に分解することにより、階層的 DP が適用できる場合もある。

隣接するエピポーラ線間での連続性の制約を評価するため、大田<sup>40)</sup>は、一対のエピポーラ線の対応探索を下位の決定過程とし、エピポーラ線にまたがる連結エッジ間の対応探索を上位の決定過程として階層的 DP を適用し、一対のステレオ画像全体としての最適対応を求める手法を開発した。

水平に配置された3台のカメラによる立体視において、3本のエピポーラ線の最適対応を求める問題も、階層的 DP により解くことができる<sup>41)</sup>。この例では、3台のカメラからみえるエッジと、2台のカメラからしかみえないエッジの2種類の対応関係が存在し、Markov 性を保証しつつ一意性の制約条件を考慮するため、前者を上位の決定過程、後者を下位の決定過程に分離して扱っている。

## 6. あとがき

パターンマッチングにおける最適化問題は、一般には、変分法の枠組みで山登り的に解くことができる。しかし、DP を用いれば、収束性の心配もなく、安定確実に最適解を得ることができる。また、計算量の見積りも容易であり、実時間性の要求されるタスクにも適している。

ある問題全体を、DP が適用可能なように構成することは、必ずしも容易でない場合が多いが、DP と他の手法を部分的に組み合わせることにより、適用可能な範囲を広げることにも可能であろう。

## 参考文献

- 1) Cohen, P.R. and Feigenbaum, E.A.: The Handbook of Artificial Intelligence, Chap. 13, Kaufmann, California (1982).
- 2) Bellman, R.: Dynamic Programming, Princeton Univ. Press, New Jersey (1957). 小田中敏夫訳: ダイナミックプログラミング, 東京図書 (1973).
- 3) Bellman, R. and Dreyfus, S.: Applied Dynamic Programming, Princeton Univ. Press, New Jersey (1962).
- 4) Ballard, D. and Brown, C.: Computer Vision, Prentice-Hall, New Jersey (1982).
- 5) 松山, 興水: Hough 変換とパターンマッチング, 情報処理, Vol. 30, No. 9, pp. 1035-1046 (1989).
- 6) 坂上, 横矢: 弛緩法と正則化, 情報処理, Vol. 30, No. 9, pp. 1047-1057 (1989).
- 7) Montanari, U.: On The Optimal Detection of Curves in Noisy Pictures, CACM, Vol. 14, No. 5, pp. 335-345 (1971).
- 8) Ballard, D. and Sklansky, J.: A Ladder-Structured Decision Tree for Recognizing Tumors in Chest Radiographs, IEEE Trans. Vol. C-25, No. 5, pp. 503-513 (1976).
- 9) Amini, A. et al.: Using Dynamic Programming for Minimizing the Energy of Active Contours in the Presence of Hard Constraints, Proc. ICCV-88, pp. 95-99 (1988).
- 10) Kass, M. et al.: Snakes: Active Contour Models, Proc. ICCV-87, pp. 259-268 (1987).
- 11) Ney, H.: Dynamic Programming as a Technique for Pattern Recognition, Proc. ICPR-82, pp. 1119-1125 (1982).
- 12) 古井貞照: デジタル音声処理, 東海大出版, (1985).
- 13) 畑岡, 市川: 音声認識におけるアルゴリズム, 信学誌, Vol. 69, No. 4, pp. 306-311 (1986).
- 14) 山田博三: 輪郭 DP 整合法と手書漢字認識への



- 適用, 信学誌, Vol. J67-D, No. 3, pp. 351-358 (1984).
- 15) Yamada, H., et al.: Rotation-invariant Contour DP Matching Method for 3D Object Recognition, Proc. 1986 IEEE Int. Conf. Syst. Man, and Cyb., pp. 997-1001 (1986).
- 16) 吉田, 迫江: スタック DP マッチングによるオンライン手書き文字認識, 信学技報, PRL 83-29 (1983).
- 17) Kovalevsky, V. A.: Sequential Optimization in Pattern Recognition and Pattern Description, Proc. IFIP-68, pp. 1603-1607 (1969).
- 18) Kovalevsky, V. A.: Image Pattern Recognition, Springer-Verlag, New York (1980).
- 19) 迫江: Rubber String Matching 法による手書き文字認識, 信学研資, PRL 74-20 (1974).
- 20) 山田博三他: 2次元 DP 整合法による超音波腎臓像の認識, 信学論, Vol. J68-D, No. 9, pp. 1649-1656 (1985).
- 21) 山田博三, 山本和彦: DP 整合法による超音波心臓動画像の認識, 信学論, Vol. J71-D, No. 4, pp. 678-684 (1988).
- 22) Yamada, H. et al.: Recognition of Kidney Glomerulus by Two-dimensional Dynamic Programming Matching Method, IEEE Trans., Vol. PAMI-10, No. 5, pp. 731-737 (1988).
- 23) Yamada, H. and Kasvand, T.: DP Matching Method for Recognition of Occluded, Reflective and Transparent Object with Unconstrained Background and Illumination, Proc. ICPR-86, pp. 95-98 (1986).
- 24) 山田博三, 山本和彦: 複雑背景下のカラーテキスト文字の認識, 信学技報, PRU 88-69 (1988).
- 25) Fischler, M. A. and Elschlager, R. A.: The Representation and Matching of Pictorial Structures, IEEE Trans., Vol. C-22, No. 1, pp. 67-92 (1973).
- 26) Moore, R. K.: A Dynamic Programming Algorithm for the Distance Between Two Finite Areas, IEEE Trans., Vol. PAMI-1, No. 1, pp. 86-88 (1979).
- 27) 磯道, 小川: 動的計画法によるパターン・マッチング, 情報処理, Vol. 16, No. 1, pp. 15-22 (1975).
- 28) Barnard, S. and Fischler, M.: Computational Stereo, Computing Surveys, Vol. 14, No. 4, pp. 553-572 (1982).
- 29) 鳥脇, 白井: 手法からみたコンピュータビジョンの動向, 情報処理, Vol. 30, No. 9, pp. 1027-1034 (1989).
- 30) 大田, 池内: 3次元情報の再構成とモデル化, OplusE, Vol. 71, pp. 86-105 (1985).
- 31) Grimson, W. E. L.: Aspects of A Computational Theory of Human Stereo Vision, Proc. IUW, pp. 128-149 (Apr. 1980).
- 32) Barnard, S. and Thompson, W.: Disparity Analysis of Images, IEEE Trans., Vol. PAMI-2, No. 4, pp. 333-340 (1980).
- 33) Marr, D.: Vision, Chap. 3, Freeman, New York (1982).
- 34) 磯道, 竹政: 航空写真からの高度図の作成, 第7回画像工学コンファレンス, pp. 31-34 (1976).
- 35) 佐々木, 古川: エッジに着目した動的計画法によるステレオ画像の対応点検出, 信学技報, IE 78-97 (1979).
- 36) Baker, H. and Binford, T.: Depth from Edge and Intensity Based Stereo, Proc. IJCAI-81, pp. 631-636 (1981).
- 37) Henderson, R., et al.: Automatic Stereo Reconstruction of Man-made Targets, SPIE, Vol. 186, pp. 240-248 (1979).
- 38) 大田, 正井, 池田: 動的計画法によるステレオ画像の区間対応法, 信学論, Vol. J68-D, No. 4, pp. 554-561 (1985).
- 39) Ohta, Y. et al.: A High-speed Stereo Matching System Based on Dynamic Programming, Proc. ICCV-87, pp. 335-342 (1987).
- 40) Ohta, Y. and Kanade, T.: Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming, IEEE Trans., Vol. PAMI-7, No. 2, pp. 139-154 (1985).
- 41) Ohta, Y. et al.: Collinear Trinocular Stereo Using Dynamic Programming, Proc. ICPR-88, pp. 658-662 (1988).

(平成元年6月5日受付)