数値的安定性を保証したシーンフロー推定法

A Scene-Flow Estimation Method with Guaranteed Numerical Stability

正会員 亀田 裕介[†], 正会員 松田 一朗[†], 正会員 伊東 晋[†]

Yusuke Kameda[†], Ichiro Matsuda[†] and Susumu Itoh[†]

あらまし 映像上の動きの算出にはオプティカルフロー推定が用いられているが、奥行き方向の速度は分からない。シーンフローは、ステレオ法とオプティカルフロー推定の拡張版から算出するステレオ映像上のベクトル場であり、奥行き方向の速度を復元できる。シーンフローにより被写体とカメラの3次元的な運動が推定でき、障害物検出や自己位置推定などに利用できる。従来のシーンフロー推定法では、シーンフローの正則条件による過平滑を防ぐために非線型方程式の数値計算が必要であった。さらに、その数値的安定性が画像と正則条件の係数に依存していたため、適切な係数の探索も不可能であった。本稿では、従来のアルゴリズムをまとめ直して単純化し、数値的安定条件の導出により、任意の計算条件で計算できる手法を提案する。これにより、適切な正則条件の係数の探索が可能になる。さらに、誤差解析により、提案法の数値解の精度向上を実証する。

キーワード：シーンフロー、変分法、ステレオ動画像、オプティカルフロー、視差、数値的安定性

1. ま え が き

映像上の動きの算出は、オプティカルフロー推定とよばれる分野で1980年代から現在に至るまで盛んに研究されている¹⁾²⁾。オプティカルフローからは、被写体の動きやおおよその自己運動を推定できる。しかし、オプティカルフローは単視点映像上の見かけの動きであり、奥行き方向の速度成分を持たないため、被写体の3次元的な運動は復元できない。一般的な状況下において、オプティカルフローだけから3次元空間的な運動を復元することは困難である。したがって、3次元的な運動を求めるためには、奥行きを推定する手法と組み合わせる必要がある。

画像上の奥行きの推定法は、レーザや赤外線パターン照射によるレンジセンサを用いる手法と、ステレオカメラを用いるステレオ法に大別される。前者は、屋内や静止物体などの好条件下では奥行き推定精度が良い一方、鏡面反射や透過、運動には不向きであり、通常のカメラより高価である。ステレオ法は、カメラ幾何学において古くから研究されており、三角測量の原理からステレオ画像上の視差とカメラ間距離から奥行きを推定する。近年、車載カメラではステレオカメラがよく用いられている。

オプティカルフロー推定の拡張による3次元運動推定として、カラー画像も撮影可能なレンジセンサに基づくRGB-Dフロー推定³⁾と、ステレオ法に基づくシーンフロー推定⁴⁾⁵⁾

の2種類がある。1990年代頃より、ステレオ法にオプティカルフロー推定法を拡張して組み合わせることで被写体の3次元的な運動を復元する研究が行われている⁴⁾。これにより生まれた分野がシーンフロー推定である。シーンフローは、画像上の視差の時間変化とオプティカルフローから成るベクトル場である⁴⁾⁵⁾。ステレオカメラ間距離を用いることで、シーンフローを被写体表面の3次元的な運動へと幾何学的に変換できる。この変換された3次元空間中の運動は3Dシーンフローとよばれる。

シーンフローによって映像上の被写体の奥行きや3次元的な運動を復元することで、障害物検出や行動分析などへの応用が可能となる。また、シーンフローからカメラの自己運動が3次元的に分かるので、ロボットや車両、航空機など自走する機械の自己位置推定にも応用が期待される。実際に、自動走行車や運転支援などの車産業や交通業界へのシーンフローの応用が始まりつつある⁵⁾⁶⁾。ステレオ映像処理では、単視点映像に比べて記憶容量と演算量がかかるが、2010年前後からの汎用並列演算器(GPGPU)などによる並列計算の高性能化により実時間処理が可能になってきている。このような背景のもと、ステレオ映像から推定されるシーンフローが注目されている。

シーンフローの算出には、オプティカルフロー推定と同様に、変分法に基づく手法が広く用いられている。シーンフローの拘束条件として、フレーム間およびステレオ間の対応する画素値の不変性を用いる。この条件だけでは、等輝度領域で一意に解が決まらないので、正則条件を付加する。正則化における過平滑を抑制するため、全変動(total variation)法⁷⁾に基づく正則条件がよく用いられている。

2013年8月および12月、年次大会および冬季大会で発表

2013年12月2日受付, 2014年4月18日再受付, 2014年5月27日採録

[†] 東京理科大学 理工学部 電気電子情報工学科

(〒278-8510 千葉県野田市山崎2641, TEL 04-7124-1501 (ext. 3722))

この全変動法の正則条件は、フローベクトルに応じて正則化の強さを変化させる手法であり、フロー駆動とよばれる。この他、Nagel 法を始め、画像に応じて正則化の強さを制御する画像駆動の正則条件もある⁷⁾。最終的に、これらの条件から成るエネルギー最小化問題を解く偏微分方程式を数値計算することで、求める解を算出する。

従来のシーンフロー推定法⁵⁾の数値計算には課題がいくつかある。まず、全変動法の正則条件では非線形で演算量の大きい処理を行う必要がある。この点について、精度を保ちつつ数値計算のアルゴリズムをできる限り単純にすることが望ましい。また、数値計算の数値的安定性が十分に議論されていないため、正則条件の係数と画像に依存して数値計算が収束しない場合がある。このため実用面で不安がある以前に、正則条件の係数の適切な値の探索すら困難であることが挙げられる。これらの数値計算上の問題を解決することで、シーンフロー推定法のハードウェア化にも道が開けると考えられる。

そこで本稿では、複雑で非線型であった従来法をまとめ直し、安定性解析を行うことで、数値的安定性が正則条件の係数と画像列に依存しないシーンフロー推定法を提案する。また、本研究で提案するアルゴリズムが、フロー駆動や画像駆動の正則条件と同様に、過平滑を抑制できることを示す。本研究により、アルゴリズムが単純化され数値計算が簡単になり、さらに任意の計算条件で数値計算の安定性が保証される。

2. シーンフローを求める変分問題

ステレオカメラの仮定として、平行化された同一仕様のカメラで同期撮影が可能であるとする。すなわち、左右のカメラの視差には縦方向の成分は存在しない。また、各時刻における視差は推定済みであるとする。この仮定のもと、輝度の不変性から成立する制約式とシーンフローの正則条件から、シーンフロー推定のためのエネルギー最小化問題について述べる。最小化問題の大まかな枠組みは先行研究⁵⁾に準ずる。

2.1 ステレオの動き制約式

$L(\mathbf{x}, t)$ を左カメラ画像の時刻 t の座標 $\mathbf{x} = (x, y)^\top \in \Omega \subset \mathbb{R}^2$ での画素値とする。ここで Ω は画像の領域である。このとき、微小時間変化 Δt に対して被写体の画素値が不変であれば、各画素の見かけの速度（フローベクトル） $\mathbf{u}(\mathbf{x}, t) = (u(\mathbf{x}, t), v(\mathbf{x}, t))^\top$ を用いて、

$$L(\mathbf{x}, t) = L(\mathbf{x} + \mathbf{u}(\mathbf{x}, t)\Delta t, t + \Delta t) \quad (1)$$

と書ける。この式はオプティカルフロー制約式とよばれる¹⁾。この式の画素対応の概念図を図 1 (a) に示す。

続いて、右カメラの画素値 $R(\mathbf{x}, t)$ について、視差 $\mathbf{d}(\mathbf{x}, t) = (d(\mathbf{x}, t), 0)^\top$ が存在すれば、左カメラの見かけの速度 $\mathbf{u}(\mathbf{x}, t)$ と視差の時間変化 $\mathbf{p}(\mathbf{x}, t) = (p(\mathbf{x}, t), 0)^\top$ を用いて、

(a) 左画像列上の画素の動き

(b) 右画像列上の画素の動き

図 1: 左右の画像列の画素対応の概念図。網掛け四角が左下方向に移動しつつ接近するとき、視差は時間的に大きくなる。

$$\begin{aligned} & R(\mathbf{x} + \mathbf{d}(\mathbf{x}, t), t) \\ &= R(\mathbf{x} + \mathbf{d}(\mathbf{x}, t) + (\mathbf{u}(\mathbf{x}, t) + \mathbf{p}(\mathbf{x}, t))\Delta t, t + \Delta t) \quad (2) \end{aligned}$$

と書ける。この式の画素対応の概念図を図 1 (b) に示す。

さらに、時刻 $t + \Delta t$ での左右のカメラの対応画素の画素値が等しいとすれば、

$$\begin{aligned} & L(\mathbf{x} + \mathbf{u}(\mathbf{x}, t)\Delta t, t + \Delta t) \\ &= R(\mathbf{x} + \mathbf{d}(\mathbf{x}, t) + (\mathbf{u}(\mathbf{x}, t) + \mathbf{p}(\mathbf{x}, t))\Delta t, t + \Delta t) \quad (3) \end{aligned}$$

が成り立つ。本稿では、簡単のため $\Delta t = 1$ とする。

シーンフローは、左カメラの見かけの速度 $\mathbf{u}$ と視差の時間変化 $\mathbf{p}$ の組であり、これ以降 $\mathbf{v} = (\mathbf{u}^\top, \mathbf{p}^\top)^\top = (u, v, p)^\top$ で表す。

2.2 変分法によるエネルギー最小化問題

シーンフローと画像は時空間 2 階微分可能とし、式 (1), (2), (3) それぞれの両辺の二乗誤差を $E_{LF}^2(\mathbf{v}, \mathbf{x}, t)$, $E_{RF}^2(\mathbf{v}, \mathbf{x}, t)$, $E_{DF}^2(\mathbf{v}, \mathbf{x}, t)$ とする。これらの制約式から

$$E_D(\mathbf{v}, \mathbf{x}, t) = E_{LF}^2 + c(\mathbf{x}, t) (E_{RF}^2 + E_{DF}^2) \quad (4)$$

をデータ項とする。ここでは簡単のため、 $(\mathbf{v}, \mathbf{x}, t)$ を略記している。また、 $c(\mathbf{x}, t)$ を視差 $d(\mathbf{x}, t)$ が存在すれば 1、オクルージョン等で視差が未定義であれば 0 とすることで、右カメラの画素対応を利用するかどうかを選択する。

式 (4) だけでは等輝度領域で $\mathbf{v}$ が一意に定まらないため、 $\mathbf{v}$ の正則化項 $E_S(\mathbf{v}, \mathbf{x}, t)$ を付加して、エネルギー関数を

$$J(\mathbf{v}, t) = \int_{\Omega} (E_D(\mathbf{v}, \mathbf{x}, t) + E_S(\mathbf{v}, \mathbf{x}, t)) d\mathbf{x} \quad (5)$$

と定義する⁵⁾。本稿での正則化項は、正の実数 λ と γ を正則条件の係数、ナブラ演算子 $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)^\top$ として、

$$\begin{aligned} E_S(\mathbf{v}, \mathbf{x}, t) &= \lambda (|\nabla u|^2 + |\nabla v|^2) + \gamma |\nabla p|^2 \\ &= \lambda (u_x^2 + u_y^2 + v_x^2 + v_y^2) + \gamma (p_x^2 + p_y^2) \quad (6) \end{aligned}$$

とする。ここでは、 $\mathbf{v}, L$ については $(\mathbf{x}, t)$ を省略し、右下添字の x, y は空間偏微分を表しており、以降同様である。これ以降、 R についても $(\mathbf{x} + \mathbf{d}, t)$ は省略する。

式 (5) の最小化問題の解は、各画素 $(\mathbf{x}, t)$ に対して成立するオイラー・ラグランジュ偏微分方程式系

$$\lambda \nabla^2 u - L_x E_{LF} - c R_x E_{RF} = 0 \quad (7)$$

Algorithm 1: 多重解像度ワーブ法によるシーンフロー推定アル

ゴリズム

Data: 左画像列 L , 右画像列 R , 視差画像列 d , 正則条件の係数 λ, γ , ピラミッド変換レベル数 $l_{\max}$, ワーブ回数 $w_{\max}$, SOR 反復回数 $k_{\max}, m_{\max}$

Result: シーンフロー v

```

1 for  $l := l_{\max}$  to 0 do
2   ガウシアンピラミッド画像列  $L^{(l)}, R^{(l)}$  とピラミッド視差画像列  $d^{(l)}$  作成;
3    $v^{(l,0)} = v^{(l)} = \{0\} \forall x$ ;
4   for  $w := 1$  to  $w_{\max}$  do
5      $d^{(l)}$  と  $v^{(l,w-1)}$  より, ワーブ画像列  $R^{(l,w)}$  作成;
6     for  $k := 1, m := 1$  to  $k_{\max}, m_{\max}$  do
7        $L^{(l)}$  と  $R^{(l,w)}$  より  $v^{(l,w)}$  を求める漸化式の計算;
8      $v^{(l)} := v^{(l)} + v^{(l,w)}$ 
9    $v^{(l)}$  をアップサンプリングして  $v^{(l-1)}$  作成.
10  $v := v^{(l)}, l = 0$ ;
```

$$\lambda \nabla^2 v - L_y E_{LF} - c R_y E_{RF} = 0 \quad (8)$$

$$\gamma \nabla^2 p - c R_x (E_{DF} + E_{RF}) = 0 \quad (9)$$

の解である⁵⁾. ここで, $\nabla^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$ はラプラシアン, 右下添字 x, y 付きの L と R は空間偏微分値を表す.

2.3 ピラミッド変換と画像ワーブ法

先行研究⁵⁾⁸⁾では, オイラー・ラグランジュ方程式系を解く際に, 非線型な式 (4) に対してテイラー展開の 1 次近似と不動点法を用い, 反復回数毎の v の微小変化に注目した複雑な式を展開する必要がある. それに対し, 本研究では先行研究と同様の理論を単純化する. そのために, 画像のピラミッド変換とワーブ法による画像の変形の方に注目し, 変形した画像間のシーンフローを求めるオイラー・ラグランジュ方程式が単純な式になることを示す.

シーンフロー推定のアルゴリズムでは, まず, 左右の画像列をガウシアンピラミッド変換により, 低解像度画像列を作成する. これにより, シーンフロー v は解像度低下と同じ比率で小さくなる. したがって, これ以降, v が微小であると仮定する. 各解像度において, 算出したシーンフローを用いて, 画像ワーブ法により $t+1$ フレームを t フレームに近づけることを繰り返す. ワーブした画像間でのシーンフロー推定の繰り返しにより, 計算精度を上げていくことができる. ワーブ法をある程度反復した後, シーンフローをアップサンプリングし, より高い解像度で反復計算し, 最終的に求めるシーンフローを得る. このアルゴリズムの概要をアルゴリズム 1 に示す.

2.4 オイラー・ラグランジュ方程式

小さな v に対して, 式 (4) のテイラー展開の 1 次近似

$$E_D \simeq \mathbf{u}^\top \nabla L + L_t + c (\mathbf{u}^\top \nabla R + R_x p + R_t) + c (\mathbf{u}^\top \nabla D + R_x p + D_t) \quad (10)$$

を用いてエネルギー関数を構築する (付録参照). ここで, L_t と R_t は時間偏微分値であり, 前後フレーム間差分

$$L_t(\mathbf{x}) = L(\mathbf{x}, t+1) - L(\mathbf{x}, t) \quad (11)$$

$$R_t(\mathbf{x}) = R(\mathbf{x} + \mathbf{d}, t+1) - R(\mathbf{x} + \mathbf{d}, t)$$

で近似し, $D = R - L$, D_t は時刻 $t+1$ のときのステレオ間差分

$$D_t(\mathbf{x}) = R(\mathbf{x} + \mathbf{d}, t+1) - L(\mathbf{x}, t+1) \quad (12)$$

とする.

画像の局所構造テンソル

$$\mathbf{S}_0(f) = \nabla f (\nabla f)^\top \quad (13)$$

を用いて 2 次の半正定値実対称行列

$$\mathbf{S}_2 = \mathbf{S}_0(L) + c (\mathbf{S}_0(R) + \mathbf{S}_0(D)) \quad (14)$$

とした対称行列

$$\mathbf{S} = \left(\begin{array}{c|c} \mathbf{S}_2 & \mathbf{s} \\ \hline \mathbf{s}^\top & 2cR_x^2 \end{array} \right), \quad (15)$$

$$\mathbf{s} = 2cR_x \nabla (R + D) \quad (16)$$

と, 定数項

$$\mathbf{b}_1 = L_t \nabla L + c (R_t \nabla R + D_t \nabla D) \quad (17)$$

$$\mathbf{b}_2 = cR_x (R_t + D_t) \quad (18)$$

を用いて, オイラー・ラグランジュ方程式系は

$$\nabla^2 \begin{pmatrix} \lambda \mathbf{u} \\ \gamma p \end{pmatrix} - \mathbf{S} \begin{pmatrix} \mathbf{u} \\ p \end{pmatrix} - \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \mathbf{0} \quad (19)$$

と簡単に書ける (付録参照).

3. 数 値 計 算

式 (19) から各解像度での各ワーブ法のための v を反復計算で求め, その反復計算の安定性について導出する.

3.1 反復計算と安定性

半陰解法⁹⁾と逐次過緩和法 (SOR 法) を用いて, 式 (19) の $\mathbf{u}$ についての漸化式を

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \frac{\omega_1}{4} \tilde{\nabla}^2 \mathbf{u}^{(k)} - \frac{\omega_1}{4\lambda} (\mathbf{S}_2 \mathbf{u}^{(k+1)} + p\mathbf{s} + \mathbf{b}_1) \quad (20)$$

とする. ここで, 右肩 (k) は $\mathbf{u}$ の反復回数, $\omega_1 \in (0, 2)$ は加速係数, $\tilde{\nabla}^2$ は空間刻み幅 1 の 2 階中心差分近似によるラプラシアンとする*. $\mathbf{u}$ の反復の際, p は定数とする.

式 (20) を整理すると,

$$\mathbf{u}^{(k+1)} = \mathbf{A} \left(\left(1 + \frac{\omega_1}{4} \tilde{\nabla}^2 \right) \mathbf{u}^{(k)} - \frac{\omega_1}{4\lambda} (p\mathbf{s} + \mathbf{b}_1) \right) \quad (21)$$

$$\mathbf{A} = \left(\mathbf{I}_2 + \frac{\omega_1}{4\lambda} \mathbf{S}_2 \right)^{-1} \quad (22)$$

となる. ここで, $\mathbf{I}_2$ は 2 次の単位行列である. $\mathbf{S}_2$ が半正定値対称行列であるので, $\omega_1 > 0$ と $\lambda > 0$ に依存せずに

* 係数行列の対角成分は 4 となるので式 (20) の第 2 項以降は 4 で割る.

$\mathbf{A}$ の固有値は 1 以下となる。したがって、式 (21) の安定条件は、 $1 + \frac{\omega_1}{4} \tilde{\nabla}^2$ から成る係数行列のスペクトル半径が 1 以下となることである。この係数行列のスペクトル半径は、 $|1 + \frac{\omega_1}{4}(-4 - 4)| = |1 - 2\omega_1|$ となるので、 $\omega_1 \leq 1$ であれば $\mathbf{u}$ は安定である。

同様に、 p についての漸化式を

$$\begin{aligned} & p^{(m+1)} \\ &= p^{(m)} + \frac{\omega_2}{4} \tilde{\nabla}^2 p^{(m)} - \frac{\omega_2}{4\gamma} \left(\mathbf{s}^\top \mathbf{u} + 2cR_x^2 p^{(m+1)} + b_2 \right) \end{aligned} \quad (23)$$

とする。ここで、 $\mathbf{u}$ と同様に、右肩 (m) は p の反復回数、 $\omega_2 \in (0, 2)$ は加速係数である。 p の反復の際、 $\mathbf{u}$ は定数とする。整理すると、

$$\begin{aligned} & p^{(m+1)} \\ &= \frac{1}{1 + \frac{c\omega_2 R_x^2}{2\gamma}} \left(\left(1 + \frac{\omega_2}{4} \tilde{\nabla}^2 \right) p^{(m)} - \frac{\omega_2}{4\gamma} (\mathbf{s}^\top \mathbf{u} + b_2) \right) \end{aligned} \quad (24)$$

となる。 $\mathbf{u}$ と同様に、 $\omega_2 \leq 1$ であれば p は安定である。

提案法の半陰解法により、 $\mathbf{S}_2$ を含む係数行列の逆行列が拡散項にかかることで、過平滑を抑制できる。したがって、 $\mathbf{S}_2$ を含む項は画像駆動正則化の性質を持つともいえる。

画像の各画素での空間偏微分値は前後のフレームでの空間偏微分値の平均とする。空間 1 階偏微分値は Scharr 作用素で近似する¹⁰⁾。時間差分と視点間差分については、標準偏差 0.8 のガウス関数によりぼかしてから行う。計算の高速化のため、 $\mathbf{u}^{(k)}$ と $p^{(m)}$ それぞれの反復計算には over-relaxed red-black Gauss-Seidel 法¹¹⁾ を用いて並列化する。

4. 実験と考察

4.1 安定な数値計算

計算条件は、比較のため先行研究⁵⁾ で提示された値と同一としており、必ずしも提案法で最も良い性能を示す値ではない。すなわち、正則条件の係数 $\lambda = 0.06$ 、 $\gamma = 0.6$ 、画像ワープ法の反復回数 $w_{\max} = 2$ 回、ピラミッドレベル $l_{\max} = 4$ 回とする。 $\mathbf{v}$ の初期条件はゼロ速度場、境界条件は斉次ノイマン条件である。対象画像列は、直進、カーブ、対向車、坂などの様々なシーンを含む車載カメラ画像列 EISATS set 2 sequences 1 と 2（それぞれ全 100 フレーム、396 フレーム）とする¹²⁾。

一回の反復での数値解の修正量（修正ベクトルのノルム）の平均値を

$$\frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} \sqrt{|\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)}|^2 + |p^{(m+1)} - p^{(m)}|^2} \quad (25)$$

と定義し、ここで $k = m$ とする。EISATS set 2 sequence 2 画像列の 395~396 フレーム間のシーンフロー

(a) $l = 4, w = 1$ のときの反復計算の平均修正量

(b) アルゴリズム 1 全体での反復計算の平均修正量

図 2: sequence 2 set 2 の 395~396 フレーム間の SOR 法 ($k_{\max} = m_{\max} = 10^3$ 回) の反復計算の平均修正量。(a) は (b) の横軸の最初の一目盛までと同一である。奇数目盛では画像ワープが、偶数目盛ではピラミッド変換が行われている。

推定において、安定性を検証するために SOR 法の反復回数 $k_{\max} = m_{\max} = 10^3$ 回としたときの、式 (25) の測定結果を図 2 に示す。図 2(a) より、修正量は 0 に収束していることが分かる。また、アルゴリズム全体の画像ワープ法 (2 回) とピラミッド変換 (4 回) のループを通しての式 (25) の測定結果を図 2(b) に示す。図 2(b) の横軸の最初の一目盛り分が図 2(a) であり、横軸の奇数番目の目盛では画像ワープが、偶数番目の目盛ではピラミッド変換が行われている。図 2(b) では、画像ワープやピラミッド変換の直後に修正量が大きくなるが、これはアルゴリズム 1 の 7 行目で新たに $\mathbf{v}$ の反復計算が開始されるためであり、その後は修正量が 0 に収束していくことがわかる。図 2(b) より、画像ワープ法とピラミッド変換の後の再度の反復計算も収束することがわかる。この傾向は、他のシーンの画像ペアでも同様である。これらの正則条件の係数と画像列に対しては、加速係数 ω_1 と ω_2 が 2 以上になると反復計算は発散した。これは、 $\mathbf{A}$ のスペクトル半径は画像と正則条件の係数に応じて 1 より小さくなり、 $|1 - 2\omega_1| \leq 1$ の条件が緩和されるためであると考えられる。以上より、理論上は加速係数が 1 以下であれば反復計算は必ず安定であり、かつ SOR 法により計算を加速させることも可能である。

図 3: EISATS set 2 sequence 1 の 11 フレーム目のシーンフロー．前方に走行中，先行車は遠ざかる画像列である．

図 4: EISATS set 2 sequence 2 の 220 フレーム目のシーンフロー．左側の対向車が右折してくる．

図 5: $u(x, t)$ の方向と大きさ

表 1: KITTI データベースのシーンフローの真値（見かけの動き u ）と推定値との画素あたりの平均誤差．

画像列 番号	平均端点誤差		平均時空間角度誤差	
	提案法	従来法	提案法	従来法
026	15.4	19.9	0.637	1.53
082	20.7	25.2	0.742	1.25
112	29.6	38.9	0.459	1.41
170	6.51	7.94	0.779	1.35
179	8.62	11.0	0.636	1.26
180	1.90	7.29	0.422	0.806
193	7.35	12.0	0.436	0.924

上述の計算条件のもと，従来法と同じく $k_{\max} = m_{\max} = 3$ 回として，提案法で計算した結果を図 3 と図 4 に示す．フローの密な描画については図 5 の色表現²⁾で表す．視差の時間変化が正の領域は大きく接近している領域に対応している．なお，図 3 の先行車は，ガラスの透過によりステレオ対応できないため， p は正しく計算できない．

4.2 性能評価

推定法の性能評価では，推定値 $u(x, t)$ と提供された真値 $u^*(x, t)$ との間の画素あたりの誤差を測定する．本稿では，平均端点誤差²⁾

$$\frac{1}{|\Omega|} \sum_{x \in \Omega} |u - u^*| \quad (26)$$

と，平均時空間角度誤差²⁾

$$\frac{1}{|\Omega|} \sum_{x \in \Omega} \arccos \left(\frac{u^\top u^* + 1}{\sqrt{|u|^2 + 1} \sqrt{|u^*|^2 + 1}} \right) \quad (27)$$

とする．ただし，真値が存在しない画素は含めない．性能評価の対象画像列には，先述の EISATS の画像列に加え，KITTI データセット¹³⁾ の画像サイズ 1238×374 の 7 つの画像列（それぞれ全 2 フレーム）を用いた．KITTI ステレオ画像の視差の推定には OpenCV ライブラリ 2.4.8 の StereoSGBM 法¹⁴⁾を用い，その推定パラメータには自動設定値を用いた．従来法の計算パラメータは当該文献で提示された最適値とし，提案法に存在しない inner iteration が 15 回であること以外，提案法のパラメータと同一である．ただし，加速係数 $\omega = 1.99$ では計算が発散したので， $\omega = 1$ とした．提案法と従来法⁵⁾の性能評価結果を図 6 と表 1 に示す．結果より，提案法は従来法より全体として精度が向上していることがわかる．以上より，提案法は従来法に数値計算の安定性を付加し，さらに計算精度を上げることができた．また，誤差が最小となるように提案手法の正則条件の係数の値を調整することで，計算精度をさらに上げることができると考えられる．

図 6: EISATS set2 sequence 1, 2 のステレオ画像列の各フレームにおける、シーンフローの真値（見かけの動き u ）と推定値との画素あたりの平均誤差。

5. む す び

本稿では、シーンフローを求める数値計算法の安定性を理論的に解析し、安定性解析の理論が正しいことを実証した。本研究により、任意の画像と正則条件の係数の組合わせで安定に数値計算を行えるようになり、また多くの場合、従来法より数値解の精度が向上した。提案法により、誤差を最小化する正則条件の係数の値の最適化を行うことが可能になった。

【文 献】

- 1) B. K. Horn and B. G. Schunck: "Determining Optical Flow", *Artificial Intelligence*, **17**, pp. 185–203 (1981)
- 2) S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black and R. Szeliski: "A Database and Evaluation Methodology for Optical Flow", *International Journal of Computer Vision*, **92**, 1, pp. 1–31 (2011)
- 3) E. Herbst and D. Fox: "RGB-D Flow : Dense 3-D Motion Estimation Using Color and Depth", *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2276–2282 (2013)
- 4) I. Patras, E. A. Hendriks and G. Tziritas: "A Joint Motion / Disparity Estimation Method for the Construction of Stereo Interpolated Images in Stereoscopic Image Sequences", *International Conference on Pattern Recognition*, pp. 359–368 (1996)
- 5) A. Wedel, T. Brox, T. Vaudrey, C. Rabe, U. Franke and D. Cremers: "Stereoscopic Scene Flow Computation for 3D Motion Un-

- derstanding", *International Journal of Computer Vision*, **95**, 1, pp. 29–51 (2011)
- 6) S. Sivaraman and M. M. Trivedi: "Looking at Vehicles on the Road: A Survey of Vision-Based Vehicle Detection, Tracking, and Behavior Analysis", *IEEE Transactions on Intelligent Transportation Systems*, **14**, 4, pp. 1773 – 1795 (2013)
- 7) J. Weickert and C. Schnörr: "Variational Optic Flow Computation with a Spatio-Temporal Smoothness Constraint", *Journal of Mathematical Imaging and Vision*, **14**, 3, pp. 245–255 (2001)
- 8) T. Brox, A. Bruhn, N. Papenberger and J. Weickert: "High Accuracy Optical Flow Estimation Based on a Theory for Warping", *European Conference on Computer Vision (ECCV)*, Vol. 3024 of *Lecture Notes in Computer Science*, pp. 25–36 (2004)
- 9) Y. Kameda, A. Imiya and N. Ohnishi: "A Convergence Proof for the Horn-Schunck Optical-Flow Computation Scheme using Neighborhood Decomposition", *12th International Workshop on Combinatorial Image Analysis*, Vol. 4958 of *Lecture Notes in Computer Science*, pp. 262–273 (2008)
- 10) H. Scharr: "Optimal Operators in Digital Image Processing", *Doctoral thesis*, Rupertus Carola University of Heidelberg (2000)
- 11) Y. Saad: "Iterative Methods for Sparse Linear Systems: Second Edition", *Society for Industrial and Applied Mathematics* (2003)
- 12) T. Vaudrey, C. Rabe, R. Klette and J. Milburn: "Differences Between Stereo and Motion Behaviour on Synthetic and Real-World Stereo Sequences", *23rd International Conference Image and Vision Computing New Zealand, IEEE*, pp. 1–6 (2008)
- 13) A. Geiger, P. Lenz, C. Stiller and R. Urtasun: "Vision Meets Robotics: The KITTI Dataset", *The International Journal of Robotics Research*, **32**, 11, pp. 1231–1237 (2013)
- 14) H. Hirschmüller: "Stereo Processing by Semiglobal Matching and Mutual Information.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30**, 2, pp. 328–41 (2008)

〈付 録〉

A. 式 (10) 導出

以下の E_{DF} の導出により, E_{LF} と E_{RF} は自明である.

$$\begin{aligned} E_{DF}(\mathbf{x}) &= R(\mathbf{x} + \mathbf{d} + \mathbf{u} + \mathbf{p}, t+1) - L(\mathbf{x} + \mathbf{u}, t+1) \\ &\simeq R + \mathbf{u}^\top \nabla R + R_x p + R_t - (L + \mathbf{u}^\top \nabla L + L_t) \end{aligned} \quad (\text{付} \cdot 1)$$

式 (11) より,

$$\begin{aligned} E_{DF}(\mathbf{x}) &\simeq \mathbf{u}^\top \nabla R + R_x p + R(\mathbf{x} + \mathbf{d}, t+1) \\ &\quad - (\mathbf{u}^\top \nabla L + L(\mathbf{x}, t+1)) \\ &\simeq \mathbf{u}^\top \nabla D + D_t + R_x p \end{aligned} \quad (\text{付} \cdot 2)$$

B. オイラー・ラグランジュ方程式

式 (10) のときの式 (5) のオイラー・ラグランジュ方程式系は

$$\begin{aligned} \lambda \nabla^2 u - L_x (\mathbf{u}^\top \nabla L + L_t) - c R_x (\mathbf{u}^\top \nabla R + R_x p + R_t) \\ - c D_x (\mathbf{u}^\top \nabla D + R_x p + D_t) = 0 \end{aligned} \quad (\text{付} \cdot 3)$$

$$\begin{aligned} \lambda \nabla^2 v - L_y (\mathbf{u}^\top \nabla L + L_t) - c R_y (\mathbf{u}^\top \nabla R + R_x p + R_t) \\ - c D_y (\mathbf{u}^\top \nabla D + R_x p + D_t) = 0 \end{aligned} \quad (\text{付} \cdot 4)$$

$$\gamma \nabla^2 p - c R_x (\mathbf{u}^\top \nabla (R + D) + 2 R_x p + R_t + D_t) = 0 \quad (\text{付} \cdot 5)$$

である. まとめると式 (19) となる.

かめだ ゆうすけ
亀田 裕介 2006 年, 千葉大学工学部情報画像工学科卒業. 12 年, 同大学大学院融合科学研究科情報科学専攻博士後期課程修了. 09~11 年, 日本学術振興会特別研究員 DC2. 民間企業を経て 13 年, 東京理科大学理工学部電気電子情報工学科助教. コンピュータビジョンと映像符号化に関する研究に従事. 博士 (工学). 正会員.

まつだ いちろう
松田 一郎 1991 年, 東京理科大学理工学部電気工学科卒業. 1996 年, 同大学大学院博士後期課程修了. 同年より, 東京理科大学理工学部電気工学科に勤務し, 現在, 同大学理工学部電気電子情報工学科教授. 画像符号化等の研究に従事. 博士 (工学). 正会員.

いとう すずむ
伊東 晋 1976 年, 東京大学工学部電子工学科卒業. 1981 年, 同大学大学院電子工学専攻博士課程修了. 同年より, 東京理科大学理工学部電気工学科に勤務し, 現在, 同大学理工学部電気電子情報工学科教授. 画像処理, 画像符号化等の研究に従事. 工博. 正会員 (フェロー).