計算論的視覚と正則化理論

Computational vision and regularization theory

Tomaso Poggio and Vincent Torre and Christoh Koch (1985)

目に見える表面の物理的特性 (距離やエッジの有無など) の記述は，一次画像データから復元する必要がある。 計算論的視覚は，本質的に曖昧でノイズの多いデータから，どのようにしてそのような記述を得ることができるかを理解することを目的としている。 この分野の最近の進展は，初期視覚を，正則化法を用いて解決できる一連の不良設定問題として捉えている。 これにより，「不良設定問題」を解くことができるアルゴリズムや並列アナログ回路が開発され，脳内の神経に相当する部分が示唆されている。

計算論的視覚は，視覚情報処理の理論的研究を中心とした人工知能の新しい分野である。 画像の入力データから自動的に情景描写を行う画像理解システムの開発と，人間の視覚を理解することを目的としている。

初期視覚とは，見る人の周囲にある表面の物理的特性，すなわち，距離，表面の向き，材料特性 (反射率，色，テクスチャ) を抽出することを目的とした一連の視覚モジュールである。 現在の研究の多くは，初期視覚の処理過程を分析している。 なぜなら，この段階では，入力と計算の目標が十分に特徴づけられるからである (レビューは文献 1-4 を参照)。 いくつかの問題が解決され，いくつかの具体的なアルゴリズムの開発に成功している。 例えば，両眼視対応，光学フローの計算，動きからの構造，陰影からの形状，表面再構成などである。

現在，これらの結果の多くを 1 つの枠組みに統合する新しい理論的展開が生まれている。 このアプローチは，初期視覚問題に共通する構造を認識したことに端を発している。 初期視覚問題は，特定のアルゴリズムと並列ハードウェアを必要とする「不良設定問題」である。 ここでは，特定の正則化アプローチを紹介し，コンピュータビジョンと並列コンピュータアーキテクチャ (生物学的視覚系で使用される可能性のある並列ハードウェアを含む) への影響を議論する。

1. 初期視覚過程

初期視覚は，2 次元の強度配列から，目に見える 3 次元の表面の物理的特性を回復する一連の処理過程で構成されている。 これらの処理過程を組み合わせた出力は，Marr の 2-1/2D スケッチ(1) や Barrow&Tennenbaum の固有画像(5) にほぼ対応している。 最近では，これらの初期視覚過程は一般的なものであり，領域依存の知識を必要とせず，物理的な単語と画像化段階に関する一般的な制約のみを必要とすると仮定するのが通例となっている (ボックスを参照)。 これらの処理過程は，概念的に独立したモジュールを表しており，第一近似的には分離して研究することができる。 しかし，それぞれの処理過程から得られる情報を組み合わせる必要がある。 さらに，異なるモジュールが早い段階で相互作用することもある。 さらに，処理は純粋に「ボトムアップ」ではない。 特定の知識が浸透して，視覚情報処理の最初の段階のいくつかに影響を与えることもある。

ボックス 初期視覚過程の例

- エッジ検出 Edge detection
- 時空間の補間と近似 Spatio-temporal interpolation and approximation
- 光学フローの計算 Computation of optical flow
- 明るさとアルベドの計算 Computation of lightness and albedo
- 輪郭からの形状 Shape from contours
- テクスチャからの形状 Shape from texture
- 陰影による形状 Shape from shading
- 両眼視対応 Binocular stereo matching

- 運動からの構造 Structure from motion
- 両眼視からの構造 Structure from stereo
- 表面再構成 Surface reconstruction
- 表面色の計算 Computation of surface colour

初期視覚モジュールの計算理論は，典型的には，表現と処理過程の二重の問題を扱う。 つまり，入力と出力の形式を指定し (表現)， 一方を他方に変換するアルゴリズムを提供しなければならない (処理過程)。 ここでは，処理過程とアルゴリズムの問題に焦点を当て，正則化理論という統一された理論的枠組みを説明する。 ここでは，処理過程とアルゴリズムの問題に焦点を当て，正則化理論という統一的な理論的枠組みを説明する。 同様に重要な問題である，各特定の処理過程の入力を表す原生トークンについては考慮しない。

初期視覚の定義は，「逆光学」であるということである。 古典的な光学やコンピュータグラフィックスでは，3 次元物体の画像を決定することが基本的な問題であるのに対し，視覚では画像から表面を復元するという逆の問題に直面する。 3 次元の世界を 2 次元の画像に投影する画像処理の過程で非常に多くの情報が失われるため，視覚では，曖昧さのない出力を得るために，自然の制約，すなわち物理的世界に関する仮定に頼らざるを得ないことが多い。 このような制約条件を特定して利用することは，特定の視覚問題を分析する際の繰り返しのテーマとなる。

初期視覚における 2 つの重要な問題は，動きの計算と，画像強度の急激な変化の検出 (物理的なエッジを検出するため) である。 これらの問題は，初期視覚の問題の難しさをよく表している。 画像内の 2 次元視野における物体速度の計算は，物体の動きと 3 次元構造を回復するためのいくつかのスキームの重要な段階である。 画像内の滑らかな輪郭に沿った各点の速度ベクトル V を決定する問題を考えてみよう。 Marr&Ullman (6) に倣い，輪郭は強度が大きく変化する場所に対応していると仮定する。 図 1 は，局所的な速度ベクトルが，曲線の法線方向と接線方向の成分に分解される様子を示している。 局所的な動きの測定では，速度の法線成分のみが得られる。 接線方向の成分は，(角などの輪郭の不連続な部分を参照しない限り) 純粋な局所計測では「見えない」ままである。 したがって，完全な速度場を推定する問題は，一般的に，画像から直接得られる測定値では決定できない。 光学フローの測定は本質的に曖昧である。 情報や仮定を追加することによってのみ，一意にすることができる。

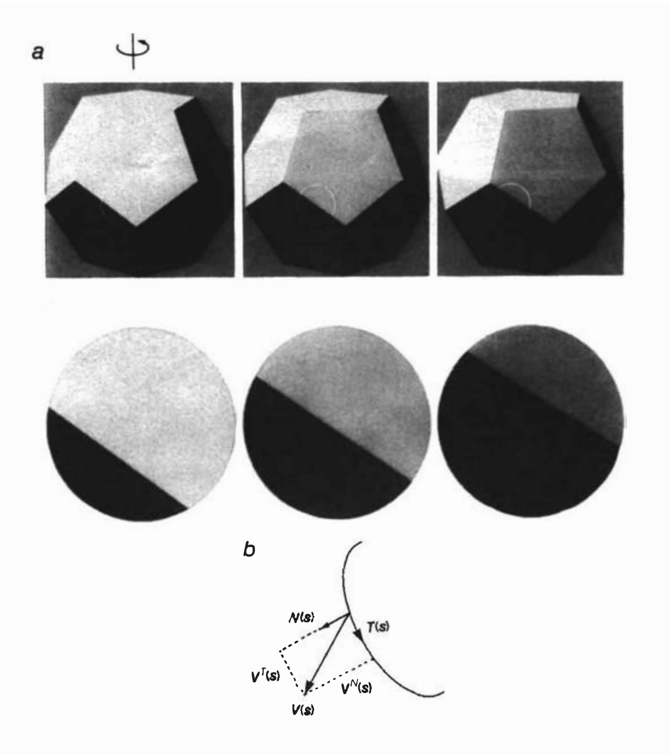

図. 1 速度場の曖昧性

- **a** 局所的な計測では， 剛体の 3 次元回転に起因する画像平面内の完全な速度場を計測することはできない (3 つのフレームを示す)。 開口部 (白丸) 内で動作する処理過程は， 輪郭に垂直な動きの成分のみを計算することができる。
- **b** 弧の長さ s で表される輪郭に沿った速度ベクトルを， 曲線に垂直な成分 ($V^N(s)$) と接線方向の成分 ($V^T(s)$) に分解している。 The computer drawing was kindly provided by Karl Sims.

エッジ検出の問題は，その難しさがやや異なる。 エッジ検出とは，画像の強度変化から 3 次元表面の物理的境界を特定する処理過程を意味する。 エッジ検出で意図しているのは，この目的のための第一歩， つまり画像の強度の急激な変化を検出して，それを局所化することである。 これは，画像データの数値微分の問題であり，撮像時やサンプリング時に避けられないノイズに悩まされる。 微分することでノイズが増幅されてしまうため，本質的に不安定な処理となってしまう。 図 3 は，エッジプロファイルとその 2 次微分の例で，ノイズが大きく増幅されている。 初期視覚の問題の多くは，同様の問題を抱えている。 光学フローの計算のように制約が少なかったり，エッジ検出のようにノイズに対してロバストでなかったりする。

2. 不良設定問題

ほとんどの初期視覚問題に共通する特徴 (ある意味，その深い構造) は，形式化することができる。 すなわち，ほとんどの初期視覚問題は，Hadamard によって定義された正確な意味での不良設定問題である (7, 8)。 この主張は制約の重要性を表しており，逆光学としての視覚の定義を反映している。

Hadamard は，偏微分方程式の分野で，不良設定問題の定義を初めて導入した (9)。 これまで，非線形問題は，ほとんど数学的な好奇心でしか考えられていなかった。 だが，現在では，非線形問題の多くは，典型的な逆問題であり，実用的な関心事であることが明らかになっている (例えば，コンピュータトモグラフィーなど)。 問題は，その解が存在し，一意であり，初期データに連続的に依存する場合には，正解となる。 逆問題とは，これらの条件のうち 1 つ以上を満たさない問題のことである。 なお，3 番目の条件は，実際に解がノイズに対して頑健であることを意味するものではない。 そのためには，問題がよく分解されているだけでなく，数値的な安定性を確保するための条件が整っている必要がある (10)。

両眼視対応，動きからの構造，光学フローの計算，エッジ検出，陰影からの形状，明るさの計算，表面の再構成など，初期視覚におけるいくつかの問題が Hadamard(8) の意味での 不良設定であることを形式的に示すことは容易い。 光学フローの計算は，輪郭に沿った法線成分から完全な速度場を復元するという「逆」問題が，一意性条件を満たさないため，不良設定である。 数値微分を目的としたエッジ検出は，解がデータに連続的に依存しないため，不良設定となる。

不良設定問題を「解決」する，つまり「良好な設定」性を回復するための主なアイデアは，適切な事前知識を導入することで，許容される解のクラスを制限することである。 事前知識とは，例えば，可能な解に制約を与える変分原理や，解空間の統計的な性質のような形で利用することができる。 ここでは，一般的な正則化という用語を，不良設定問題を正則化するために用いられるあらゆる方法に使用する。 変分正則化とは，変分原理の観点から不良問題を再定式化する正則化手法を示す。 ここでは，主に Tikhonov (11,12) に由来する，標準的な正則化法と呼ぶべき変分正則化法を紹介する (文献 13,14 参照)。 また，標準的な理論を将来的に拡張するために，初期視覚の観点からも説明する。

データ y から z を求めるという不良設定問題を正則化すると

$$\mathbf{Az} = \mathbf{y}, \tag{1}$$

は，ノルム $|\cdot|$ と安定化汎関数 $|Pz|$ の選択が必要である。 標準正則化理論では，$\mathbf{A}$ は線形演算子であり，ノルムは 2 次であり，$\mathbf{P}$ は線形である。 適用可能な 2 つの方法がある (8,13)。 (1) $\|Az - Y\| < \epsilon$ を満たす z の中から，ϵ を最小化する z を見つける。 ϵ は推定測定誤差に依存し，データがノイズ無しであればゼロとなる。

$$\|\mathbf{Pz}\|^2 \tag{2}$$

(2) は次式を最小化する z を探索する:

$$\|\mathbf{Az} - \mathbf{y}\| + \lambda\|\mathbf{Pz}\|^2 \tag{3}$$

ここで λ はいわゆる正則化パラメータである。

第 1 の方法は，データを十分に近似し，かつ最も「規則的」な，すなわち「基準」$\|\mathbf{Pz}\|^2$ を最小化する関数 z を計算することである。2 つ目の方法は，λ は解の正則化の度合いとデータへの近さの間の妥協点を制御する方法である。 標準正則化理論は，最適な λ (12,15) を決定する技術を提供する。 このように，標準正則化法は，式 (3) のコスト関数のような変分原理によって問題に制約を与える。 最小化されるコストは，何が良い解を表すかという物理的な制約を反映し，データに近く，かつ，量 $\||\mathbf{Pz}\|^2$ を小さくする必要がある。 $\mathbf{P}$ は問題の物理的制約を具体化したものである。 二次変分原理では，穏やかな条件の下では，解空間は凸であり，唯一の解が存在することが示される。 標準正則化法は，問題の非負の性質を注意深く分析した後に適用しなければならないことを指摘しなければならない。 ノルム $\|\cdot\|$ の選択，安定化関数 $\|\mathbf{Pz}\|$ の選択，関係する

関数空間の選択は，数学的特性と物理的妥当性の両方によって決定される。 これらは，正しい正則化のための正確な条件が，特定のケースで成立するかどうかを決定する。

変分原理は，物理学，経済学，工学の分野で広く使われている。 例えば物理学では，基本法則のほとんどが変分原理によってコンパクトに定式化されており，エネルギーやラグランジアンなどの適切な関数を最小化することが求められる。

3. 例

式 (3) の形式の変分原理は，過去に初期視覚で使用されていた (16-25)。 その他の問題については，標準正則化手法を用いてアプローチされている (表 1 参照)。 これまで初期視覚で使用されてきた安定化関数のほとんどは Tikhonov 型であり，目的の解 z の最初の p 個の導関数の線形結合である (文献 12)。 これらの安定化関数から得られる解は，補間スプラインまたは近似スプラインに対応する。 ここで，動きとエッジの検出の例に戻り，標準正則化技術がどのように適用できるかを示す。

表 1. 初期視覚における正則化

- エッジ検出 $\int \left[(Sf-i)^2+\lambda(f_{xx})^2\right]dx$
- 光学フロー (領域ベース) $\int \left[(i_x\mu+i_y\nu+i_t)^2+\lambda(\mu_x^2+\mu_y^2+\nu_x^2+\nu_y^2)^2\right]dx\,dy$
- 光学フロー (輪郭ベース) $\int \left[(\mathbf{V}\cdot\mathbf{N}-\mathbf{V}^N)^2+\lambda((\partial/\partial_s)\mathbf{V})^2\right]ds$
- 表面再構成 $\int \left[(S\cdot f-)^2+\lambda(f_{xx}^2+2f_{xy}^2+f_{yy}^2)^2\right]dx\,dy$
- 時空間近似 $\int \left[(S\cdot f-i)^2+\lambda(\nabla f\cdot\mathbf{V}+ft)^2\right]dx\,dy\,dt$
- 色 $\|\mathbf{I}^\nu-\mathbf{Az}\|^2+\lambda\|\mathbf{Pz}\|^2$
- 陰影からの形状 $\int \left[(E-R(f,g))^2-\lambda(f_x^2+f_y^2+g_x^2+g_y^2)\right]dx\,dy$
- 両眼視 $\int \left\{\left[\nabla^2 G*(L(x,y)-R(x+d(x,y),y))\right]^2+\lambda(\nabla d)^2\right\}dx\,dy$

初期の視覚問題の中には変分原理の観点から解決される。 最初の 5 つは， 標準的な二次正則化原理である。 エッジ検出 (26, 27) では， 画像強度のデータ $(i=i(x))$ (演算子 S は，復元されるべき連続分布 f に対するサンプリング演算子である。 同様の関数を用いて， 時間変動する画像を近似することができる。 データ $i(x,y,t)$ から復元される時空間的な強度は $f(x,y,t)$ であり，安定化項は画像平面内での一定の速度 V という制約を課している (61)。 領域に基づくオプティカルフロー (18)では， i は画像強度，u と v とは速度場の 2つの成分である。 表面再構成 (21, 22) では，表面 $f(x,y)$ は， 疎な深さデータ $d(x,y)$ から計算される。 色 (32) の場合， 明るさは 3 つの適切な色座標 $I_v(v=1,2,3)$ のそれぞれで測定される。 解ベクトル z は，照明成分とアルベド成分を別々に含み， A によって理想データに写像される。 適切な安定化項の最小化は， 空間的に滑らかな照明と， 一定または急激に変化するアルベドの制約を強いる。 陰影シェーディング (19) と立体視からの形状 (T.P. and A. Yuille, 未刊行) については， 2 つの非二次正則化関数を示す。 R は反射率地図，f と g とは表面勾配の成分に関係し，E は輝度分布である (19)。 視差の視野 d の正則化には，左画像 (L) と右画像 (R) のガウス分布の 2 階微分ラプラシアンと，視差勾配に対応するTikhonov 安定化関数との畳み込みが必要である。

直感的には，拡張された輪郭上の速度の法線成分の測定値の集合は，輪郭の大域的な動きにかなりの制約を与えるはずである。 しかし，異なる場所での局所的な測定値を組み合わせるためには，実世界の性質に関するいくつかの追加的な仮定が必要である。 例えば，V を一意に決定するためには，画像平面上での剛体運動の仮定で十分である (23,4)。 この場合，異なる場所での法線成分の局所的な測定値を直接使用して，どこでも同じであるオプティカルフローを求めることができる。 しかし， この仮定は，3 次元空間内の剛体の動きの場合をカバーしていないので，過度に制限されている (図 1)。 Hildreth は Hom&Schunck (18) に続いて，速度場に対するより一般的な滑らかさの制約を提案した (23, 24)。 実世界は滑らかな表面を持つ固体で構成されており，その投影された速度場は通常滑らかである，という物理的な考察が背景にある。 安定化項の具体的な形 (Tikhonov 安定化項) は，数学的な考察，特に解の一意性によって決定された。 2 つの正則化法は，Hildreth(23) によって提案され，実装された 2 つのアルゴリズムに対応する。 最初のものは，法線方向の速度成分 $V_N(s)$ の測定値が正確であることを前提とし，次のように最小化する:

$$\|\mathbf{PZ}\|^2 = \int \left(\frac{\partial \mathbf{V}}{\partial s}\right)^2 ds, \tag{4}$$

は，速度の法線成分の測定値を対象とする (s は弧長)。 積分は輪郭に沿って評価される。 正確ではないデータの場合，2 番目の方法は，次のように最小化することで解を得る。

$$\left\|\mathbf{V}\cdot\mathbf{N} - \mathbf{V}^N\right\|^2 + \lambda \int \left(\frac{\partial \mathbf{V}}{\partial s}\right)^2 ds \tag{5}$$

ここで，N は輪郭に対する法線方向の単位ベクトル，$\lambda^i s{-}1$ はデータの信頼性を表す。 図2a は，第 1のアルゴリズムでオプティカルフローの計算に成功した例である。

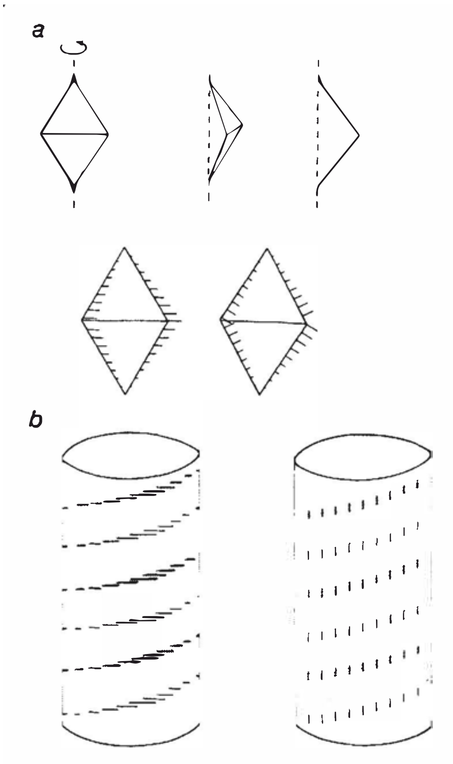

図 2 輪郭に沿って最も滑らかな速度場を計算

- **a**. Wallach (62) が最初に使用した 3 次元刺激。 物体の 2 次元の動きを投影して 3 次元の構造を導き出す人間の視覚システムの能力 (運動奥行き効果) を示す。 上の部分は， ある図形を垂直軸の周りに回転させたときの 3 つの見えを示している。 右下には， 法線方向の速度成分 V_f の初期測定値を示している。 左下には 式(4) を用いて計算された速度場が示されている。 最終的な解は， 物理的に正しい速度分布に対応している。 最近の電気生理学的証拠によると， サルの中側頭領域は，同様の運動統合が起こる部位であると考えられている (63)。
- **b**. 垂直軸を中心に回転する架空の 3 次元円柱上の円形螺旋 (床屋の円柱)。 曲線の画像平面への投影と， その結果としての 2 次元速度ベクトルが左に描かれている。 真の速度場 V は厳密には水平 (左) であるが， 最も滑らかな速度場 (右) は垂直である。 この例は， アルゴリズムと人間の視覚系の両方が同じ目の錯覚に陥る場合を示している。 文献 23 より引用。

近年，正則化技術がエッジ検出に応用されている (26,27)。 数値微分の問題は，画像に滑らかさの制約を反映させた Tikhonov 安定化項を用いた 第 2 の方法で正則化することができる (表 1)。 物理的に正当化されるのは，高い空間周波数をカットする光学系の帯域制限特性のため，画像が有界導関数であることである。 この正則化された解は，穏やかな条件の下では，以前に提案されたガウシアン (26, 図 3) に似たフィルタの導関数で強度データを畳み込むことと同等である (28,29)。

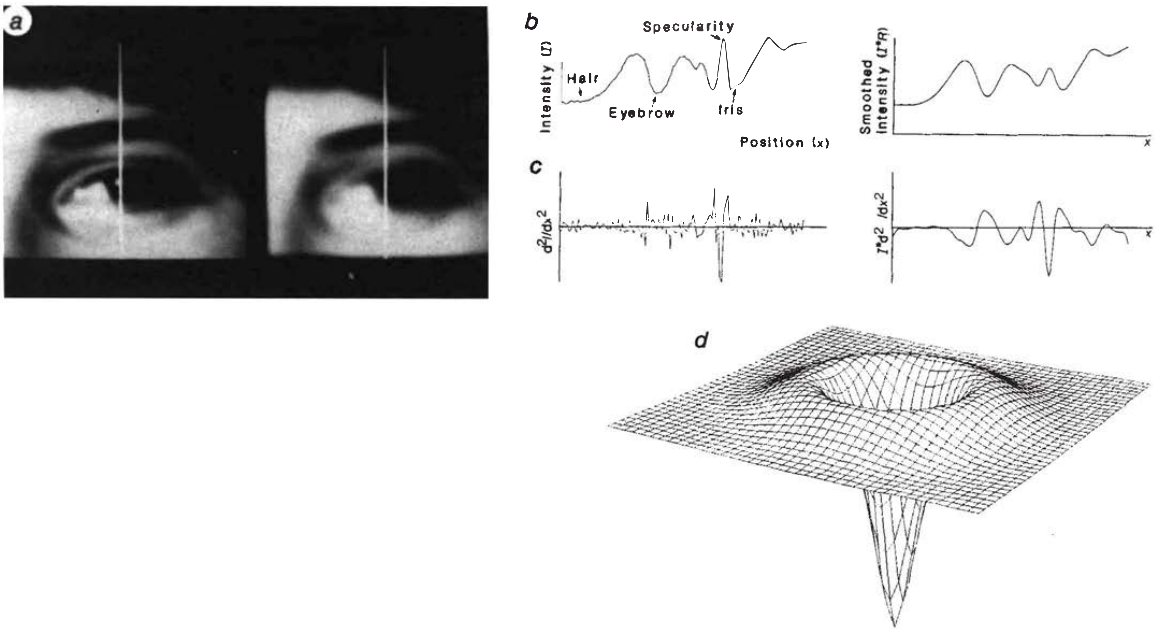

図 3. エッジ検出のための正則化解。

- **a**. デジタル画像 (256×256 画素) に何もフィルタリングしない状態 (左) と，d に示した標準的な正則化理論 (26) による 2 次元演算子でフィルタリングした状態 (右)。.
- **b**. 一次元正則化フィルタ (26) によるフィルタリングを行っていない場合 (左) と行った場合 (右) の， a で示した走査線に沿った強度プロファイル。
- **c**. 一次元正則化フィルタリングを行っていない場合 (左) と行った場合 (右) の b で示されたプロファイルの二階微分。
- **d**. エッジ検出の非正規配置問題を正規化して得られた 2 次元フィルタ (26)。 これは， ガウス分布によく似た五角形のスプラインである。 パラメータ A は，フィルタのスケールを制御する。 その値は，画像の S/N 比に依存する。 脊椎動物の網膜にあるほとんどの神経節細胞の空間受容野は， 中央に興奮性領域があり， 周囲に抑制性領域があるという非常に似た構造を持っている (64,65)。 図は Harry Vorhees 氏よりご提供いただいた。

その他の初期視覚問題は，標準正則化技術によって解決することができる。 例えば，表面の再構成は，表面の滑らかさを仮定することで，疎な深度値の集合から実行することができる (20-22)。 オプティカルフローは，Tikhonov 安定化項の形で，滑らかな変化の制約を用いることで，輪郭に沿ってではなく，画像の各点で計算することができる(17)。 正確には二次式ではないが，式 (3) の形をした変分原理は，初期視覚の他の問題にも利用できる。 実際，Tikhonov の主要な結果は，ある条件を満たせば，演算子 A と P が非線形である場合にも拡張できる(30)。 物体の明るさの変化は，その形状を知る手がかりとなる。 表面の向きは，表 1 に示す変分原理を用いて，強度画像から計算することができる。 この変分原理は，平滑性制約と照度制約に違反する向きに罰則を与える(18)。 両眼視対応とは，一対の両眼画像から，一方の画像のどの特徴が他方の画像のどの特徴に対応しているかを見つけて，正しい両眼視差 (したがって奥行き) を推定する問題である。 この問題は非可決であり，隠蔽がない場合の制限的な条件の下では，2 つの画像から抽出された特徴地図間の不一致を測定する項を含む変分原理と， 大きな視差勾配に罰則を課して実質的に視差勾配制限を課す安定化項によって正則化することができる (表 1)。 このアルゴリズムは，視差勾配が小さければ，西原型の面積ベースの相関アルゴリズム (31) に還元できる。 色画像において，空間的に変化する照明から物質の反射を分離する問題を解決するために，標準正則化原理が提案されている (32)。 このアルゴリズムは，視覚心理物理学で「色の恒常性」として知られている問題を解決するものである (33)。

4. 物理的妥当性と錯覚

正則化解析では，解の一意性よりも，解の物理的妥当性が最も重要な関心事である。 問題とその重要な制約条件の物理的分析が主な役割を果たす (8)。 正則化された解が物理的な解に対応していない特定のケースでは，不良設定問題を解くために必要な先験的な仮定が破られる可能性がある。 アルゴリズムが目の錯覚を起こしてしまうのである。 その良い例が，運動視の

計算である。 式 (5) の滑らかさの仮定は，いくつかの一般的な条件下では正しい結果を与える (例えば，物体の画像が連結した直線で構成されている場合など, 34)。 動きや輪郭のクラスによっては，平滑性の原理では正しい速度場が得られないことがある。 しかし，これらのいくつかのケースでは，人間の視覚系も同様に正しくない速度場を導き出しているようで，脳が世界について作っている先験的な仮定を明らかにしている可能性がある。 顕著な例としては，床屋の柱錯視 (23, 図 2b) がある。

5. アナログネットワーク

生物の視覚の謎の一つは，その処理速度である。 この問題を解決する方法として，並列処理がしばしば提唱されてきた。 しかし，ニューロンが単純なデジタルスイッチとは全く異なる複雑なデバイスであるという証拠が増えていることを考えると，デジタル処理によって提供される計算モデルは満足のいくものではない。 したがって，初期視覚に対する正則化アプローチが， 別種の並列計算につながるかどうかを考えるのは興味深いことである。 我々は最近，線形のアナログネットワーク (電気的または化学的) が，標準正則化理論 (7) で規定されている変分原理を解く自然な方法であることを提案した (文献 22,35 参照)。

変分原理と電気・化学ネットワークとの間にこのような対応関係が成立する基本的な理由は，Hamilton の最小作用原理にある。 アナログネットワークで計算可能な変分原理のクラスは，Kirchhoff の電流法と電圧法によって与えられ，ネットワークの各構成要素が満たす保存性と連続性の制約を表している (適切な変数は，電気ネットワークでは通常，電圧と電流，化学システムでは親和性と回転速度 (36, 37)。 一般に，ユニークなネットワークは存在しないが，同じ変分原理を持つネットワークは多数存在する可能性がある。 例えば，連想記憶の文脈で Hopfield が提案したタイプの段階的ネットワーク (38) は，標準的な正則化原理を解くことができる (39)。

Kirchhoff の法則から，ユニークな解を持つ二次変分問題 (通常はそうなる, 8) には，同じ解を持つ抵抗と電圧または電流源からなる対応する電気ネットワークが存在することが証明できる (7)。 つまり，ネットワークの定常電流 (または電圧) 分布は，標準正則化問題の解，たとえば接線速度分布 $v^T(s)$ に対応している(図 4)。 さらに，系にキャパシタンスを追加してダイナミクスを導入しても，系は安定している。 データは，電流を注入したり，電池を導入したりして，つまり定電流源や定電圧源によって供給される (7)。

このアナログ並列計算モデルは，ニューロン，膜，シナプスの生物物理学を理解する上で，特に興味深い。 多くのニューロンでは，電位が主要な役割を果たしていることを示す証拠が増えている(40)。 樹状突起のシナプス (41,42)，ギャップ結合 (43)，異なる時間と距離で作用する神経伝達物質(44)，神経ペプチドで調節可能な電圧依存性チャネル(45)，シナプスのコンダクタンス変化の相互作用 (46) などの多様な機構が，神経細胞にさまざまな異なる回路要素を提供している。 神経膜のパッチは，抵抗，容量，現象論的インダクタンスに相当する (47)。 樹状突起上のシナプスは電圧源に似ており，太い樹状突起や体幹上のシナプスは電流源として機能する (48,49)。 このように，単一のニューロンや小さなニューロンネットワークは，正則化原理のアナログ解を実装することができる。 図 4 のアナログ回路の仮想的なニューロン実装が考案されているが， これには 1 本または 2 本の樹状突起しか関与していない(7)。

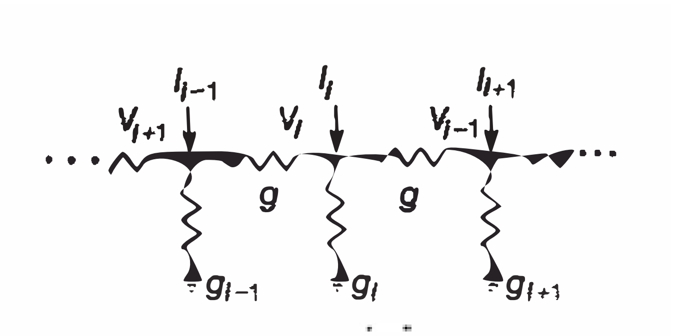

図 4. アナログネットワーク。 最も滑らかな速度場を計算する抵抗性ネットワーク(23)。 このネットワークは，法線方向の速度成分 $\mathbf{V}^N$ の測定値が正確であると仮定した場合に対応する。 関連する変分方程式 (4) を輪郭に沿って離散化すると，オイラー・ラグランジェ方程式 $(2+\kappa_i^2)\mathbf{V}_i^T - \mathbf{V}_{i+1}^T = d_i$, ここで κ_i は位置 i における輪郭の曲率,d_i はデータ V_i^N と輪郭の関数, $\mathbf{V}_i^T$ は輪郭に沿った位置 i における速度, $\mathbf{V}$ の未知の接線成分である。 電気回路の i 番目のノードを記述する方程式は， $(2g+g_i)\mathbf{V}_i^T - g\mathbf{V}_{i+1} - g_i = \mathbf{I}_i$ であり， ここで， $\mathbf{V}_i$ は未知の $\mathbf{V}_i^T$ に対応する電圧，$\mathbf{I}_i$ は計測値 $\mathbf{V}_i^N$ に応じてノード i に注入される電流である。 ここで V_i^T の測定値

が正確でない場合には，もう少し複雑な回路を設計することができる (7,式(5))。 正則化された解の一意性は，容量が導入された場合でも，対応するネットワークの安定性を常に保証する。 同等のアナログネットワークは，拡散反応系で実装することができる。 ここでは，近隣の場所間の相互作用を，一次の速度論による拡散または化学反応を用いて模倣する。 仮想的なニューロンの実装を想定することもできる。 コンダクタンス g は，樹状突起の小さなセグメントに対応し，可変コンダクタンス g は，樹状突起の静止電位に近いか等しい反転電位を持つシナプス入力 (すなわち，サイレントまたはシャント抑制) に対応し，電流源は， 樹状突起に電流Iを注入する従来の化学シナプスに対応している。 出力は化学シナプスによって位置 i でサンプリングされる。文献 7 より引用

6. 標準正則化理論を超えて

初期視覚の新しい理論的枠組みは，標準正則化理論の Tikhonov 形式に内在する魅力と限界を明確に示している。 主な問題は，復元しなければならない未知の関数に必要な滑らかさの度合いである。 例えば，表面補間において，いわゆる薄板スプラインに対応する滑らかさの度合いは，深さ方向の不連続性を滑らかにしすぎ，しばしば非現実的な結果をもたらす (20 ただし，不連続性は検出され，その後，第 2 の正則化ステップで使用されることがある, 66)。

標準正則化理論は，線形問題を扱い，二次安定化項に基づいている。 そのため，二次関数の最小化と線形オイラー=ラグランジュ方程式が導き出される。 正しい物理的制約を与えるためには，非二次関数が必要となる場合がある (表 1 は 陰影からの形状復元の非二次の場合を示している)。 この場合でも，標準正則化理論の手法を用いることができるが (30)，解空間はもはや凸形ではなく，最小化の過程で多くの局所的な最小値が見つかる可能性がある。

奥行きから表面を再構成する際に不連続性を保存する問題に対して， 非二次の安定化項が提案されている (50)。 この安定化項は，基本的には Geman&Geman (51) に起因するもので (同様の原理が Blake(52) によって提案されているが，厳密な正当化はされていない)，Terzopoulos(67) の変分連続性制御も参照)，不連続性の幾何学的性質 (ライン仮定)，特に連続的でしばしば直線的な輪郭であることについての事前知識を埋め込む。 標準正則化原理では，探索空間には 1 つの局所的な最小値しかなく，適切なアルゴリズムは常にそこに収束する。 非二次関数の場合，探索空間は多くの局所最小値を持つ山脈に似ているかもしれない。 最近，この種の最小化問題を解くための確率的アルゴリズムが提案されているが，これは，単純な山登りアルゴリズムが引っかかるような局所最小値から逃れるためである (53-55)。 その基本的な考え方は，探索アルゴリズムに強制的なノイズ項を追加することに多少似ている。 非二次変分原理を非線形アナログネットワークで表現することができれば (文献 39 のように)，適切なガウシアンノイズ源がアナログネットワークを駆動することができる。 システムのダイナミクスは，拡散過程を表す非線形確率微分方程式で記述される。

視覚の正則化理論にとって現在の課題は，標準正則化法を超えて拡張することである。 二次関数の観点から実行できる計算の世界はかなり限定されている。 これを理解するには，二次損失関数の最小化が線形正則化演算子，つまり入力データの解空間への線形写像につながることを理解すれば十分である。 データが規則的なグリッド上にあり，適切な条件に従っている特殊な場合は， 線形演算子は畳み込み，つまりデータに対する単純なフィルタリング操作になる。 物理学における線形モデルと同様に，標準正則化理論は，多くの場合，非常に有用な近似だが，視覚の複雑さを完全に扱うことはできない。

7. 正則化への確率的経路

ベイズ推定とマルコフ確率場モデルに基づく正則化の別の厳密なアプローチがある。 このアプローチでは，事前知識は適切な確率分布で表現されるが，標準正則化では事前知識は解空間の制限につながる。 例として，表面の再構成の場合を考えてみよう。 知識は，表面のマルコフ確率場 (MRF) モデルの観点から定式化することができる。 MRF では，ある離散的な場所での値は，与えられた近隣の値にのみ依存する。 この方法では，最適な表面は，最大事後推定値や MRF の事後平均値など，ある種の尤度基準を最大化する。 MRF の最大事後推定は，式 (3) の一般的な形式の変分原理と同等であることが指摘されている (50)。 第 1 項はデータと解の間の不一致を測定し，第 2 項は解の任意のポテンシャル関数 (離散格子上で定義される) となる。 一般的には二次関数ではない変分原理は，ノイズが加法性でガウス性であり，場の一次差分がゼロ平均で独立したガウス性の確率変数である場合，標準的な正則化タイプの二次関数に減少する。 この場合，最大事後推定値 (MAP) はすべての推定値と一致し，特に事後平均と一致する。 しかし，Marroquin(56) は最近，一般的にはそうではないことを示した。 ほとんどの場合，MAP 推定値は自然な誤差測定に関して最適ではなく，事後平均のようなより良い推定値を見つけることが

できる。 このような場合，この問題はエネルギー関数の大域最小値を求めることと同等ではない。 シミュレーテッドアニーリングは必要なく，代わりにメトロポリス・アルゴリズム(55) を使用することができる。

Hildreth の運動視の計算 (23) の場合，滑らかさの仮定は，輪郭に沿った近隣点間の速度変化がゼロ平均の独立したガウス分布の確率変数であるという仮説に対応する。 確率論的アプローチと標準正則化法の間のこの接続は，制約条件の性質と安定化項の選択について興味深い視点を与えてくれる。 視覚の逆問題を解くのに使われる変分原理は，もっともらしい解を生成するマルコフ構造に対応している。

今後の研究課題として，正則化演算子の学習問題が挙げられる。 標準正則化の場合，データを解に写像する対応する線形演算子は，生物学的記憶 (58) に関連して提案されたタイプの連想学習スキーム (57) によって学習することができる。

8. 記号的描写を目指して

ここまでは，初期視覚段階である，見る人の周りにある物理的な 3 次元の表面を画像のように表現することに限定して説明してきた。 固有画像 (5) や 2-1/2Dスケッチ (1) と呼ばれるこれらの表現を超えるステップは，大きなものである。 固有画像は，まだ画像的な数値表現であり，物体としては記述されていない。 固有画像は，操作やナビゲーションなど，視覚系の高レベルな課題の一部にはすでに十分である。 しかし，より記号的な表現の生成と使用を必要とする認識や記述の課題には，直接使用することはできない。 非正則問題の正則化という観点から，記号的表現の計算がどのように適合するのか，最初は理解できないように思われれる。

すべての正則化手法の基本的な考え方は，可能な解の空間を制限することである。 この空間が有限の次元を持つように制限されていれば，逆問題が正則化される可能性が高くなる。 このように，離散的な記号の有限集合に基づく表現は，不良設定問題を正則化する。 この観点から，知覚の問題 (物理世界の一般的な制約を用いて，制約の少ない問題を正則化する) は，古典的な人工知能の解法と推論の問題と実質的に同等になる。 すなわち，解法の探索を制限することで，難解な問題 (チェスなど) を解く方法を見つけることができる。

9. おわりに

我々は，現在開発中の並列デジタルコンピュータアーキテクチャに自然に対応する視覚アルゴリズムの分類を提案する。 標準正則化が十分であれば，2 つのクラスの並列アルゴリズムになる。 最急降下法のような凸関数の最小値を求めるアルゴリズムや，ビジョン用に開発されたより効率的なマルチグリッドアルゴリズム (59) は常に使用することができる。 データが規則的なグリッド上に与えられ，式 (1) の A が空間不変である場合，これらのアルゴリズムは畳み込みアルゴリズムで置き換えることができる。 後者の場合，正則化解は，あらかじめ計算されたフィルタを通してデータを畳み込むことで得られる。

これらのアルゴリズムはすべて，局所的接続しか持たない多数のプロセッサからなる並列アーキテクチャで実装できる可能性がある。 正則化の観点からアプローチできない問題や，記号的な表現とその演算を必要とする問題では，現在開発中の Connection Machine (60) のような大域的な通信機能を備えた並列アーキテクチャが必要になるかもしれない。

初期視覚の問題の多くは，不良設定問題という概念と，それに関連する新旧の正則化理論によって，満足のいく理論的枠組みが提供されてるようである。 また，この新しい視点は，初期視覚問題の計算上の性質 (非負の問題) と，それを解くアルゴリズムの構造，そして効率的な視覚情報処理に使用できる並列ハードウェアとの間に関連性をもたらす。 また，初期視覚でこれまで使われてきた変分原理の本質的な限界を示し，同時に正則化分析を標準化理論を超えて拡張する方法を示している。