

解説



弛緩法と正則化†

坂上 勝彦†† 横矢 直和††

1. はじめに

大型の連立方程式のための数値解析法として、反復法あるいは弛緩法と呼ばれる一連の手法が広く使われている¹⁾。これらは、近似解から出発し、それを修正する計算を反復することによって真の解に収束させる手法である。弛緩法という方法論の厳密な定義は明確ではないが、その考え方は数値計算法を離れていろいろな問題に適用されている。コンピュータビジョンの分野では、弛緩法=「①全体に対して定義されるなんらかの基準を最小(または最大)にするように、②局所的な拘束(並列演算の)反復によって全体に伝播させる手法の総称」という定義がコンセンサスを得ているようである。

1970年代のWaltzのフィルタリング²⁾や、Rosenfeldらによる弛緩ラベリング(relaxation labeling)³⁾は、広義の弛緩法の適用例といえよう。その後、コンピュータビジョンの初期視覚(early vision)⁴⁾に関するいくつかの分野で、連立方程式を(本来の)弛緩法によって解くことが行われた^{5), 6)}。これらは、いずれも初期視覚を最適化問題として定式化し、変分法の枠組みで解いている。特に、Poggioらが初期視覚を正則化理論で説明⁷⁾、統一的な枠組みを提唱してからその流れはさらに大きなものとなった。

一方、ニューラルコンピューティングの分野では、ボルツマンマシン⁸⁾と呼ばれる統計力学へのアナログを基本原理としたネットワークが提案され、最適化機械としてある種のエネルギー最小化原理に基づく問題解決に有効であることが示された。このアプローチは、Geman & Geman⁹⁾によって拡張され、統計的弛緩法(stochastic relaxation)として画像復元に適用され、その後の研究に大きな影響を与えた。

こうした、弛緩法をとりまく研究の動向を図示した

のが図-1である。このようにコンピュータビジョンのいろいろな分野・問題で、広義・狭義の弛緩法が利用されている。本稿では、この弛緩法にスポットを当て、その原理と応用例を紹介する。2.では3種類の弛緩法について原理を概説する。3., 4.では共通概念であるエネルギー最小化原理について述べ、特に現在脚光を浴びている正則化理論について応用例とともに詳しく解説する。5.では正則化の弱点と指摘されている不連続性(discontinuity)への対処法について述べる。

2. 弛緩法

本章では、図-1の中の3本の大きな流れに対応する3種類の弛緩法、(1)弛緩ラベリング、(2)決定的弛緩法、(3)統計的弛緩法、について手法の違いを簡単に説明する。本解説では触れていないが、Hopfield型ニューラルネットワークで行われる処理もさらに広い意味では弛緩法の一つとみなせる。ニューラルコンピューティングについては多くのよい解説がある(たとえば文献¹⁰⁾)のでそちらを参照されたい。

2.1 Waltzのフィルタリングと弛緩ラベリング

コンピュータビジョンの分野での弛緩法といえば、Rosenfeldら³⁾によって提案された弛緩ラベリングと呼ばれる手法が有名である。原理および応用については文献¹¹⁾に解説があるため、本稿では概略を述べるにとどめる。

弛緩ラベリングとは、互いに制約し合っている対象物に矛盾が起こらないようにラベルを代入する枠組みであり、制約条件から生じる局所的な矛盾を減少させる操作を各対象物に対し並列的に反復することによって実現される。Waltzのフィルタリング²⁾がそのルーツであろう。線画中の接続点が対象物であり、接続点での可能なラベル候補から線画に対する首尾一貫したラベル付けを見つける問題である(図-2)。Waltzの用いたアルゴリズムは、それまで行われていた能率の悪いしらみつぶし法ではなく、「線画中の各線分は全長にわたって唯一のラベルしかもちえない」という

† Relaxation and Regularization by Katsuhiko SAKAUE and Naokazu YOKOYA (Machine Understanding Division Electrotechnical Laboratory).

†† 電子技術総合研究所情報情報部

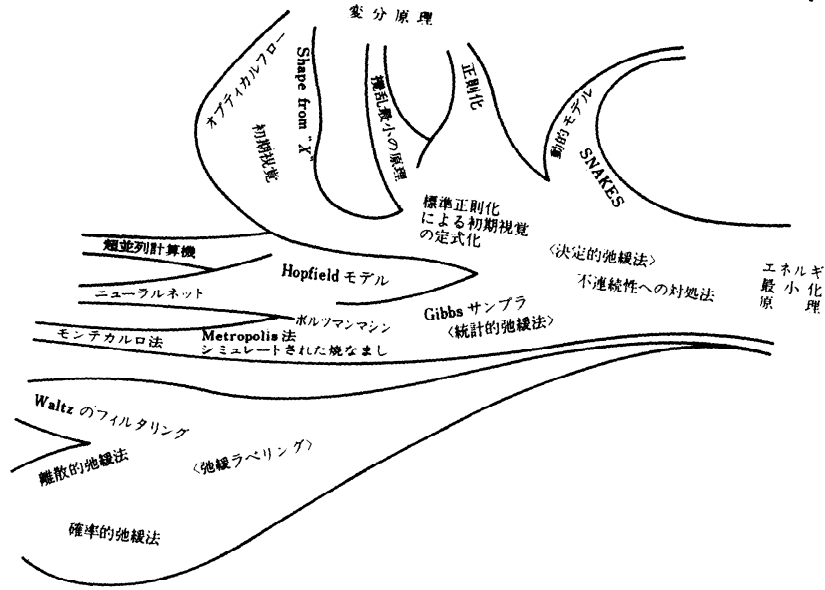


図-1 コンピュータビジョンにおける「弛緩法と正則化」研究の流れ

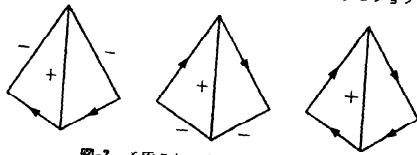


図-2 矛盾のない3種類のラベル付け

原理に基づいて、各接続点について、隣合う接続点と適合しないラベルの候補を取り除くというものである。Tenenbaum & Barrow¹²⁾によるシーン解釈も同じ枠組みとみることができる。Waltzのフィルタリングをラベル付け問題として一般化し、並列アルゴリズムとして定式化したのがRosenfeldらの離散的弛緩法(discrete relaxation)である³⁾。

離散的なラベル値は、確率を表す連続量に拡張することができる。文献³⁾ではこれを確率的弛緩法(probabilistic relaxation)と呼んでいる。画像から取り出された対象物の集合を $A = \{a_1, \dots, a_n\}$ 、これらの対象物に対する可能な解釈を示すラベルの集合を $I = \{\lambda_1, \dots, \lambda_m\}$ とする。各対象物 a_i にラベル付けの可能性を示す確率ベクトル $[p_i(\lambda_1), \dots, p_i(\lambda_m)]$ を割り当てると、 $p_i(\lambda)$ の更新ルールは、

$$p_i^{(k+1)}(\lambda) = \frac{p_i^{(k)}(\lambda) [1 + q_i^{(k)}(\lambda)]}{\sum_j p_i^{(k)}(\lambda) [1 + q_i^{(k)}(\lambda)]} \quad (1)$$

である。ただし、

$$q_i^{(k)}(\lambda) = \sum_j d_{ij} [\sum_{\lambda'} r_{ij}(\lambda, \lambda') p_j^{(k)}(\lambda')] \quad (2)$$

$$0 \leq p_i(\lambda) \leq 1, \quad d_{ij} \geq 0,$$

$$\sum_{\lambda} p_i(\lambda) = 1, \quad \sum_j d_{ij} = 1$$

ここで、 $r_{ij}(\lambda, \lambda')$ が適合係数 (compatibility coefficient) と呼ばれる重要な係数であり、線の強調の問題で使われたZucker¹³⁾による線ラベル間の適合係数(図-3)が有名である。確率的弛緩法は画像データのもつあいまい性をうまく式に乗せる手法として当時大きな流行となった。更新ルールに関しては、その後も多くの提案がある¹⁴⁾⁻¹⁶⁾。しかし、1983年ごろになると、

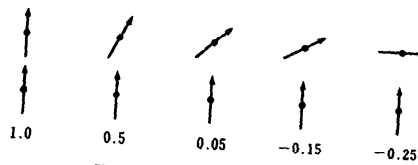


図-3 線の強調における適合係数¹³⁾

適合係数がアドホックに定義され、収束の保証がなくしばしば発散するという欠点が指摘され始めた¹⁷⁾。すなわち、前述の定義の①の部分が明確でない。

2.2 決定的弛緩法

連立方程式の数値解法としての弛緩法（あるいは緩和法）には、ガウスザイデル法、ヤコビ法、加速緩和（SOR）法などがあり¹⁾、あらゆる分野で広く使われている。コンピュータビジョンでは、なんらかの基準（たとえば、ある種のエネルギー）を画像全体で最小（あるいは最大）化する場面において、その基準を表す汎関数の極値問題を扱う変分法に対する解法として登場する。この問題の解法には直接法とオイラ方程式に訴える方法があるが、必要条件であるオイラの方程式に持ち込む場合は、この偏微分方程式を画素（あるいはデータ点）ごとに離散化することによって得られる連立方程式を解く問題に帰着する。コンピュータビジョンの問題では、この連立方程式の係数行列は大規模かつ疎となることが多いため、数値解法として弛緩法が一般に用いられる。こうしたコンピュータビジョンの分野における“変分法+連立方程式+弛緩法”の枠組みを本稿では決定的弛緩法と呼ぶことにする。

具体例によって決定的弛緩法の概要を紹介しよう。いま、まばらな観測標高データを最小曲率原理に基づく曲面（地表面）で内挿する問題¹⁸⁾を考える（図-4）。ここで求めるべき曲面 $u(x, y)$ を、観測データ点で固定された薄い弾性板とみなし、そのたわみエネルギー（2乗曲率とも呼ばれる）

$$C(u) = \iint_{\Omega} (\partial^2 u / \partial x^2 + \partial^2 u / \partial y^2)^2 dx dy \quad (3)$$

を最小にするような曲面を求めることによって内挿を行う。これは、観測データを境界条件とした汎関数の極値問題であり、変分問題になっている。式(3)に対するオイラの方程式は、次のような重調和方程式となる。

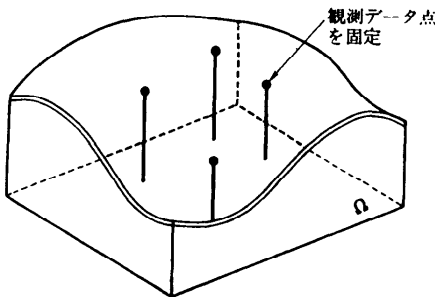


図-4 薄い弾性板モデルによる曲面の内挿

$$(\partial^4 u / \partial x^4) + 2(\partial^4 u / \partial x^2 \partial y^2) + (\partial^4 u / \partial y^4) = 0 \quad (4)$$

ところで、座標 (i, j) でのラプラシアン $c_{i,j} = \partial^2 u / \partial x^2 + \partial^2 u / \partial y^2$ を、格子間隔を h として離散的に表現すると、

$$c_{i,j} = (u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j}) / h^2 \quad (5)$$

となる。よって、式(3)を全格子点での総和に置き換えると（ただし端効果は無視する）次式のようになる。

$$C = \sum_{j=1}^N \sum_{i=1}^M (c_{i,j})^2 \quad (6)$$

C が極小となるとき、

$$\partial C / \partial u_{i,j} = 0 \quad (i=1, \dots, M; j=1, \dots, N) \quad (7)$$

となる。式(5)と(6)より、 C は $u_{i,j}$ の2次形式だから式(7)は $M \times N$ 元連立一次方程式である。さらに C の中で $u_{i,j}$ が現れるのは、 $c_{i,j}, c_{i+1,j}, c_{i-1,j}, c_{i,j+1}, c_{i,j-1}$ の項だけだから、式(7)は、

$$\begin{aligned} &u_{i+2,j} + u_{i,j+2} + u_{i-2,j} + u_{i,j-2} \\ &+ 2(u_{i+1,j+1} + u_{i-1,j+1} + u_{i+1,j-1} + u_{i-1,j-1}) \\ &- 8(u_{i+1,j} + u_{i-1,j} + u_{i,j-1} + u_{i,j+1}) + 20u_{i,j} \\ &= 0 \end{aligned} \quad (8)$$

となる。Briggs¹⁸⁾ は、

$$\begin{aligned} u_{i,j}^{(k+1)} = &(1/20) \cdot [8(u_{i+1,j}^{(k)} + u_{i-1,j}^{(k)}) \\ &+ u_{i,j-1}^{(k)} + u_{i,j+1}^{(k)} \\ &- 2(u_{i+1,j+1}^{(k)} + u_{i-1,j-1}^{(k)}) \\ &+ u_{i+1,j-1}^{(k)} + u_{i-1,j+1}^{(k)} \\ &- (u_{i+2,j}^{(k)} + u_{i,j+2}^{(k)} + u_{i-2,j}^{(k)} \\ &+ u_{i,j-2}^{(k)})] \end{aligned} \quad (9)$$

として、ヤコビ法で解いている。ただし、観測点での $u_{i,j}$ は固定である。初期値は、観測点からの線形補間などが使われる。文献¹⁹⁾の前半では、別の解法が採用されている。どちらも局所並列演算の反復で実現できる。

式(8)は、実はオイラの方程式(4)の離散表現である。結局これらの解法は、オイラの方程式を観測データを境界条件として解いたことにほかならない。

2.3 統計的弛緩法

最小化すべきエネルギー汎関数が2次形式で与えられない問題については、解くべき連立方程式は非線形となり（非凸問題）、通常の弛緩法による数値解法では最適解に到達しない。同じ問題は、相互接続型のニューラルネットワーク（たとえば Hopfield 型ネットワーク）におけるエネルギー最小化の過程でも生じる。初期値の選び方によっていろいろな局所的最小に落ち込み、最適解が得られないことが多い¹⁰⁾。そこで、ニューラ

ルネットワークの動作を、統計力学からのアナロジに基づいた確率的な状態遷移でモデル化したボルツマンマシンが提案された⁹⁾。

ボルツマンマシンは、入力信号の大きさによって発火の「確率」が変化するニューロン素子（状態は0と1の2値）によって構成されたニューラルネットワークである。確率的な状態遷移を十分に繰り返した後の定常状態では、ネットワークの状態が x となる確率は、

$$p(x) = (1/Z) \cdot \exp(-U(x)/T) \quad (10)$$

である¹⁰⁾。Zは正規化のための定数、Tは統計力学における「温度」に対応するパラメータである。U(x)は状態 x でのネットワークのエネルギーである。この式の意味は、エネルギーの小さい状態ほど大きな確率で現れ、温度が低いほどエネルギー最小の状態の出現確率が支配的になる、ということである。問題において最小にしたい基準がUと一致するように前もってネットワークを構成しておけば基準を最小にする変数の値の組合せを求めることができる。しかし、いきなり $T \rightarrow 0$ とするとつまらない局所的最小につかまるため、はじめは高い温度で自由に状態を遷移させ、ゆっくりTを0に近づければ、最終的に状態は確率1でU(x)の真の最小点へ収束する。このプロセスはシミュレートされた焼きなまし(simulated annealing)^{20), 21)}と呼ばれる。膨大な計算時間がかかるのが欠点である。

ボルツマンマシンのように、離散的なラベルへのラベル付けを確率的に行うメカニズムを統計的弛緩法^{9), 22)}と呼ぶ。原画像の確率モデルとしてマルコフ確率場モデル(MRF: Markov random field)を採用した場合、すなわち画素の状態がその近傍だけに依存する場合は、ボルツマンマシンはGibbs サンプラというより一般的な形で定式化される⁹⁾。

ここでは、画像データを想定し画素ごとに確率変数が付随しているとし、画素値 X_i のとり値の組合せ(すなわち画像)を ω とする。実は、MRF、すなわち他の画素がある状態のとき X_i がある値をもつ確率は i の近傍にしか依存しないという仮定の下では、 ω の確率分布 $\pi(\omega)$ は次のようなGibbs分布となる。

$$\pi(\omega) = (1/Z) \cdot \exp(-U(\omega)/T) \quad (11)$$

ただし、

$$U(\omega) = \sum_{c \in C} V_c(\omega) \quad (12)$$

Zは正規化のための定数である。式(12)の右辺は、画像の近傍系における局所的なエネルギー $V_c(\omega)$ の画像全体での総和である。最小化すべき画像全体でのエネルギーが式(12)の形で定式化できれば、Gibbs サンプラ

を、ボルツマンマシンと同様シミュレートされた焼きなましによって動作させることにより、最適解を得ることができる。

Gibbs サンプラの統計的弛緩法アルゴリズムを以下に示す。Metropolisの方法に基づいたシミュレートされた焼きなまし法による反復演算である。文献⁹⁾には、Tを反復の回数 k とともに $1/\log(1+k)$ のオーダーで下げていけば、厳密に最適解が達成できることが示されている。文献²³⁾ではステップ(b)でGibbs分布の代わりに一様分布が使われた例も示されている。

(a) 温度パラメータの初期値 T_0 を設定する。

(b) 状態変化 $\omega \rightarrow \omega'$ を、 ω' に対応するGibbs分布に従って確率的に選択する。(画素ごとの並列処理が可能である。逐次型アルゴリズムでは画素はランダムシーケンスで選ばれる。)

(c) $\Delta U = U(\omega') - U(\omega)$ を求め、 $\min\{1, \exp(-\Delta U/T)\}$ の確率で ω' に遷移する。

(d) 定常状態になるまで(b)~(c)を反復する。

(e) 定常状態でのエネルギーが大きく変化しなくなるまでTを下げて(b)~(d)を反復する。

重要なことは、 $\pi(\omega)$ がGibbs分布であるため、一つの状態変化による全体のエネルギーの変化量 ΔU は、局所的なエネルギー変化量と同じであるという点である。これにより、このプロセスは局所並列演算で実現することができる。応用例は5.2で紹介する。

3. エネルギー最小化原理と正則化

決定的弛緩法と統計的弛緩法は、どちらもエネルギー最小化原理に基づく問題の解決手段とみなすことができる。仮説ではあるが、人間の視覚はこのエネルギー最小化原理にある程度基づいていると考えられている。図-5は、回転する床屋のポールのオプティカルフロー(optical flow)である。実際のポールの動きは(a)であるにもかかわらず、(b)のような上下運動をしているように見える。これはもちろん錯視であるが、逆に、脳内でエネルギー最小化の計算が行われていることを予想させる。

最近のコンピュータビジョン研究では、特に画像理解の初期の段階(初期視覚)を、エネルギー最小化原理に基づいて問題を定式化する考え方が一つの動向となっている。本章では、コンピュータビジョンにおけるエネルギー最小化原理の代表的な導入例を紹介する。いずれも基本的に凸問題であり、決定的弛緩法の枠組みで解くことができる。

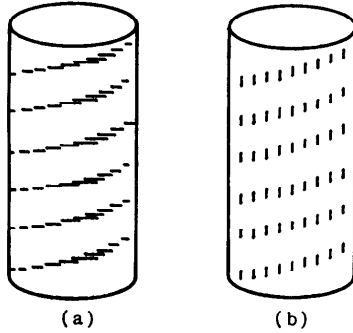


図-5 床屋のポールのオプティカルフロー

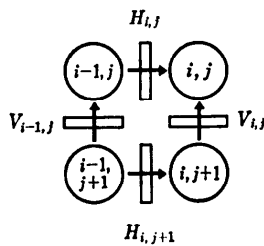


図-6 画素とエッジの位置関係

3.1 攪乱最小の原理

Blake¹⁷⁾は、濃淡画像におけるエッジの強調と検出を目的として、エッジに関する二つの拘束 (continuity constraint と minimum strength constraint) を考え、攪乱最小の原理 (least disturbance principle), すなわち初期ラベルと最終ラベルとの差 (ユークリッド距離) を最小にする最適化問題を定式化した。直感的な言い方をすれば、必要以上にラベルの更新を行わないという基準である。

エッジは連続でギャップがなく、二つの領域の共通境界上では強度が一定である。画像 $I = \{I_{i,j}\}$ に対して、画素間にエッジがあると考え、水平、垂直のエッジ強度を次式で計算する (図-6)。

$$H_{i,j} = I_{i,j} - I_{i-1,j} \quad (13)$$

$$V_{i,j} = I_{i,j} - I_{i,j+1} \quad (14)$$

このとき、continuity constraint は、

$$S_{i,j} = H_{i,j} - V_{i,j} - H_{i,j+1} + V_{i-1,j} = 0 \quad (15)$$

で表される。これは、閉ループ上でのベクトル和が0であることからきている。この H と V を初期ラベル ($H^{(0)}, V^{(0)}$) とする。minimum strength constraint とは、原画像上で、差分値 (エッジ強度) があるしきい値以下のところではエッジが存在しないという拘束である。

二つの拘束を満足させながら、攪乱最小原理にしたがって、エッジ強度の初期ラベル ($H^{(0)}, V^{(0)}$) を修正する。これは、

$$T = \|H - H^{(0)}\|^2 + \|V - V^{(0)}\|^2 \quad (16)$$

を最小化する拘束条件つき最適化問題である。ただし、 $\|M\| = \sum_{i,j} (M_{i,j})^2$ とする。

Blake は、ラグランジュの未定乗数法を用いてこの問題を拘束のない最適化問題に変換している。この最適化問題の解は、画素数の2倍の変数 (H と V) を含む線形連立方程式として与えられ、並列型の弛緩法で解を求める。反復演算の式は以下のとおりである。

$$H_{i,j}^{(n+1)} = \begin{cases} 0 & (\text{しきい値以下の弱いエッジの場合}) \\ H_{i,j}^{(n)} - (1-k)(S_{i,j}^{(n)}/r_{i,j} - S_{i+1,j}^{(n)}/r_{i+1,j}) & (\text{強いエッジの場合}) \end{cases} \quad (17)$$

ただし、 $r_{i,j}$ は $S_{i,j}$ に関係する強いエッジの数である ($1 \leq r_{i,j} \leq 4$)。 V についても同様の反復演算式が求まる。この反復演算は、パラメータ k が $0 < k < 1$ の範囲なら、必ず収束する。収束を速めるためには、 k を小さく設定すればよい。実験は並列アレイプロセッサのエミュレータで行われた。CLIP 4 の数値だと、 96×96 の画像で、50回の繰り返しに約 210 ms 要する。

最適化問題として問題をはっきりさせている点で、従来の弛緩ラベリングと明らかに異なっており、価値の高い論文である。数学的に明解であり、収束が保証されている。式(17)の補正項が、式(1)の $q_i(\lambda)$ に相当するが、弛緩ラベリングでは補正項をアドホックに前もって与える点が大きな相違点である。

3.2 正則化理論

コンピュータビジョンの初期段階における処理は、光学系を通して得られた投影像である画像から対象の物理的性質を推定する問題である。これは光学の逆問題であり、Hadamard の言う意味での不良設定問題 (ill-posed problem)²⁴⁾であることが多い。ある計算問題が良設定 (well-posed) とは、(1)解が存在し、(2)解が一意に決まり、(3)入力データの変化に対して解が連続的に変化することである。上記の条件の一つでも満たさない問題は不良設定である。

初期視覚では、特に、与えられたデータからだけでは解が一意に決まらない問題が多い。たとえば、画像からの雑音除去、両眼ステレオ、部分的に与えられた距離データからの曲面の再構成、明るさの変化から面

の傾きを求める問題などである。これまで、この種の問題では、対象に関するなんらかの仮定や拘束条件を用いることによって解を特定するというアプローチが採られてきた。

近年、これらのアプローチを正則化理論 (regularization theory) の観点から統一的に論じようという研究が盛んである⁷⁾。正則化とは、ある適切な拘束条件を用いて解の存在可能空間を制限することによって、不良設定問題を良設定問題に変える手法である。標準正則化 (standard regularization) では、未知の量 z から線形操作を受けて、データ

$$y = Az \quad (18)$$

が得られているとし、データ y から z を推定する問題の解法として、

$$E = \|Az - y\|^2 + \lambda \|Pz\|^2 \quad (19)$$

を最小にする z を求めるという最小化問題にもち込む。 $\|Az - y\|^2$ はデータと解との差を表し、ペナルティ汎関数と呼ばれる。 $\|Pz\|^2$ は、安定化汎関数と呼ばれ、解に対する現実世界の一般的な拘束条件 (たとえば滑らかさの拘束) を表す項であり、オイラ方程式がラプラス方程式や重調和方程式になる汎関数がよく用いられる。 λ は両者を考慮する割合を決める正則化パラメータである。標準正則化では、2次形式のノルムを採用しており、また P は線形作用素であるとしている。すなわち、変分法の枠組みにもち込んだ場合に、解くべき連立方程式が線形となるため、弛緩法による数値解法で容易に解が得られる。

代表的な標準正則化の適用例でもあるオブティカルフロー⁶⁾や shape from shading⁵⁾ は、正則化理論がコンピュータビジョンに適用される以前の成果であり、すでに個別に取り組みられていた問題である。それらを包含するより大きな枠組みとして正則化理論が導入された点は興味深い。いろいろな初期視覚問題を同一のアーキテクチャで解決する汎用正則化マシンなるものの開発が期待される。

実例として、オブティカルフローと曲面の再構成を取り上げ、4. で詳しく紹介する。また、解の存在可能空間の制限は、しばしば滑らかすぎる不自然な解の原因となる。5. ではこの問題の解決法について触れる。

3.3 エネルギー最小化原理に基づく動的モデル

最近、柔物体についてのマッチング問題が、力学系でのエネルギー最小化問題のアナログとして定式化されている。力学方程式に基づく動的モデルと呼ぶことのできるアプローチである。

対象のモデルの振るまい (変形) を、モデル自体の傾向と外からの拘束の線形結合によるエネルギーとして表現し、エネルギー最小の安定状態を見つける。モデルの傾向が正則化における安定化汎関数に対応し、外からの拘束がペナルティ汎関数に対応しており、解き方は正則化に似ている。ただし、不良設定問題の正則化という意味合いは薄い。

SNAKES²⁵⁾ は、外部からの強制力 (external constraint) と画像特徴 (線やエッジ) に引き寄せられる力 (image force) および、滑らかさなどの曲線内部の力 (internal energy) を最小にする曲線として得られる対話的で動的な輪郭モデルである。今、曲線 snake のパラメータ表現を $v(s) = (x(s), y(s))$ とし、上記の3種の力に対応するエネルギーを、 E_{con} 、 E_{image} 、 E_{int} とすると、

$$E^*_{snake} = \int_0^1 E_{int}(v(s)) + E_{image}(v(s)) + E_{con}(v(s)) ds \quad (20)$$

を最小化する問題を解くことによって snake すなわち輪郭の位置を決定することができる。 E_{int} には1次および2次のスプライン汎関数の和が使われる。

$$E_{int} = (\alpha |v_x(s)|^2 + \beta |v_y(s)|^2) / 2 \quad (21)$$

もっとも簡単な E_{image} は snake と線との関係であり、 I を画像の濃淡レベルとして以下の E_{iine} が定義できる。

$$E_{image} = E_{iine} = I(x, y) \quad (22)$$

外部からの強制力 E_{con} は必要に応じて使えばよい。たとえば、反発力や引っ張り力がある。

輪郭検出問題を解くには蛇 (snake) を別の方法で輪郭のそばに置かなくてはいけない。snake はその初期位置から変形を始めて局所的な最小の位置で止まる。その動きは、人間あるいは上位プログラムが E_{con} を使って制御することができる。

式(20)の離散表現を最小化する問題を変分法によって決定的弛緩法の枠組みで解くことができる。snake の初期位置が弛緩法の初期値に対応し、snake の動きは反復演算によって順次求まる近似解である。図-7は、動画像処理への応用例である。前フレームでの輪郭を初期位置として現フレームの輪郭が検出される。喜多²⁶⁾は、似た考え方の動的モデルを使って胃形状の予測を行っている。なお、snake のエネルギー最小化問題を変分法によらず、動的計画法で解こうという試みもある²⁷⁾。

Symmetry-Seeking Model²⁸⁾ は、一枚の単眼画像か

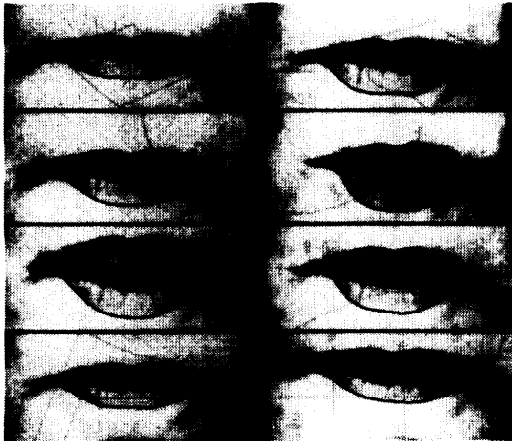


図-7 SNAKES による唇の動きの追跡²⁵⁾

ら3次元の物体を復元する方法である。次のようなモデルを考える。「弾性のあるシートを丸めたチューブに変形可能なワイヤの spine (背骨) を通し、放射状のスプリングでチューブの軸の位置に spine を保つ。チューブを spine のまわりにほぼ対称の位置に拘束する力を加える。最後に、spine から発する拡大・収縮の力を加え、形を整える。」このモデルは、画像から求められる外部からの力に応じて、その2次元画像平面への投影が画像上の対象の輪郭と一致するように変形する。解法の枠組みは SNAKES と同じである。初期状態は人間が指示しなくてはならない。図-8 に処理例を示す。

その他、コンピュータアニメーションの問題であるが、ゴム、布、紙、柔らかい金属などの変形可能な物質をモデル化した deformable model²⁹⁾ も興味深い。動きをモデル化した微分方程式を立て、微分方程式の数値解法によってモデルの形状変化の時系列を計算し、アニメーションを作成する。

4. 正則化理論の適用事例

各種の初期視覚問題に対して正則化がどのように適用できるかを表-1 に示した⁷⁾。それぞれの問題について最小化すべきエネルギー汎関数を示している。本章では、このうちオプティカルフローと曲面の再構成を取り上げ、標準正則化の初期視覚への代表的な応用例として紹介する。

4.1 オプティカルフロー

オプティカルフローとは、観測者と物体との間の相対的な運動によって生じる画面上の見かけの速度分布である。画像上の点 (x, y) の時刻 t における明るさ

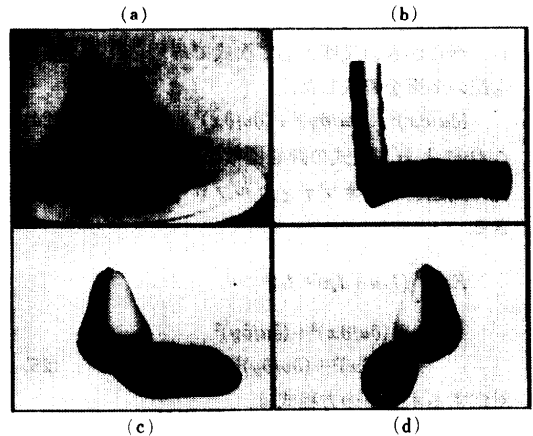


図-8 Symmetry-Seeking Model による濃淡画像からの3次元モデルの抽出²⁶⁾

- (a) 原画像
- (b) 人間が指定した3次元モデルの初期状態
- (c) 抽出されたモデルの陰影表示
- (d) モデルの側面像

表 1 初期視覚問題における標準正則化適用例⁷⁾

Problem	Regularization principle
Edge detection	$\int [(Sf-i)^2 + \lambda(f_{xx})^2] dx$
Optical flow (area based)	$\int [(i_x u + i_y v + i_t) + \lambda(u_x^2 + u_y^2 + v_x^2 + v_y^2)] \times dx dy$
Optical flow (contour based)	$\int [(V \cdot N - V^N)^2 + \lambda((\partial/\partial s) V)^2] ds$
Surface reconstruction	$\int [(S \cdot f - d)^2 + \lambda(f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2)] dx dy$
Spatiotemporal approximation	$\int [(S \cdot f - i)^2 + \lambda(f \cdot V + f_t)^2] dx dy dt$
Colour	$\ I - Az\ ^2 + \lambda \ Pz\ ^2$
Shape from shading	$\int [(E - R(f, g))^2 + \lambda(f_x^2 + f_y^2 + g_x^2 + g_y^2)] \times dx dy$
Stereo	$\int [(\nabla^* G * (L(x, y) - R(x+d(x, y), y)))^2 + \lambda(\nabla d)^2] dx dy$

を $I(x, y, t)$ で表す。今、微小時間 dt の間に物体が画面上で dx, dy だけ移動したとし、物体上の明るさが不変であると仮定すると近似的に次式が成立する。

$$(\partial I / \partial x) \cdot (dx/dt) + (\partial I / \partial y) \cdot (dy/dt) + (\partial I / \partial t) = 0 \quad (23)$$

以下では、 $(dx/dt) = u, (dy/dt) = v$ とする。式(23)は、画像中の点における明るさの時間的変化率が、明るさの空間的変化率とオプティカルフロー (u, v) との積に等しいことを意味している。しかし、視野が限られている場合は、オプティカルフローは、いわゆる密接問題によって一意には決定できない。そこで Horn ら⁶⁾

は、「速度分布は空間的に滑らかに変化する」と仮定し、滑らかさの尺度として次式で与えられる速度場の勾配の2乗を導入した。

$$(\partial u/\partial x)^2 + (\partial u/\partial y)^2 + (\partial v/\partial x)^2 + (\partial v/\partial y)^2 \quad (24)$$

これにより、次式の評価関数を最小化する標準正則化の問題としてオプティカルフローを解くことができる。

$$E = \iint (I_x u + I_y v + I_t)^2 + \lambda [(\partial u/\partial x)^2 + (\partial u/\partial y)^2 + (\partial v/\partial x)^2 + (\partial v/\partial y)^2] dx dy \quad (25)$$

対応するオイラの方程式は、

$$I_x^2 u + I_x I_y v - \lambda \nabla^2 u + I_x I_t = 0 \\ I_x I_y u + I_y^2 v - \lambda \nabla^2 v + I_y I_t = 0 \quad (26)$$

である。ただし、 ∇^2 はラプラシアンを表す。 I_x, I_y, I_t は画像データから差分によって求めた値を使い、式(27)のようなヤコビ型の弛緩法で解くことができる。

$$u^{(k+1)} = \bar{u}^{(k)} - I_x (I_x \bar{u}^{(k)} + I_y \bar{v}^{(k)} + I_t) / (\lambda + I_x^2 + I_y^2) \\ v^{(k+1)} = \bar{v}^{(k)} - I_y (I_x \bar{u}^{(k)} + I_y \bar{v}^{(k)} + I_t) / (\lambda + I_x^2 + I_y^2) \quad (27)$$

ただし、 \bar{u}, \bar{v} は4近傍での速度の平均値である。また、Kochら³⁰⁾は、この問題が立体抵抗回路によるアナログ回路で解けることを示している。

オプティカルフローによって2次元画像上で求められた動き情報から、3次元構造・運動を復元するには、次元を一つ補うためになんらかの拘束条件を用いなければならない。この条件を安定化汎関数とすれば正則化の枠組みを使うことができる。これについては文献^{31), 32)}に報告がある。

4.2 曲面の再構成

ある曲面 u にサンプリング操作 S を施した結果としてのデータ d が与えられたとして、元の曲面を推定する曲面再構成の問題を考えてみる (図-9)。まばらな観測標高データや DTM (Digital Terrain Map) データから密に標高を算出する問題と同じである。2.2 で紹介した内挿問題では観測データが保存されたが、実際には、観測データは不正確なことが多いため、図-10 のように観測点データからの許容度を与える。画素数を $m \times n$ とし、滑らかさの拘束としては2次分散 (quadratic variation) と呼ばれるものがよく用いられる。式(3)と同様、法線ベクトルに連続性をもたせた薄板スプライン (thin plate spline) である。式(19)は次式のように書くことができる。

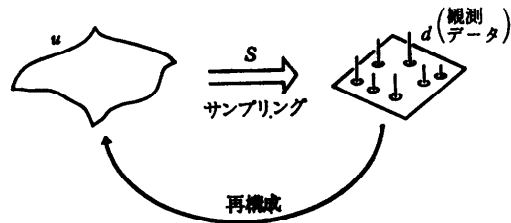


図-9 曲面の再構成

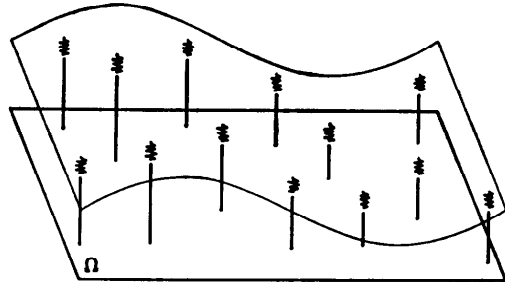


図-10 薄い弾性板モデルによる曲面の近似

$$E = \sum_{j=1}^n \sum_{i=1}^m S_{i,j} (u_{i,j} - d_{i,j})^2 + \lambda \iint_{\Omega} [(\partial^2 u/\partial x^2)^2 + 2(\partial^2 u/\partial x \partial y)^2 + (\partial^2 u/\partial y^2)^2] dx dy \quad (28)$$

これを最小にする曲面 $u_{i,j}$ ($1 \leq i \leq m, 1 \leq j \leq n$) を求めるためには、変分法により、

$$\partial E / \partial u_{i,j} = 0 \quad (1 \leq i \leq m, 1 \leq j \leq n) \quad (29)$$

という $m \times n$ 元連立方程式を解けばよい。式(28)の中の $\partial^2 u/\partial x^2$ などの偏微分を、

$$\partial^2 u/\partial x^2 = (u_{i+1,j} - 2u_{i,j} + u_{i-1,j})/h^2 \quad (30)$$

のように差分に置き換えると (h は格子間隔)、式(29)は画像 Ω 上のすべての点での標高値に関する連立一次方程式、

$$\partial E / \partial u_{i,j} = 2S_{i,j}(u_{i,j} - d_{i,j}) + \lambda [(25/h^4)u_{i,j} - (8/h^4)(u_{i+1,j} + u_{i-1,j}) - (8/h^4)(u_{i,j+1} + u_{i,j-1}) + (3/2h^4)(u_{i+2,j} + u_{i-2,j}) + (3/2h^4)(u_{i,j+2} + u_{i,j-2}) + (1/4h^4)(u_{i+2,j+2} + u_{i+2,j-2} + u_{i-2,j+2} + u_{i-2,j-2})] = 0 \quad (31)$$

となり、時間さえかければ解くことができる。Grimson は文献¹⁹⁾の後半で共役勾配法を採用している。

結果的にはスプライン曲面を求めているのにほかな

らないが、正則化の枠組みを用いてエネルギー最小化を明確に式として表現しているため、その他の拘束条件を加えるときの見通しがきわめてよいという大きな特徴をもっている。村木ら³³⁾は部分的に等高線が得られている場合を想定し、標高の高さ変化が地図の等高線上で0となる拘束、

$$\iint_{\Omega} [(\partial u/\partial x)\Delta x(x, y) + (\partial u/\partial y)\Delta y(x, y)]^2 dx dy \quad (32)$$

を式(28)に加えて良好な結果を得ている。ただし、 $(\Delta x(x, y), \Delta y(x, y))$ は画像上の点 (x, y) での等高線の方向余弦である。数値解法としては加速緩和法が採用されている。

5. 不連続性への対処法

標準正則化やエネルギー最小化原理において、解くべき連立方程式を線形にして安定に解を求めるために、滑らかさの拘束を一概に適用する例をいろいろ紹介してきた。しかし、このような方法では、しばしば滑らかになりすぎて直感に合わない結果が得られることがある。たとえば曲面の再構成では、奥行き(高さ)の不連続および面の方向の不連続は、むしろ頻繁に存在する。再構成アルゴリズムはそれらを滑らかに変えてしまうのではなく、正しく検出し明示しなくてはならない。また、オプティカルフローは、物体のオクルージョン部で不連続である。不連続な場所が未知の場合には、エネルギー汎関数が非凸になり、局所的最小をもつ。不連続性への対処法とは非凸なエネルギー関数の最小化問題への対処法と考えてもよい。本章では、決定的な手法と統計的な手法に分けて簡単に解説する。

5.1 決定的方法

Terzopoulos^{34), 35)}は、「制御された連続性の拘束 (controlled-continuity constraint)」を導入した。この拘束は、連続性制御関数 (continuity control function) による複数の次数の一般化スプラインで表現される。曲面再構成の問題に対しては、制御された連続性の拘束を表す次のような汎関数が採用された。

$$S_{\rho, \tau}(u) = \frac{1}{2} \iint_{\Omega} \rho(x, y) \{ \tau(x, y) (u_x^2 + 2u_x u_y + u_y^2) + [1 - \tau(x, y)] (u_x^2 + u_y^2) \} dx dy \quad (33)$$

ただし、 $\rho(x, y)$ と $\tau(x, y)$ は連続性制御関数である ($0 \leq \rho \leq 1, 0 \leq \tau \leq 1$)。この汎関数は、薄い弾性板のエネルギー汎関数と薄い膜のエネルギー汎関数との加重結合であり、任意点でその点での制御関数を調整することに

よって拘束の滑らかさ特性を制御することができる。基本的な制御法は、①すべての非不連続点 (x, y) において、 $\rho(x, y), \tau(x, y)$ を非零にする。②方向の不連続を形成すべきところでは、 $\tau(x, y)$ を0(あるいは0に近い値)に設定する。③奥行きの不連続点では $\rho(x, y)$ を0(あるいは0に近い値)にする。

あらかじめ $\rho(x, y), \tau(x, y)$ が分かっていたれば式(33)は2次形式になり、標準正則化の枠組みで解くことが可能である。不連続を前もって検出するには、いわゆるエッジ検出手法が使えるが、Terzopoulos³⁵⁾は曲面再構成の途中で、薄い弾性板のたわみモーメントが0のところ奥行きの不連続を検出し、極大点で方向の不連続を検出した。同様のアプローチは文献³⁶⁾にもみられる。 $\rho(x, y), \tau(x, y)$ 自体を未知として変分法に組み込むと2重の最小化過程が生じ非常にやっかいである。

Blake^{37), 38)}は、部分的に不連続があってもよいが、その場合にはペナルティを課するという考え方で、連続な部分での変動と不連続な部分でのペナルティの総和を最小にする拘束条件を weak constraint と名付けた。この問題は非凸問題である。解法として GNC (graduated non-convexity algorithm) 法が提案されている。これは、もとの目的関数 F (非凸) と F の凸包絡 F^* (凸) の間に落ちる関数列を作り、 F^* から順次、局所最適解を山登り法で求め、徐々に F での解に近づける手法である。

そのほか、Kochらは、画素の値および画素と画素の間のエッジの値に実数を割当て、Hopfield型のニューラルネットワークで最適解を求めることを試みている。パラメータの調整が大変困難なようである³⁹⁾。

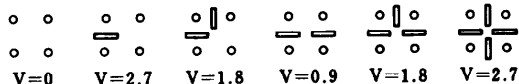
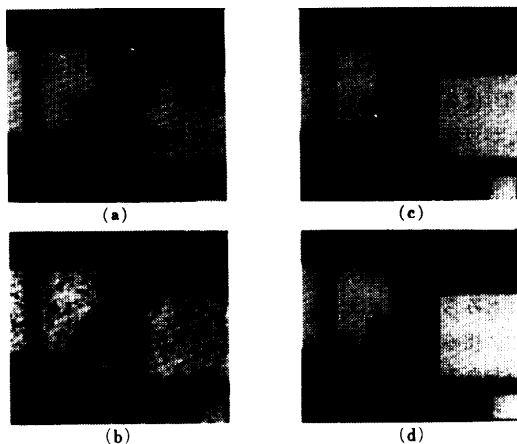
5.2 統計的方法

Geman & Geman⁹⁾は、ライン過程 (line process) というエッジの存在を表す確率変数を、濃淡レベルの確率変数とは別に準備することにより不連続の問題を解決している。区分的に一定なマルコフ確率場モデルに基づいて、観測画像から事後確率最大の意味でもっとも確からしい原画像を推定する問題である。最小化すべきエネルギーは、

$$U(I) = \sum_{i,j=1}^4 V(i, j, L_{ij}) + \sum_k V_k(L) + \sum_i (D_i - L_i)^2 / 2\sigma^2 \quad (34)$$

である。ただし、

$$V(i, j, L_{ij}) = \begin{cases} V_2(i, j) & \text{if } L_{ij} = \text{off} \\ 0 & \text{if } L_{ij} = \text{on} \end{cases}$$

図-11 エッジの配置とエネルギー⁹⁾図-12 統計的弛緩法による画像復元⁹⁾

(a) 原画像 (c) 復元結果(ライン過程なし)
 (b) 雑音による劣化画像 (d) 復元結果(ライン過程あり)

$$V_s(i, j) = \begin{cases} -1 & \text{if } I_i = I_j \\ +1 & \text{otherwise} \end{cases}$$

j は画素 i の 4 近傍画素である。 I_{ij} は画素 i, j 間に存在するエッジを意味する。このように、近傍画素間の濃淡レベルについての関係が、間にあるエッジによって分断されることにより不連続が適度に保存される。 $\sum_s V_s(I)$ はエッジの配置に関するエネルギーで、画像中のすべての近傍系についての総和である。具体的な値は図-11 に示す。 D_i は観測濃淡レベル、 I_i は原画像の濃淡レベル、 σ は加えた雑音の標準偏差である。式(34)を最小化するように、2.3 の Gibbs サンプラを動作させれば解を得ることができる。図-12 に処理結果を示す。

統計的弛緩法の欠点は計算量が膨大になることである。特に、濃淡レベルの過程は、ライン過程 (on/off の 2 値) と比べて状態数が多く、統計的弛緩法には不向きである。そこで、Marroquin ら⁴⁰⁾ は、ライン過程をまず固定し濃淡レベルのエネルギーを決定的な手法で最小化し、次にライン過程を統計的に解くことを繰り返す方式を提案している。

6. おわりに

コンピュータビジョンにおける弛緩法と正則化について、原理と応用例を紹介した。エネルギー最小化原理

に基づいた一連の手法は、得られた結果の意味が明解である。すなわち、なんらかの意味での最適解が得られる。アドホックな手法の組合せが大部分だった従来技術から一步前進したと言えよう。膨大な計算を行うのにやぶさかでないという立場も魅力的である。将来、初期視覚問題だけでなく、広く応用が試みられることを期待したい。

筆者らの経験では、線形連立方程式であっても初期値の与え方によっては正しく収束しないことがある。また、シミュレートされた焼きなまし法は論文などで示されている温度制御法 (annealing schedule) ではあまりに時間がかかり実用的でない。こういったインプリメントの際の現実的な問題についても研究の余地が残されている。

ここで紹介した一連の手法は、本質的に高並列性と拘束伝播に伴う協調型メカニズムを備えており、分散協調型計算メカニズムの典型例でもある。今後は、このような計算メカニズムに適したアーキテクチャやプログラミング・パラダイムの研究も必要となろう。

最後に、本解説を執筆するにあたり、ご討論いただいた電総研画像研究室村木茂氏、久保田鉄工山崎祐一氏に感謝いたします。

参考文献

- 1) 例えば、森 正武：数値解析、共立出版 (1973)。
- 2) Waltz, D.: Understanding Line Drawings of Scenes with Shadows, In *The Psychology of Computer Vision*, P. Winston (Ed.), McGraw-Hill, New York, pp. 19-92 (1975)。
- 3) Rosenfeld, A., Hummel, R. A. and Zucker, S. W.: Scene Labeling by Relaxation Operations, *IEEE Trans.*, Vol. SMC-6, No. 6, pp. 420-433 (1976)。
- 4) Marr, D.: *Vision*, Freeman, San Francisco, (1982).
乾, 安藤 (訳) : ビジョン—視覚の計算理論と脳内表現, 産業図書 (1987)。
- 5) Ikeuchi, K. and Horn, B. K. P.: Numerical Shape from Shading and Occluding Boundaries, *Artificial Intelligence*, Vol. 17, Nos. 1-3, pp. 141-184 (1981)。
- 6) Horn, B. K. P. and Schunck, B. G.: Determining Optical Flow, *Artificial Intelligence*, Vol. 17, Nos. 1-3, pp. 185-203 (1981)。
- 7) Poggio, T., Torre, V. and Koch, C.: Computational Vision and Regularization Theory, *Nature*, Vol. 317, No. 6035, pp. 314-319 (1985)。
- 8) Farlman, S. E., Hinton, G. E. and Sejnowski, T. J.: Massively Parallel Architectures for AI: NETL, Thistle, and Boltzman Machines, *Proc. AAAI-83*, pp. 109-113 (1983)。

- 9) Geman, S. and Geman, D.: Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images, *IEEE Trans.*, Vol. PAMI-6, No. 6, pp. 721-741 (1985).
- 10) 麻生英樹: ニューラルネットワーク情報処理, 産業図書 (1988).
- 11) 坂上勝彦: 画像処理における反復演算の応用, 情報処理, Vol. 23, No. 7, pp. 641-650 (1982).
- 12) Tenenbaum, J. M. and Barrow, H. G.: IGS: A Paradigm for Integrating Image Segmentation and Interpretations, *Proc. 3 IJCPR*, pp. 504-513 (1976).
- 13) Zucker, S. W., Hummel, R. A. and Rosenfeld, A.: An Application of Relaxation Labeling to Line and Curve Enhancement, *IEEE Trans.*, Vol. C-26, No. 4, pp. 394-403 (1977).
- 14) Peleg, S.: A New Probabilistic Relaxation Scheme, *IEEE Trans.*, Vol. PAMI-2, No. 4, pp. 362-369 (1980).
- 15) Faugeras, O. D.: Relaxation Labeling and Evidence Gathering, *Proc. 6 ICPR*, pp. 405-412 (1982).
- 16) Hummel, R. A. and Zucker, S. W.: On the Foundations of Relaxation Labeling Processes, *IEEE Trans.*, Vol. PAMI-5, No. 3, pp. 267-287 (1983).
- 17) Blake, A.: Relaxation Labelling—The Principle of "Least Disturbance", *Pattern Recognition Letters*, Vol. 1, Nos. 5-6, pp. 385-391 (1983).
- 18) Briggs, I. C.: Machine Contouring Using Minimum Curvature, *Geophysics*, Vol. 39, No. 1, pp. 39-48 (1974).
- 19) Grimson, W. E. L.: An Implementation of a Computational Theory of Visual Surface Interpolation, *Comput. Vision, Graphics, Image Processing*, Vol. 22, pp. 39-69 (1983).
- 20) Kirkpatrick, S., Gelatt, C. D. Jr. and Vecchi, M. P.: Optimization by Simulated Annealing, *Science*, Vol. 220, No. 4598, pp. 671-680 (1983).
- 21) 深尾 毅: アニーリング法 —一つの最適化手法—, 信学誌, Vol. 70, No. 12, pp. 1247-1250 (1987).
- 22) Landy, M. S.: A Brief Survey of Knowledge Aggregation Methods, *Proc. 8 ICPR*, pp. 248-252 (1986).
- 23) Murray, D. W., Kashko, A. and Buxton, H.: A Parallel Approach to the Picture Restoration Algorithm of Geman and Geman on an SIMD Machine, *Image and Vision Computing*, Vol. 4, No. 3, pp. 133-142 (1986).
- 24) Poggio, T. and Koch, C.: Ill-posed Problems in Early Vision: from Computational Theory to Analogue Networks, *Proc. Royal Society London*, Vol. B 226, pp. 303-323 (1985).
- 25) Kass, M., Witkin, A. and Terzopoulos, D.: SNAKES: Active Contour Models, *Proc. ICCV-87*, pp. 259-268 (1987).
- 26) 喜多泰代: 胃の2次元モデルによるX線像の胃領域形状の予測, 情報処理学会研究会資料, CV 60-5 (1989).
- 27) 大田, 山田: 動的計画法によるパターンマッチング, 情報処理, Vol. 30, No. 9, pp. 1058-1066 (1989).
- 28) Terzopoulos, D., Witkin, A. and Kass, D.: Symmetry-Seeking Models for 3D Object Reconstruction, *Proc. ICCV-87*, pp. 269-276 (1987).
- 29) Terzopoulos, D., Witkin, A. and Kass, M.: Energy Constraints on Deformable Models: Recovering Shape and Non-rigid Motion, *Proc. AAAI-87*, pp. 755-760 (1987).
- 30) Hutchinson, J., Koch, J., Luo, J. and Mead, C.: Computing Motion Using Analog and Binary Resistive Networks, *IEEE Computer*, Vol. 21, No. 3, pp. 52-62 (1988).
- 31) 森川, 原島: 正則化を用いる剛体・非剛体運動からの3次元構造復元, 画像理解の高度化と高速化シンポジウム講演論文集, pp. 97-102 (1989).
- 32) Yasumoto, Y. and Medioni, G.: Robust Estimation of Three-dimensional Motion Parameters from a Sequence of Image Frames Using Regularization, *IEEE Trans.*, Vol. PAMI-8, No. 4, pp. 464-471 (1986).
- 33) 村木, 山田, 横矢, 山本: 等高線との整合性を考慮した3次元曲面の生成法, 信学技報, PRU 88-77 (1988).
- 34) Terzopoulos, D.: Regularization of Inverse Problems Involving Discontinuities, *IEEE Trans.*, Vol. PAMI-8, No. 4, pp. 413-424 (1986).
- 35) Terzopoulos, D.: The Computation of Visible-surface Representations, *IEEE Trans.*, Vol. PAMI-10, No. 4, pp. 417-438 (1988).
- 36) Lee, D. and Pavlidis, T.: One-dimensional Regularization with Discontinuities, *IEEE Trans.*, Vol. PAMI-10, No. 6, pp. 822-829 (1988).
- 37) Blake, A.: The Least-Disturbance Principle and Weak Constraints, *Pattern Recognition Letters*, Vol. 1, Nos. 5-6, pp. 393-399 (1983).
- 38) Blake, A.: Comparison of the Efficiency of Deterministic and Stochastic Algorithms for Visual Reconstruction, *IEEE Trans.*, Vol. PAMI-11, No. 1, pp. 2-12 (1989).
- 39) 池田, 川人, 三宅, 乾, 淀川, 鈴木: 画像復元をする神経回路モデルのエネルギー学習, テレビジョン学会技術報告, VVI '88-25 (1988).
- 40) Marroquin, J., Mitter, S. and Poggio, T.: Probabilistic Solution of Ill-posed Problems in Computational Vision, *Proc. DARPA Image Understanding Workshop*, pp. 293-309 (1985).

(平成元年6月8日受付)