

# 正則化の拡張定式化 Extended Regularization Formula

武川直樹  
Naoki Mukawa

NTT ヒューマンインタフェース研究所  
NTT Human Interface Laboratories

**Abstract** This paper presents an extended regularization formula to solve computer vision problems. Conventional standard regularization theory assumes that measurement noise has a fixed variance. This often produces insufficient estimation results since the variance of measurement noise actually varies in space and time. The conventional formula cannot be applied to sensor data integration since measurement noise differs among data. In this paper, a covariance matrix of measurement noise is introduced into the regularization formula as a weighting term of the evaluation function. This formula improves estimation, especially for sensor data integration. The formula can be applied to many computer vision problems. Here, it is applied to motion field estimation. The motion field in an image sequence is estimated using data obtained from two filters: one is a low-pass-filtered image and the other is a derivative image with low-pass-filtering. This procedure is equivalent to sensor data integration. Simulation results show that the extended regularization formula exhibits excellent performance.

## 1 まえがき

正則化理論は、測定データから元のデータを推定する、いわゆる逆問題を解く理論的枠組みとして確立され [3],[4]、ビジョンの問題を扱う手法として整備されつつある。この理論は、これまでのビジョンのアルゴリズムが、環境変動、ノイズなどにより推定結果が不安定になり、現実には適用範囲が限定されるという指摘 [1],[2] に対する一つの解決手法を示したものである。

標準正則化理論においては、未知の量  $\mathbf{x}$  から線形操作  $A$  を受けたデータ  $\mathbf{y}$  が観測され、その関係が、

$$\mathbf{y} = A\mathbf{x}, \quad (1)$$

と表わせるとき、データ  $\mathbf{y}$  から  $\mathbf{x}$  を推定する問題を次の目的関数  $E$  を最小化する最適化問題として定式化する。

$$E = \|A\mathbf{x} - \mathbf{y}\|^2 + \lambda \|P\mathbf{x}\|^2. \quad (2)$$

ここで、 $P$  は線形作用素である。式 (2) の第 1 項は観測

データと解の差を表わし、ペナルティ汎関数という。また、第 2 項は、安定化汎関数と呼ばれる解に対する現実世界の制約条件を示す。 $\lambda$  は二つの条件の割合を決めるパラメタである。この式の最小化は、観測データとできるだけ一致し、かつ制約条件をよく満たす解を求めることを意味している。

標準正則化理論の枠組により種々のビジョンの問題は最適化問題を解くことに帰着される。しかしながら、次のような問題点がいまだ解決されていない。(1) 観測誤差が適切に定式化されていない。ペナルティ汎関数と安定化汎関数のバランスを示す単一の係数  $\lambda$  では、個々のセンサ、時々刻々のセンサの観測誤差の評価を表現できない。いま、係数  $\lambda$  の物理的意味を考えると、 $\lambda$  は二つの汎関数の重みの比と、システムの平均的な観測誤差の分散の積になっている。このうち、後者は一定な平均値としてのみの扱いをしているため最小化により得られる推定精度は不十分である。(2) 複数センサのデータ統合への適用が難しい。係数  $\lambda$  のみではセンサの違いによる観測

誤差を適切に評価できない。

本論文では上記の問題の解決法を提案する。そのため、個々の観測ベクトルに対する観測雑音、安定化汎関数におけるシステム雑音の評価の項を正規化評価関数の重みづけの項として導入する。即ち、従来の定式化を拡張して、(1) 従来手法における誤差の評価量  $\lambda$  がスカラーであるのに対し、観測雑音の評価を各センサ、各観測時刻における共分散行列で与える。これにより、評価関数の精度の向上を狙う。(2) 同様にシステムの変動要因も共分散行列で与える。(3) 複数センサのデータ統合がその定式化に含まれる形式とする。この定式化により、時々刻々のセンサデータの信頼性の変動に適應して、精度のよい推定を可能にするとともに、複数センサのデータの統合を実現した。

上の定式化の適用例として、動画解析における複数センサのデータ統合の手法を提案する。具体的には、同一の画像から異なるフィルタを施して得られた二つの画像（複数センサ出力）に対して拡張定式化を適用し、従来の式を適用するのに比べより精度のよい解が得られることを示す。

本論文の構成は以下のようになっている。第2章では、観測過程に雑音を含む場合、標準正規化における定式化をより一般化した式で扱うべきことを主張する。また、拡張定式化がビジョンの様々の問題を統一的に解くための有効なアプローチであり、特に、複数のセンサを用いてセンサデータを統合する場合に従来解けなかった問題を解くことができることを示す。第3章では、拡張定式化を動画解析における動き場推定に適用するための準備として、従来の正規化の定式化に基づく動き場推定手法について述べる。第4章では、拡張定式化を用いた動き場推定アルゴリズムを提案する。複数のフィルタにより得られた画像シーケンスに対し、画素毎に観測雑音の評価値を算出し、その値に基づいて拡張定式化を適用するアルゴリズムを示す。第5章では、CG 画像および実際のカメラからの画像を対象にして提案アルゴリズムをシミュレーションにより評価する。

## 2 正規化理論の定式化の拡張

まず、観測システム、システム生成に関する順プロセスとしての定式化を行う。時刻  $t$  におけるセンサによる観測過程を、

$$\mathbf{y}_t = \mathbf{A}_t \mathbf{x}_t + \mathbf{v}_t, \quad (3)$$

と線形システムにより表現できると仮定する。但し、 $\mathbf{y}_t$  は観測ベクトル、 $\mathbf{A}_t$  は観測行列、 $\mathbf{x}_t$  はシステムの状態ベクトル、 $\mathbf{v}_t$  は観測におけるガウス雑音ベクトルであ

る。ここでは、雑音の分布の形状は既知であるが、その分散は時間的、空間的に変化するものとする。次に、対象とするシステムが次式のように表されると仮定する。

$$P_t \mathbf{x}_t + \mathbf{w}_t = 0. \quad (4)$$

ここで  $P_t$  はシステムの拘束を記述するシステム行列、 $\mathbf{w}_t$  はガウス雑音ベクトルである。この式は、システムの拘束  $P_t \mathbf{x}_t = 0$  からの変動を雑音により記述している。逆問題は、以上の二つの式が与えられたとき、 $\mathbf{x}_t$  を推定する問題である。

いま、観測雑音ベクトル  $\mathbf{v}_t$ 、システム雑音ベクトル  $\mathbf{w}_t$  の共分散行列を  $R$ 、 $Q$  とすると、共分散行列で重みづけをした

$$E_v = E \left[ (\mathbf{A}\mathbf{x} - \mathbf{y})^T R^{-1} (\mathbf{A}\mathbf{x} - \mathbf{y}) \right] + \lambda E \left[ (\mathbf{P}\mathbf{x})^T Q^{-1} (\mathbf{P}\mathbf{x}) \right], \quad (5)$$

を  $\mathbf{x}$  を求めるための評価関数とすることができる [9]。この式が従来の正規化の定式化を拡張した式である。但し、 $E(\cdot)$  は、その期待値を表す。

さて、いま、観測データ間に相関がない場合を考える。このとき、共分散行列の  $mn$  成分は、

$$(R)^{mn} = (\sigma_r^m)^2 \delta^{mn}, \quad (6)$$

となる。但し、 $m = n$  に対して  $\delta^{mn} = 1$ 、 $m \neq n$  に対して  $\delta^{mn} = 0$  である。同様に、システム雑音間に相関がない場合、その共分散行列の  $mn$  成分は、

$$(Q)^{mn} = (\sigma_q^m)^2 \delta^{mn}. \quad (7)$$

と表すことができる。さらに、各  $\sigma_r^m$ 、 $\sigma_q^m$  の値が  $m$  について等しいとき、従来の標準正規化理論における評価関数、

$$E_v = \|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 + \lambda \|\mathbf{P}\mathbf{x}\|^2, \quad (8)$$

と一致する。

拡張定式化は、従来の式に対して以下のような特長を有する。(1) 従来の式のパラメタ  $\lambda$  がスカラーで、かつ固定値であるのに対し、拡張定式化では各センサの各時刻、各空間位置における観測雑音の共分散行列で重みづけする。その結果、従来より精度のよい解が得られる。(2) 同様にシステムの変動要因を共分散行列で与えることにより精度のよい解が得られる。(3) 複数センサのデータ統合がその定式化のなかに自然に扱える形式となる。この点については、以下で更にくわしく述べる。

さて、標準正規化理論が適用できる多くの初期視覚の

問題は、拡張定式化を適用できる。観測ベクトルの誤差の共分散行列の知識が事前にあれば、あるいは、観測を通じて推定可能であれば、従来の式を用いるときに比べ、よりよい推定が可能となる。また、システムの安定化拘束についても変動（雑音）の共分散行列が推定可能であればよりよい推定が実現できる [8] が、ここではこれについてはこれ以上述べない。

拡張定式化の長所は (3) で示した複数センサデータの統合ができることにある。センサデータ統合は、センサ一つでは十分なデータ復元ができない場合、センサを複数備えてそれらのセンサデータ出力を統合することにより精度のよい解を得ることを狙うものである。センサデータ統合は、正則化理論の適用を前提にした場合、安定化汎関数を持つシステムにおいて、複数のセンサからの観測データに基づいてある目的関数を最小化する問題となる。従って、複数観測データに対する目的関数を定式化する必要がある。いま、観測過程を

$$\begin{aligned} y_1 &= A_1 x + v_1 \\ y_2 &= A_2 x + v_2 \\ &\vdots \\ y_m &= A_m x + v_m, \end{aligned} \quad (9)$$

とすることにより複数センサの観測システムを記述する。ここで、式 (9) は、次のように次元の大きい単一の観測行列により記述することができる。即ち、観測過程を、

$$y_M = A_M x + v_M \quad (10)$$

のように考える。但し、ここで、

$$y_M = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \quad (11)$$

$$A_M = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} \quad (12)$$

$$v_M = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \quad (13)$$

である。観測関数を式 (10) とすることにより式 (5) の拡張定式化がそのまま適用できる。

いま、簡単のため各センサ間の観測雑音には相関がないとして、センサの観測雑音の分散を考える。分散の値はセンサによって異なるから複数センサにおける観測の式 (10) の観測ベクトルに対する共分散行列は対角要素が等しくならない。そのため従来の定式化では推定誤差が

大きくなる。また、センサの観測雑音間に相関があるときにも問題は同様である。即ち、拡張定式化は精度のよい推定を行うために必須となる。

本論文の以下の章では、動画像シーケンスから動き場を推定する手法として、画像に複数のフィルタを施して得られた複数画像（これが複数センサ出力に対応する。）に対して拡張定式化を適用することを提案する。次章はその準備として、勾配法の原理、標準正則化理論に基づく動き場推定手法について述べる。

### 3 標準正則化理論に基づく動き場推定の原理—従来手法

勾配法により 2 次元動き場を求めるためには、画像の輝度勾配と速度の局所的な拘束関係（第 1 の拘束）と、2 次元動き場の空間的な変動が小さいという動きの滑らかさの拘束（第 2 の拘束）が利用される。第 1 の拘束からペナルティ汎関数、第 2 の拘束から安定化汎関数を構成し、標準正則化の手法を適用して 2 次元の動き場を得る [5],[6]。

勾配法においては、 $x-y$  平面からなる 1 枚の画面の時間  $t$  における点  $(i, j)$  の輝度を  $I(i, j, t)$  と記述し、動き量を  $u, v$  ( $x$  方向の動きを  $u$ 、 $y$  方向の動きを  $v$ ) とする。また、点  $(i, j, t)$  における画像の  $x$  方向の輝度勾配を  $I_x = \partial I(i, j) / \partial x$ 、 $y$  方向の輝度勾配を  $I_y = \partial I(i, j) / \partial y$  とする。このとき、動きと輝度勾配の間の拘束式は、

$$f = I_x u + I_y v + I_t = 0 \quad (14)$$

と表せる。 $I_x, I_y, I_t$  は画像から計算できるので、この式は未知数 2 個の線形方程式になる。式 1 つに対して 2 つの未知数があることから、この式のみでは輝度勾配の方向の動きしか原理的に求めることができないが、2 次元動き場の空間的な変動の滑らかさを仮定することにより勾配方向に垂直な成分をも求める手法が提案されている [5][6]。

この手法は、時間間隔  $T$  だけ離れた 2 フレームを用い、画面内の 2 次元動き場の空間的な変動（滑らかさ）のパワーの評価関数  $E_p$  と 2 次元動き場の輝度勾配と速度場の拘束条件である式 (14) からの誤差のパワーの評価関数  $E_d$  を定義し、 $E_p$  と  $E_d$  の和が最小となるフローを求めるものである。 $E_d$  が標準正則化理論におけるペナルティ汎関数、 $E_p$  が安定化汎関数に対応する。即ち、評価関数  $E_o$  を、

$$\begin{aligned} E_o &= E_p + E_d \\ &= \sum_{i,j} (u_x^2 + u_y^2 + v_x^2 + v_y^2) \\ &\quad + \lambda \sum_{i,j} (I_x u + I_y v + I_t)^2. \end{aligned} \quad (15)$$

と記述する。ここ  $u_x = \partial u(i, j) / \partial x$ 、 $v_x = \partial v(i, j) / \partial x$ 、 $u_y = \partial u(i, j) / \partial y$ 、 $v_y = \partial v(i, j) / \partial y$  であり、また、二つの汎関数のバランスを定めるパラメタ  $\lambda$  は画像に含まれる平均の誤差、想定される動きの複雑さなどにより事前に設定するものである。この評価関数  $E_v$  を最小化することにより一意な 2 次元動き場  $u(i, j)$ 、 $v(i, j)$  を求めることができる。

動き場推定における式 (15) の正則化の定式化について調べるため、観測システムについて検討する。勾配法の観測過程は、

$$f + r = I_x u + I_y v + I_t r = 0, \quad (16)$$

で与えられる。ここで、 $r$  は観測雑音である。今、観測雑音  $r$  の期待値が画面全体にわたって一定であると仮定すると、従来の式 (15) により良好な推定結果を得ることができる。しかしながら一般に雑音の性質は局所的に変化するためこの仮定は正しいとはいえない。例えば、輝度勾配が小さく、その変化の小さい領域では  $r$  の分散の期待値も小さいと考えられる。一方、勾配の大きい領域では、式 (16) の各観測値  $I_x$ 、 $I_y$ 、 $I_t$  の値の誤差が大きくなり、 $r$  の分散も大きくなる。また、輝度の極値の領域では、輝度の空間勾配がフレーム間で変化するため観測誤差は非常に大きくなるものと考えられる。このように観測雑音は画像の局所的な性質により変化することがわかる。従って、従来のようにペナルティ汎関数と安定化汎関数のバランスを示す単一の係数  $\lambda$  を用いたのでは、最小化のための適切な制御を行っているといえない。

#### 4 拡張定式化に基づく動き場推定法 – 複数センサーデータ統合への適用

本章では、拡張定式化を適用して動き場を推定するための二つのアイデアを提案する。その一つは観測した輝度勾配に応じて観測誤差の分散を推定する方法である。もう一つは、同じ画像に対して異なる特性のフィルタを施すことにより実質的に二つのセンサー出力を得、これらに対して正則化の拡張定式化を適用することによってセンサーデータ統合を実現する方法である。

まず、観測誤差の分散を推定する方法について述べる。勾配法は、輝度勾配が緩やかな領域では式 (16) の観測誤差  $r$  は小さいが、前章で述べたように輝度の勾配の大きい領域、極値の周辺の領域では観測誤差が大きくなる。そこで勾配の値に基づいて観測誤差の共分散行列を推定したうえで、式 (5) に対応する評価関数を最小化すればより精度の高い動き場が推定できる。

一方、複数センサーデータを統合することにより、極値周辺の推定精度向上が以下の方法により実現できる。まず輝度値に空間微分を施すと、輝度の極値の周辺では輝度勾配（微分値）は、その符号が変化して零を横切る。元の画像の輝度の変動が小さいと仮定すると微分値はその近傍で線形と見なすことができる。微分画像出力に対し、さらにローパスフィルタを施せばより雑音の影響を受けにくい出力  $J$  を得る。ここでは、元の画像にローパスフィルタを施した画像  $I$  と、元の画像を空間的に微分した後ローパスフィルタを施した画像  $J$  の二つの画像から結果を統合して動き場を推定することを考える。

まず、ローパスフィルタ画像  $I$  についての観測系は、

$$f + r_1 = I_x u + I_y v + I_t r_1 = 0 \quad (17)$$

で与えられる。ここで、 $r_1$  は画像  $I$  に関する観測系の雑音である。また、“微分 + ローパスフィルタ” 画像  $J$  についての観測系は、

$$g + r_2 = J_x u + J_y v + J_t r_2 = 0 \quad (18)$$

で与えられる。但し、 $r_2$  は  $J$  に関する観測系の雑音である。さて、本章では画面全体における動き場を求めることが目的であるから、式 (5) を画面全体に適用して、

$$E_v = \sum_{i,j} \left[ (A\mathbf{x} - \mathbf{y})^T R^{-1} (A\mathbf{x} - \mathbf{y}) \right] + \lambda \sum_{i,j} \left[ (P\mathbf{x})^T Q^{-1} (P\mathbf{x}) \right]. \quad (19)$$

を最小化の評価関数とする。

次に、この評価関数を最小にする  $\mathbf{x}$  を求める。まず、観測系は全微分の形式となっているため見かけ上、観測ベクトルが現れないが、 $I$ 、 $J$  の空間時間微分が観測値であり、

$$\begin{pmatrix} I_x & I_y & I_t \\ J_x & J_y & J_t \end{pmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = 0. \quad (20)$$

と書ける。一方、システムの安定化関数は、形式的に、

$$\begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} \begin{bmatrix} u & u & v & v \end{bmatrix} + \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = 0, \quad (21)$$

と表すことができる。これより最小化すべき評価関数は、

$$\begin{aligned}
E_v &= \sum_{i,j} \left( \begin{bmatrix} f \\ g \end{bmatrix}^T \begin{pmatrix} \sigma_{r1}^2 & 0 \\ 0 & \sigma_{r2}^2 \end{pmatrix}^{-1} \begin{bmatrix} f \\ g \end{bmatrix} \right) \\
&+ \lambda \sum_{i,j} \left( \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \end{bmatrix}^T \begin{pmatrix} \sigma_{q1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{q2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{q3}^2 & 0 \\ 0 & 0 & 0 & \sigma_{q4}^2 \end{pmatrix}^{-1} \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \end{bmatrix} \right) \\
&= \sum_{i,j} \left[ \frac{(I_x u + I_y v + I_t)^2}{\sigma_{r1}^2} + \frac{(J_x u + J_y v + J_t)^2}{\sigma_{r2}^2} \right] \\
&+ \lambda \sum_{i,j} \frac{u_x^2}{\sigma_{q1}^2} + \frac{u_y^2}{\sigma_{q2}^2} + \frac{v_x^2}{\sigma_{q3}^2} + \frac{v_y^2}{\sigma_{q4}^2}, \quad (22)
\end{aligned}$$

で与えられる。λは、シーンにおける事前の知識に基づいて決定されるパラメタである。また、 $\sigma_{rn}$ 、( $n=1,2$ )  $\sigma_{qn}$ 、( $n=1,2,3,4$ )は、事前の知識、および観測を通じて推定できるものである。2次元動き場  $u, v$  は、λを事前に与えた後、各  $\sigma$  を評価し、輝度値  $I, J$  を観測しながら上記評価関数を最小にすることにより得られる。

この最小化問題は式 (22) を  $u, v$  について偏微分して零とする連立方程式を解けばよい。これは緩和法を用いて求めることができる。即ち、第  $n$  回目のステップにおける画素  $(i, j, t)$  の動き場の  $x$  成分  $u_{i,j}^{(n)}$ 、 $y$  成分  $v_{i,j}^{(n)}$  は、

$$u_{i,j}^{(n)} = \frac{I_x f_x / \sigma_{r1}^2 + J_x g_x / \sigma_{r2}^2 - \lambda U_{i,j} / \sigma_q^2}{4\lambda / \sigma_q^2 + (I_x)^2 / \sigma_{r1}^2 + (J_x)^2 / \sigma_{r2}^2} \quad (23)$$

$$v_{i,j}^{(n)} = \frac{I_y f_y / \sigma_{r1}^2 + J_y g_y / \sigma_{r2}^2 - \lambda V_{i,j} / \sigma_w^2}{4\lambda / \sigma_w^2 + (I_y)^2 / \sigma_{r1}^2 + (J_y)^2 / \sigma_{r2}^2} \quad (24)$$

を繰り返し用いることにより  $u_{i,j}^{(n)}, v_{i,j}^{(n)}$  が最終解に近づいていき、推定値が得られる。但し、ここで、

$$\begin{aligned}
f_x &= \frac{\partial f}{\partial x} = I_y v_{i,j}^{(n-1)} + I_t \\
f_y &= \frac{\partial f}{\partial y} = I_x u_{i,j}^{(n-1)} + I_t \\
g_x &= \frac{\partial g}{\partial x} = J_y v_{i,j}^{(n-1)} + J_t \\
g_y &= \frac{\partial g}{\partial y} = J_x u_{i,j}^{(n-1)} + J_t \\
U_{i,j}^{(n-1)} &= u_{i-1,j}^{(n-1)} + u_{i+1,j}^{(n-1)} + u_{i,j-1}^{(n-1)} + u_{i,j+1}^{(n-1)} \\
V_{i,j}^{(n-1)} &= v_{i-1,j}^{(n-1)} + v_{i+1,j}^{(n-1)} + v_{i,j-1}^{(n-1)} + v_{i,j+1}^{(n-1)}, \quad (25)
\end{aligned}$$

である。

## 5 シミュレーション

### 5.1 CG 画像に対する動き場推定結果

正則化の拡張定式化に基づく動き場推定アルゴリズムの効果を確かめるため図1に示す回転球画像シーケンスについてシミュレーションをおこなった。画像のパラメタ

はサイズが、256×256画素、量子化が8ビットである。垂直軸まわりに回転する球を斜め上から見たものである。球の表面には反射率の違いによりテクスチャを生じている。図2にシミュレーション画像に与えた動きを示す。この回転球画像にローパスフィルタを施した画像を図3に示す。また、微分処理を行ったあとローパスフィルタを施した画像を図4に示す。この二つの画像に基づいて評価式 (22) を最小化して、動き場を求めた。λは、想定される被写体の形状の複雑さなどから定めた。また、観測雑音の分散は、輝度の勾配、勾配の時間変化に基づき画素毎に決定している。実験では、比較のため、従来手法である式 (15) に基づく手法についても実験を行った。

実験結果の一例を図5に示す。図5(a)はローパスフィルタを施した画像に対して、従来手法、即ち、式 (15) を適用して得た動き場である。図5(b)は、式 (22) のうち  $J$  に関する項を除いた評価式を最小化することにより得られた動き場である。これは、単一のセンサ出力、即

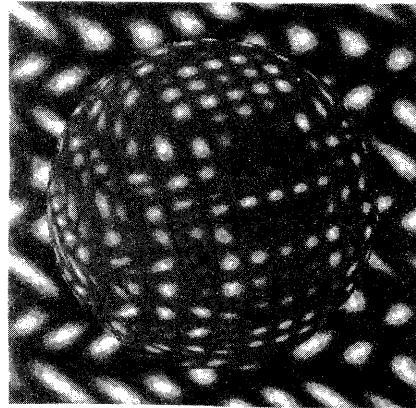


図1 シミュレーションに用いたCG画像。  
右から左へ回転している。

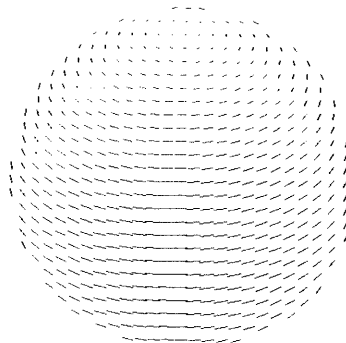


図2 シミュレーション画像における真の動き場。

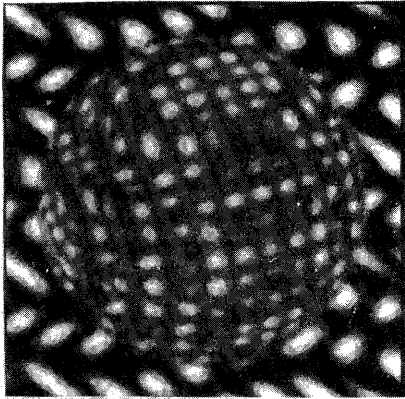


図3 ローパスフィルタ (センサ1) 出力画像。

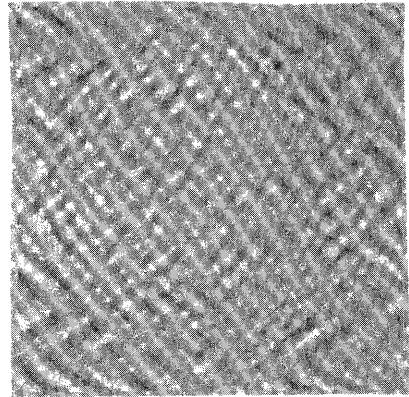
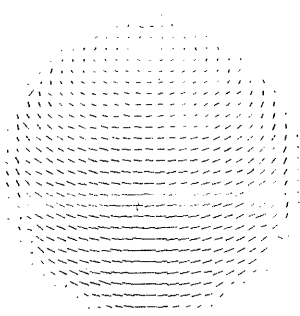
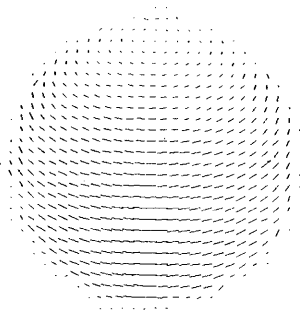


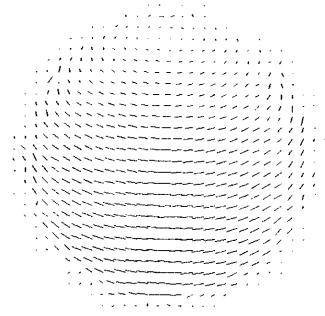
図4 微分フィルタ (センサ2) 出力画像。



(a) 従来方法。式 (15) に基づく。

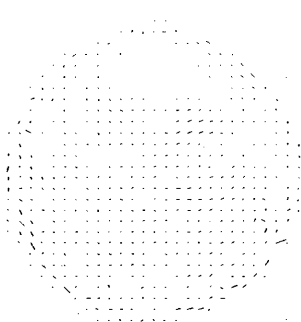


(b) 提案手法。単一画像を用い、式 (22) に基づく。

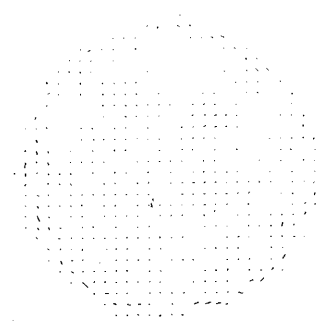


(c) 提案手法。複数画像を用い、式 (22) に基づく。

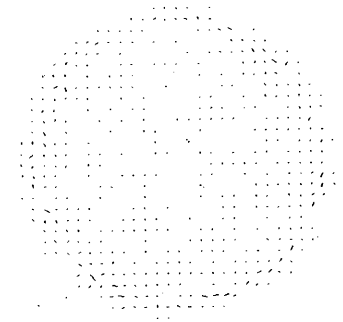
図5 シミュレーション結果。推定された動き場。



(a) 従来方法。式 (15) に基づく。



(b) 提案手法。単一画像を用い、式 (22) に基づく。



(c) 提案手法。複数画像を用い、式 (22) に基づく。

図6 真の動き場との誤差のパターン。

ち、ローパスフィルタを施した画像に対し、画素毎に変化する観測雑音により重みづけを変化させている。図5(c)は、式(22)をそのまま最小化するもので、2種のセンサ出力、即ち、ローパスフィルタを施した画像、および微分処理を行ったあとにローパスフィルタを施した画像に対して拡張定式化を適用している。また、図6(a),(b),(c)に推定されたフローと真のフローの誤差ベクトルのボタンを示す。図6(a)より従来法では、輝度の平坦部、極値周辺部の推定誤差が大きくなっていることがわかる。一方、図6(b),(c)から提案手法が従来法に比べて精度が向上していることが確かめられた。図7に実際のフローの値と推定値との誤差を示す。図より提案手法では、画素毎に推定観測雑音に基づいて重みを変えることにより、従来手法に比べて、動きの推定精度が約30%向上し、さらに複数の画像を統合することにより精度がさらに約40%向上した。

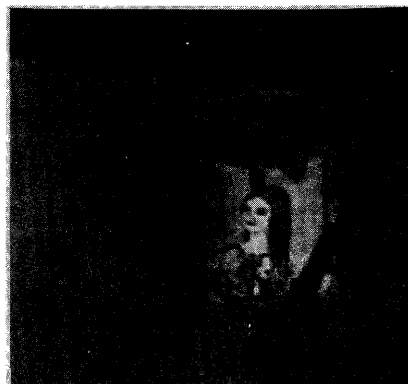


図8 シミュレーションに用いた実画像。回転するシリンダ。

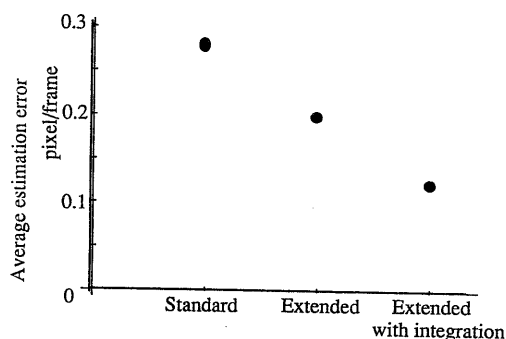


図7 シミュレーション結果。真の動きと推定値との誤差。

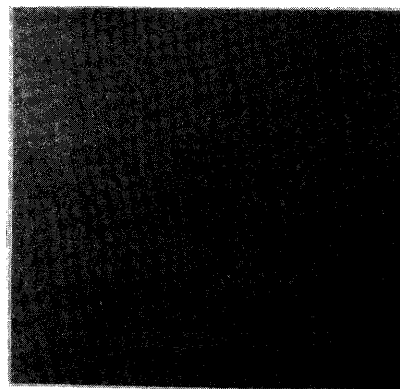
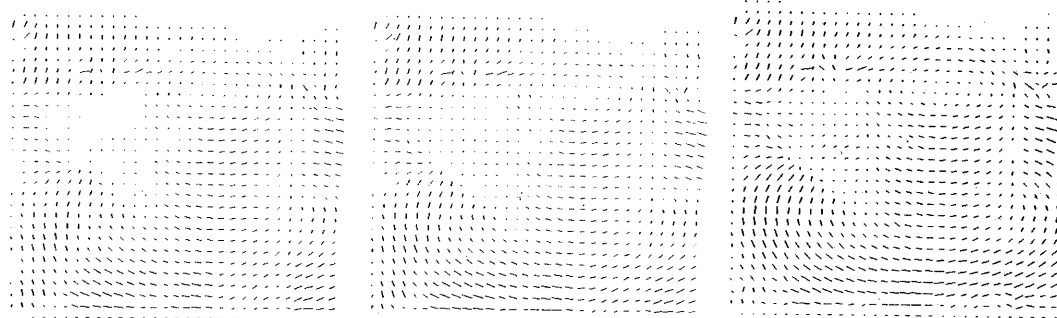


図9 微分フィルタ出力画像。



(a) 従来方法。式(15)に基づく。

(b) 提案手法。単一画像を用い、式(22)に基づく。

(c) 提案手法。複数画像を用い、式(22)に基づく。

図10 シミュレーション結果。推定された動き場。

## 5.2 リアル画像に対する結果

最後に、テレビカメラにより撮影したリアル画像に対して動き場推定を行った結果を示す。図8に示したような被写体をターンテーブルに載せて回転させたものである。光源を斜め上に配置した。光源に不変な動き場推定アルゴリズム [7] に拡張正則化を組み合わせる実験を行った。図9に微分処理後ローパスフィルタをフィルタを施した画像を示す。図10に推定結果を示す。図10(a)の従来法では回転しているシリンダの部分の動き場の誤差が大きいことが認められた。図10(b)はローパスフィルタを施した画像に対し、画素毎に変化する観測雑音により重みづけを変化させ得られた動き場、図10(c)は2種類のフィルタを施した画像に対して拡張定式化を適用して得られた動き場である。以上により、実画像に対しても提案手法の効果が確かめられた。なお、回転台の部分の動き場は、どの手法においても推定誤差が大きくなっているが、これは、勾配法が動きの小さいところでのみ適用できることを示すものである。

## 6 まとめ

従来の標準正則化における定式化を拡張し、観測時の雑音の評価、複数センサの信頼度の差などを最適化の定式化に取り入れる手法を提案し、これがビジョンの種々の問題に適用できることを述べた。この手法では、センサの観測雑音が観測する対象に依存して画素毎に異なることを利用しているもので、従来の観測雑音の分散を一定とする手法に比べ、より精度が高い推定が可能となることを示した。また、提案した定式化は、複数センサデータの統合を行うために必須の手法であることを述べた。

シミュレーション実験により、動き場推定における複数センサデータ統合の簡単な例について提案手法の効果を確かめた。そこでは、ローパスフィルタ画像とハイパスフィルタ画像の統合を取り上げ、標準正則化理論の従来の定式化を適用した場合に比較して効果が大きいことを確かめた。

今回は、拡張定式化において、簡単のために観測雑音間の相関を無視している。また、観測雑音の分散推定は簡単な予備実験を通じて設定している。実際のセンサデー

タの雑音は相関があり、この項まで考慮した最適化は今後の課題としたい。さらには、観測雑音の共分散行列の自動推定へと検討を進める予定である。

謝辞 研究の機会を与えていただいた釜江尚彦ヒューマンインタフェース研究所長、高野陸男知能ロボット研究部長、日頃御討論頂く金子博主幹研究員、尺長健主幹研究員を始め研究グループの皆様感謝する。

## 参考文献

- [1] 杉原、金谷、“数理的手法による3次元情報抽出”，人工知能学会誌，4,1, pp.13-20 (1989)
- [2] 松山隆司、“画像理解における推論方式”，人工知能学会誌，4,1, pp.21-29 (1989)
- [3] T. Poggio, V. Torre and C. Koch, “Computational vision and regularization theory”, *Nature* 317,26 (Sep. 1985)
- [4] A.N. Tikhonov, “The regularization of ill-posed problems,” *Dokl. Akad. Nau. SSR*, 153, 1, pp.49-52 (1963)
- [5] B. Horn and B. Schunck, “Determining optical flow”, *Artificial Intelligence* 17, pp.185-203 (1981)
- [6] E.C. Hildreth, “The measurement of visual motion”, MIT press, 1983.
- [7] N. Mukawa, “Estimation of shape, reflection coefficient and illuminant direction from image sequence”, *Proc. 3rd Int. Conf. Computer Vision*, pp.507-512, Osaka (Dec. 1990)
- [8] D. Terzopoulos, “Regularization of inverse visual problems involving discontinuities”, *IEEE Trans. on Pattern Anal. Machine Intell.*, PAMI-8,4, pp.413-424 (July 1986)
- [9] A. Tarantola, “Inverse problem theory”, Elsevier Science Publisher (1987)