

Subspace 拘束を用いた三次元シーンフローの修正とモーション推定

水野 宏基[†] 藤吉 弘亘^{†a)} 岩堀 祐之[†]

A Method for Regularizing Three Dimensional Scene Flow by Subspace Constraints and Estimating Rigid Motion

Hiroki MIZUNO[†], Hironobu FUJIYOSHI^{†a)}, and Yuji IWAHORI[†]

あらまし 本論文では、ジェスチャ認識に応用が可能である技術として、三次元シーンフローを用いたモーション推定法を提案する。本手法は、画像中の二次元オプティカルフローを入力とすることで、三次元空間中の物体形状を復元することなく、三次元シーンフローを復元する。次に、三次元空間中の物体運動は短時間において剛体運動であると仮定することで、Subspace 拘束を用いて三次元シーンフローを修正し、最後に RANSAC を用いて剛体運動パラメータを推定する。シミュレーションと実画像を用いた実験より、推定したモーションパラメータを評価し、提案手法の有効性を示す。

キーワード 三次元シーンフロー、Subspace 拘束、修正、モーション推定

1. ま え が き

近年、ユビキタス社会におけるユーザビリティの向上として、人の自然動作に近いインタフェースであるジェスチャ認識の実現が期待されている。従来、人の運動を計測する手法として、モーションキャプチャが用いられている。モーションキャプチャは人間の関節などに特殊なマーカを装着し、それをカメラや磁気などのトラッカを用いて検出することで運動計測を行う。しかし、特殊なマーカを装着しなければならないという問題があり、非接触で人間の動作を計測する手法が期待されている。非接触で人間の動作を計測する手法として、カメラを用いた三次元シーンフローを用いることが考えられる。三次元シーンフローとは、三次元空間における物体の運動を三次元のベクトルで表現したものである。三次元シーンフローの復元には、あらかじめ三次元構造を復元する手法 [12] と、画像中の二次元オプティカルフローを用いる手法 [11] が提案されている。前者の手法では、三次元構造を復元するために、特殊な装置を必要とする。このような問題に対し、後者では、特殊な装置を必要とせず三次元シー

ンフローを復元することが可能であるというメリットがある。しかし、二次元オプティカルフローの精度が悪いと、高精度に三次元シーンフローを復元することが難しいという問題がある。そこで本研究では、複数のカメラ画像の二次元オプティカルフローから三次元シーンフローを高精度に復元し、そのシーンフローから物体のモーションを推定することを目的とする。

三次元空間中の物体の運動は短時間では剛体であると仮定した場合、その各点から発生するシーンフロー集合からなる「観測行列」には Subspace 拘束が存在し、そのランクは行列のサイズに対し非常に小さいものとなる。観測行列とは、 M フレーム間のシーンフロー集合 N 点からなる $3M \times N$ の行列である。しかし、ノイズを含んだシーンフロー集合から観測行列を作成した場合、そのランクは大きなものとなる。本研究では、この観測行列に対し特異値分解を用いたランク制限を施すことにより、シーンフローの修正を行い、精度を向上させる手法を提案する。

Subspace 拘束を適用するには、設定する問題において、観測行列がどのようなランクをもつのかを考える必要がある。観測行列は生成される運動モデルが簡単であるほど、そのランクを小さくすることができる。Tomasi らは Shape from motion の問題に対し、Subspace 拘束を用いることで高精度に物体の形状を復元する手法を提案している [4]。彼らの手法では、カ

[†] 中部大学大学院工学研究科情報工学専攻, 春日井市
Dept. of Computer Science, Chubu University, Kasugai-shi,
487-8501 Japan

a) E-mail: hf@cs.chubu.ac.jp

メラが平行投影であるという条件を用いることで、運動モデルを簡単化する。また、Irani は Subspace 拘束を用い、移動するカメラ映像において、短時間ではシーン構造は変化しないことを仮定することで、オプティカルフローを高精度に推定する手法を提案している [6]。彼女の手法では、実空間におけるカメラの運動を画像中で扱うため、各点の運動を表現するために多くのパラメータを必要とする。そのため、ある程度多くのフレームを観測行列にスタックしなければ、期待される精度を得ることは難しい。本手法では、実空間における物体の運動をそのまま扱うことができるため、より簡単な運動モデルで物体運動を表現することが可能となる。よって、観測行列のランクを小さくすることが可能であり、少ないフレームから精度の高いシーンフローを復元することができる。

次に、本論文では、復元した三次元シーンフローから、物体のモーションを推定する手法を提案する。復元された三次元シーンフローから、RANSAC を用いて剛体運動のパラメータを求めることで、物体のモーションを高精度に推定できることを示す。

2. 三次元シーンフローの復元

本研究では、複数のカメラ画像から得られる二次元オプティカルフローを入力とし、そこから得られる三次元シーンフローに対し修正を行うことで、精度を向上させる。三次元シーンフローの復元には、Vedulaらの提案する手法を用いる [11]。本章では、その復元手法について述べる。

シーンフロー復元は、二つのステップからなる。第1のステップは、シーンフローの存在位置の判定である。まず、三次元空間を一定間隔の格子 (Voxel) に分割し、その各点において、シーンフローが存在するかどうかを判定する。シーンと画像間の幾何的な関係から得られるオプティカルフローとシーンフローの制約を用いることで、三次元シーンフローの存在位置を判定する。第2のステップは、シーンフロー存在位置におけるフローの大きさや方向の復元である。以下に各ステップにおける手順を示す。

2.1 シーンフロー存在位置の判定

シーンとカメラ間のカメラ行列が与えられた場合、シーンフローと二次元画像上のオプティカルフローは次式のように表現できる。

$$\frac{d\mathbf{u}_k}{dt} = \frac{\partial \mathbf{u}_k}{\partial \mathbf{x}} \frac{d\mathbf{x}}{dt} \quad (1)$$

$\frac{d\mathbf{u}_k}{dt}$ は k 番目のカメラ中の点 \mathbf{u}_k におけるオプティカルフローであり、 $\frac{d\mathbf{x}}{dt}$ はシーン内の点 \mathbf{x} におけるシーンフローである。 $\frac{\partial \mathbf{u}_k}{\partial \mathbf{x}}$ はシーン内の点 $\mathbf{x} = [X_x, X_y, X_z]^T$ と、それをカメラ k に投影した画像座標 $\mathbf{u}_k = [u_k, v_k]^T$ との微分関係である。ここで式 (1) を次式のように変形する。

$$\frac{d\mathbf{x}}{dt} = \left(\frac{\partial \mathbf{u}_k}{\partial \mathbf{x}} \right)^* \frac{d\mathbf{u}_k}{dt} + \mu \mathbf{r}_k(\mathbf{u}_k) \quad (2)$$

$$\left(\frac{\partial \mathbf{u}_k}{\partial \mathbf{x}} \right)^* : \frac{\partial \mathbf{u}_k}{\partial \mathbf{x}} \text{ の一般化逆行列}$$

$\mathbf{r}_k(\mathbf{u}_k)$: 点 \mathbf{u}_k を通る光線ベクトル μ : 不定値

光線と画像中のオプティカルフローのなす面に垂直なベクトルは、シーンフロー $\frac{d\mathbf{x}}{dt}$ に対し垂直となるため、式 (2) を以下のように変形する。

$$\begin{aligned} \mathbf{m}_k(\mathbf{x}) \cdot \frac{d\mathbf{x}}{dt} \\ \equiv \left[\left(\frac{\partial \mathbf{u}_k}{\partial \mathbf{x}} \right)^* \frac{d\mathbf{u}_k}{dt} \times \mathbf{r}_k(\mathbf{u}_k) \right] \cdot \frac{d\mathbf{x}}{dt} = 0 \quad (3) \end{aligned}$$

式 (3) におけるベクトル $\mathbf{m}_k(\mathbf{x})$ はシーンフローに垂直なベクトルである。もし三次元空間中の点 \mathbf{x} においてシーンフローが存在する場合、図 1 に示すように、すべてのカメラから計算される $\mathbf{m}_k(\mathbf{x})$ は同一の平面上に存在する。各カメラから得られた $\mathbf{m}_k(\mathbf{x})$ が同一の平面上に存在するかどうかの判定には次式を用いる。

$$M(\mathbf{x}) = \sum_k \hat{\mathbf{m}}_k \hat{\mathbf{m}}_k^T \quad (4)$$

ここで、 $\hat{\mathbf{m}}_k$ は $\mathbf{m}_k(\mathbf{x})$ を単位ベクトルに正規化したものである。ベクトル $\mathbf{m}_k(\mathbf{x})$ がすべて同一の平面上に存在する場合、式 (4) における行列 $M(\mathbf{x})$ の最小の

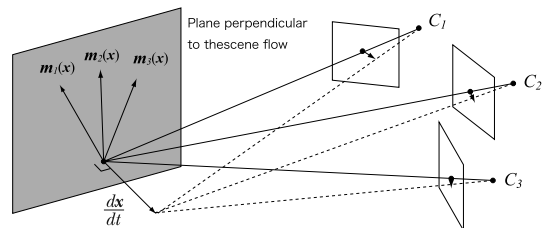


図 1 $\mathbf{m}_k(\mathbf{x})$ の性質
Fig. 1 Property of $\mathbf{m}_k(\mathbf{x})$.

固有値は 0 となる．したがって，行列 $M(x)$ の最小の固有値が 0 となるような点を探すことで三次元シーンフローの存在位置を判定することが可能となる．

シーンフローの存在位置を判定するには，三次元空間中を Voxel に区切り，その各点を順に探索する必要がある．この探索には，注目点が投影されるカメラにおいて可視か不可視かを考える必要があるため，Seitz らが提案した Voxel Coloring アルゴリズムを用いる [7]．Voxel Coloring アルゴリズムは探索の順番と Z バッファを用いることで，可視か不可視かを判定する．本研究では，Voxel Coloring アルゴリズムの探索法を用い，カメラ間のオプティカルフローの一致性から，三次元シーンフローの存在位置を判定する．

2.2 シーンフローの大きさと方向の復元

シーンフローが存在する点では，オプティカルフローとシーンフローの関係式 (1) は， n 台のカメラに適用すると，次式のように表現できる．

$$\begin{bmatrix} \frac{\partial u_1}{\partial t} \\ \frac{\partial v_1}{\partial t} \\ \vdots \\ \frac{\partial u_n}{\partial t} \\ \frac{\partial v_n}{\partial t} \end{bmatrix} = \begin{bmatrix} \frac{\partial u_1}{\partial X_x} & \frac{\partial u_1}{\partial X_y} & \frac{\partial u_1}{\partial X_z} \\ \frac{\partial v_1}{\partial X_x} & \frac{\partial v_1}{\partial X_y} & \frac{\partial v_1}{\partial X_z} \\ \vdots & \vdots & \vdots \\ \frac{\partial u_n}{\partial X_x} & \frac{\partial u_n}{\partial X_y} & \frac{\partial u_n}{\partial X_z} \\ \frac{\partial v_n}{\partial X_x} & \frac{\partial v_n}{\partial X_y} & \frac{\partial v_n}{\partial X_z} \end{bmatrix} \frac{dx}{dt} \quad (5)$$

シーンフロー $\frac{dx}{dt}$ の未知数は 3 であるのに対し，式 (6) は $2n$ の式を作ることができるため，点 x が 2 台以上のカメラで観測された場合，最小二乗法を用いることで，シーンフローの大きさと方向を復元することができる．

3. Subspace 拘束によるシーンフローの修正

一般に，二次元オプティカルフローには，多くのノイズや異常値が含まれることが知られている．そのため，オプティカルフローから復元される三次元シーンフローにも多くのノイズや異常値が含まれる可能性がある．そこで本論文では，物体の運動が短時間において剛体運動であることを仮定し，三次元シーンフローを修正し，精度を向上させる手法を提案する．

3.1 運動モデルと観測行列

シーン内の運動が剛体であると仮定した場合，ある時刻 j における対象物体中の点 x_i のシーンフロー $\frac{dx_{ij}}{dt}$ は，次式のように，平行移動ベクトル t_j と，回転速度ベクトル ω_j を用いて表現することができる

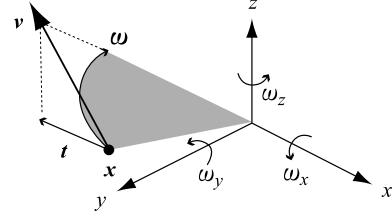


図 2 運動モデル
Fig. 2 Motion model.

(図 2 参照)．ここで，シーンフローは運動を表したベクトルとして扱うため，点 x におけるシーンフロー $\frac{dx_{ij}}{dt}$ を v_{ij} とする．

$$\begin{aligned} v_{ij} &= t_j + \omega_j \times x_i & (6) \\ v_{ij} &= [v_{x_{ij}}, v_{y_{ij}}, v_{z_{ij}}]^T, \quad x_i = [X_{x_i}, X_{y_i}, X_{z_i}]^T \\ t_j &= [t_{x_j}, t_{y_j}, t_{z_j}]^T : \text{平行移動ベクトル} \\ \omega_j &= [\omega_{x_j}, \omega_{y_j}, \omega_{z_j}]^T : \text{回転速度ベクトル} \\ i &: \text{座標番号}, \quad j : \text{フレーム番号} \end{aligned}$$

式 (6) を展開すると次式となる．

$$\begin{bmatrix} v_{x_{ij}} \\ v_{y_{ij}} \\ v_{z_{ij}} \end{bmatrix} = \begin{bmatrix} t_{x_j} & 0 & 0 & 0 & -\omega_{z_j} & \omega_{y_j} \\ 0 & t_{y_j} & 0 & \omega_{z_j} & 0 & -\omega_{x_j} \\ 0 & 0 & t_{z_j} & -\omega_{y_j} & \omega_{x_j} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ X_{x_i} \\ X_{y_i} \\ X_{z_i} \end{bmatrix} \quad (7)$$

ここで，式 (7) を次式に置き換える．

$$\begin{bmatrix} v_{x_{ij}} \\ v_{y_{ij}} \\ v_{z_{ij}} \end{bmatrix} = \begin{bmatrix} s_{x_j} \\ s_{y_j} \\ s_{z_j} \end{bmatrix}_{(3 \times 6)} q_i(6 \times 1) \quad (8)$$

$$\begin{aligned} s_{x_j} &= [t_{x_j}, 0, 0, 0, -\omega_{z_j}, \omega_{y_j}] \\ s_{y_j} &= [0, t_{y_j}, 0, \omega_{z_j}, 0, -\omega_{x_j}] \\ s_{z_j} &= [0, 0, t_{z_j}, -\omega_{y_j}, \omega_{x_j}, 0] \\ q_i &= [1, 1, 1, X_{x_i}, X_{y_i}, X_{z_i}]^T \end{aligned}$$

ここで，式 (7) は空間中の 1 点のみの運動を表現した式であるため，式 (8) を以下のように拡張する．

$$\begin{bmatrix} v_{x_j} \\ v_{y_j} \\ v_{z_j} \end{bmatrix}_{(3 \times N)} = \begin{bmatrix} s_{x_j} \\ s_{y_j} \\ s_{z_j} \end{bmatrix} Q_{(6 \times N)} \quad (9)$$

$$Q_{(6 \times N)} = [q_1, \dots, q_N] = \begin{bmatrix} 1 & \dots & 1 \\ 1 & \dots & 1 \\ 1 & \dots & 1 \\ X_{x_1} & \dots & X_{x_N} \\ X_{y_1} & \dots & X_{y_N} \\ X_{z_1} & \dots & X_{z_N} \end{bmatrix}$$

更に式 (9) の s_{x_j} , s_{y_j} , s_{z_j} をそれぞれ次式のように M フレームへと拡張し, 式 (10) を得る.

$$S_x = \begin{bmatrix} s_{x_1} \\ \vdots \\ s_{x_M} \end{bmatrix}, \quad S_y = \begin{bmatrix} s_{y_1} \\ \vdots \\ s_{y_M} \end{bmatrix}, \quad S_z = \begin{bmatrix} s_{z_1} \\ \vdots \\ s_{z_M} \end{bmatrix}$$

$$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}_{(3M \times N)} = \begin{bmatrix} S_x \\ S_y \\ S_z \end{bmatrix}_{(3M \times 6)} Q_{(6 \times N)} \quad (10)$$

ここで, $[V_x/V_y/V_z]$ を観測行列と呼ぶ. Q の行列は, 上 3 列がすべて 1 であるため, 下 3 列と合わせるとランクは 4 以下となる. また, 瞬間的な物体の運動は等速であると仮定した場合, S_x, S_y, S_z 各行列のランクはすべて 1 以下となるため, 行列 $[S_x/S_y/S_z]$ のランクは 3 以下となる. 観測行列 $[V_x/V_y/V_z]$ は行列 Q と行列 $[S_x/S_y/S_z]$ の積で表現されており, 小さい方のランクに制限されるため, そのランクは 3 以下となる. 以下では, 対象とする運動を等速であると仮定し, 観測行列のランクが 3 以下となることを利用する.

3.2 Subspace 拘束によるフローの修正

二次元オプティカルフローにノイズや異常値が含まれる場合, シーンフローから得られる観測行列のランクは 3 以上となる. そこで復元されたシーンフローに対し, Subspace 拘束によるランク制限を施すことで精度を向上させる.

Subspace 拘束を用いてランク制限を施すために, 複数フレームのシーンフローの成分を v_x, v_y, v_z それぞれ観測行列に当てはめる. 次に, 観測行列に対し, 特異値分解を行い次式を得る.

$$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}_{(3M \times N)} = U_1 D U_2^T \quad (11)$$

U_1 は $3M \times 3M$ の直交行列, U_2 は $N \times N$ の直交行列, D は特別な対角行列であり. D の対角要素は

特異値と呼ばれ, ベクトル $d = [d_1, d_2, \dots, d_{3M}]^T$ で表す. ここで, 観測行列のランクを 3 に制限するには対角成分が $d' = [d_1, d_2, d_3, 0, \dots, 0]^T$ となる行列 D' を次式のように求める.

$$D' = \begin{bmatrix} d_1 & 0 & \dots & \dots & \dots & 0 \\ 0 & d_2 & 0 & \dots & \dots & 0 \\ 0 & 0 & d_3 & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{bmatrix}_{(3M \times N)}$$

この D' から, U_1 と U_2 を用いて新しい観測行列を以下のように計算する.

$$\begin{bmatrix} V_x' \\ V_y' \\ V_z' \end{bmatrix}'_{(3M \times N)} = U_1 D' U_2^T \quad (12)$$

得られた新しい観測行列 $[V_x'/V_y'/V_z']'$ は, ランクが 3 の行列となり, 式 (6) に示す運動モデルに従うシーンフローに修正することができる.

4. シーンフローを用いたモーション推定

本章では, 修正した三次元シーンフローからのモーション推定について述べる.

4.1 モーション推定

シーン内のある点 x_i における運動 v_{ij} が得られた場合, 時刻 j における式 (6) は次式のように変形できる.

$$v_{ij} = \begin{bmatrix} 0 & X_{z_i} & -X_{y_i} & 1 & 0 & 0 \\ -X_{z_i} & 0 & X_{x_i} & 0 & 1 & 0 \\ X_{y_i} & -X_{x_i} & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_j \\ t_j \end{bmatrix}$$

$$= \begin{bmatrix} -[x_i]_{\times} & I_{(3 \times 3)} \end{bmatrix} \begin{bmatrix} \omega_j \\ t_j \end{bmatrix} \quad (13)$$

$[x_i]_{\times}$ x_i のひずみ対称行列,

$I_{(3 \times 3)}$: 3×3 の単位行列

これを N 点に拡張すると次式を得る.

$$\begin{bmatrix} v_{1j} \\ \vdots \\ v_{Nj} \end{bmatrix} = \begin{bmatrix} -[x_1]_{\times} & I_{1(3 \times 3)} \\ \vdots & \vdots \\ -[x_N]_{\times} & I_{N(3 \times 3)} \end{bmatrix} \begin{bmatrix} \omega_j \\ t_j \end{bmatrix} \quad (14)$$

ここで, 未知数は $\omega_j = [\omega_{x_j}, \omega_{y_j}, \omega_{z_j}]^T$ と $t_j =$

$[t_{x_j}, t_{y_j}, t_{z_j}]^T$ の六つとなる．式 (14) からは， $3 \times N$ の式を得ることができるが，ひずみ対称行列 $[x_i]_{\times}$ のランクは 2 であるため， $N = 2$ の場合，式 (14) は一意に解を決定することができない．よって，3 点以上の v_{ij} と x_i の組から最小二乗法を用いることで，剛体運動の各パラメータを推定する．

4.2 RANSAC

3. で述べた Subspace 拘束により修正した三次元シーンフローは，運動モデルのランクに従った運動となる．しかし，Subspace 拘束ではシーンフローの大きさと方向を修正することは可能であるが，本来存在しないところから得られたシーンフローを除去することはできない(図 3 参照)．そこで，このようなアウトライアを除去してパラメータを推定するために，RANSAC を用いる [3]．以下に RANSAC に基づくモーション推定の手順を示す．

- Step 1 同時刻のすべてのシーンフローからランダムに 3 点選択
- Step 2 最小二乗法を用い各パラメータを推定
- Step 3 推定されたパラメータから各点のフローを計算
- Step 4 計算されたフローとシーンフローとの誤差を計算
- Step 5 誤差の中央値を選択
- Step 6 Step 1 ~ Step 5 を繰り返し，ソーティング後の誤差の中間値が最小となるパラメータを最終的なパラメータとする

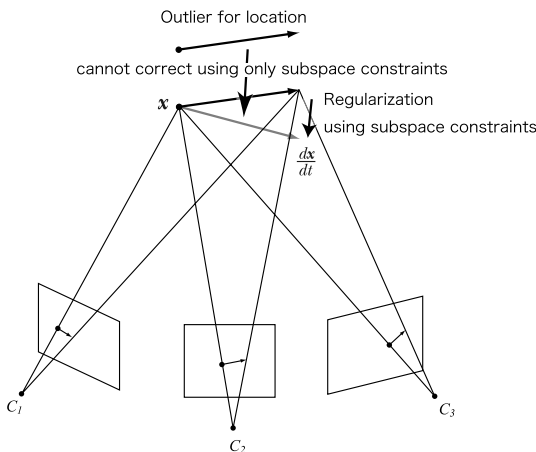


図 3 位置に対するアウトライア
Fig. 3 Outlier for location of scene flow.

5. 実験

以上に述べた手法により，三次元シーンフロー復元とモーション推定の実験を行う．シミュレーションでは，形状が既知である物体に対し，大きさと方向に対しノイズを加え，位置に対しアウトライアを加えたシーンフローを作成し，Subspace 拘束による修正とモーション推定の実験を行う．次に，実際に 5 台のカメラから得られる実画像を対象とした実験を行い，本手法の有効性を示す．また，本実験では，基準フレームから前後 1 フレームのシーンフローから観測行列を作成し，修正を行う ($M = 2$)．

5.1 シミュレーションによるシーンフロー復元実験

シミュレーションでは，立方体，球，任意の形状の 3 種類の物体に対し，平行移動，回転運動，平行移動と回転が同時に発生した場合の計 9 シーケンスのシーンフローを用いる．すべての形状は 0~1 の大きさに正規化されている．生成したシーンフローの大きさと方向に対し，正規乱数を用いてノイズを与え．更に，シーンフローの存在位置に対し，10 点に 1 点の確率でアウトライアを生成する．シミュレーションにより生成したシーンフローと，Subspace 拘束を用いた修正結果の例を図 4 に示す．シーン内の運動する物体を剛体であると仮定した場合，そのシーンフローを局部的に観測すると，全体的に一樣なシーンフローとなると考えられる．図 4(a) では，フローの向きが分散しているが，Subspace 拘束を用いて修正することで，図 4(b) に示すように一樣なシーンフローを得ることができた．

シーンフローの精度の評価として，シーンフローの

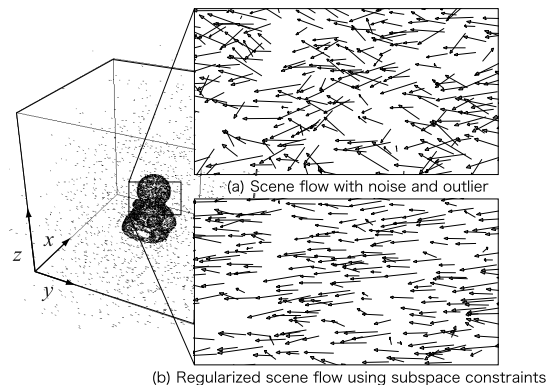


図 4 シミュレーション実験結果の例
Fig. 4 Example of simulation.

真値 v_t と復元されたシーンフロー v との方向の類似度を式 (15) より、大きさに対して、真値とのノルムの差を式 (16) より計算する。

$$\text{方向に対する類似度} : \cos \theta = \frac{v \cdot v_t}{\|v\| \|v_t\|} \quad (15)$$

$$\text{真値とのノルムの差} : \frac{|\|v\| - \|v_t\||}{\|v_t\|} \quad (16)$$

実験に用いる 9 シーケンスに対し、方向の類似度を平均した結果を図 5 に、大きさの差を平均した結果を図 6 に示す。図 5 はシーンフローの方向の類似度のヒストグラムであり、修正前は、誤差 10 度以内のシーンフローは全体の 56.7% であるのに対し、修正後では 94.2% まで向上した。図 6 は大きさに対するヒストグラムであり、修正前は、誤差 10% 以内のシーンフローは全体の 46.3% であるのに対し、修正後では、58.3% まで向上した。

5.2 シミュレーションモデルを用いたモーション推定実験

5.1 のシミュレーションで復元された三次元シーンフローからモーション推定の実験を行う。修正後のシーンフローを入力とし、それぞれのモーションを推定し、評価を行う。また、修正前のシーンフローの全点から推定した結果と比較する。推定されたモーションパラメータの誤差をプロットしたものを図 7 に示す。図 7(a) の結果より、平行移動のみのシーケンスでは、全点を用いた結果が最も良い結果を示している。これは、平行移動のみの場合、各点のシーンフローは、存在位置に依存しないことが原因であり、ノイズを含んだシーンフローから RANSAC を適用した場合、フローの組合せにより精度が下がるということが考えられる。これに対し、修正後の RANSAC では高い精度で推定が可能であることが分かる。これは、修正を行うことで、シーンフローに含まれるノイズの影響を低減することができたためであると考えられる。

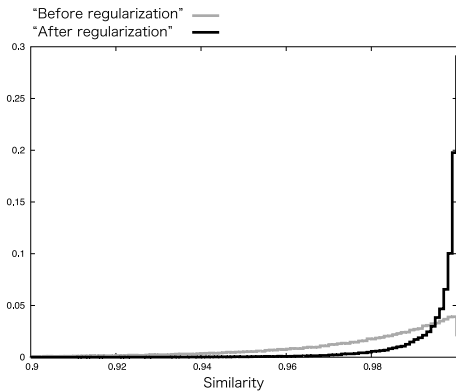


図 5 類似度のヒストグラム
Fig. 5 Histogram of similarity.

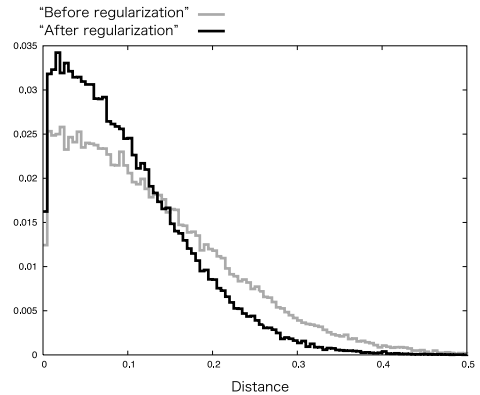


図 6 真値とのノルムの差のヒストグラム
Fig. 6 Histogram of difference from true magnitude.

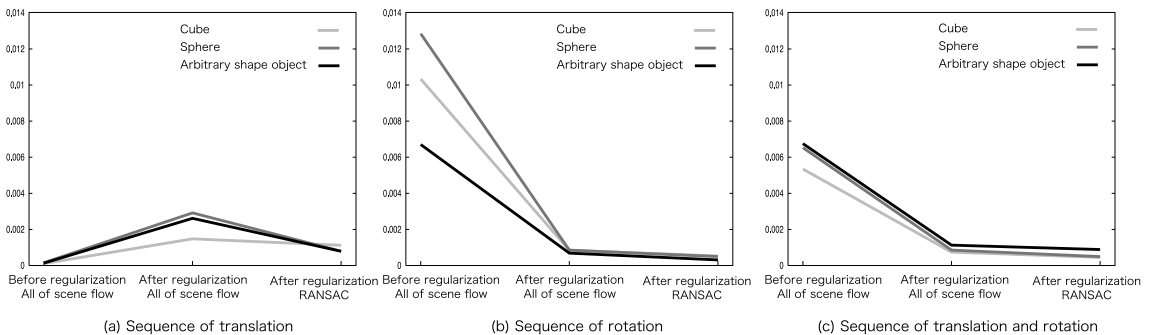


図 7 モーションパラメータの推定誤差
Fig. 7 Estimating errors of motion parameters.

図 7 (b), (c) は、回転を含んだシーケンスであり、全点を用いた場合に精度が極端に落ちていることが分かる。これは、回転運動はシーンフローの存在位置に依存したパラメータであり、アウトライアが影響しているためだと考えられる。これに対し、RANSAC を用いることで、高精度にパラメータを推定することが可能となる。

5.3 実画像を用いた実験

実画像を用いたモーション推定実験を行う。実験環境は、図 8 に示すように、三次元空間中に配置した 5 台のカメラを用いる。各カメラのパラメータは既知である。各カメラ間のシャッタータイミングは同期している。また、各カメラは固定してあり、背景は運動しないとする。対象とする物体は、図 9 に示すように、

ルービックキューブをもち、空間内で運動する人間の腕部とする。オプティカルフローの計算にはブロックマッチング法を用いる。本実験では、 y 軸方向に 1 秒間に約 15 cm 平行移動した場合と、 x 軸を中心に約 180 度回転した場合の二つのシーケンスを対象とし、入力画像は基準フレームから前後 1 フレームの計 15 枚 (3 フレーム \times 5 台) を用いる。

図 9 に示す平行移動における入力オプティカルフローから復元されたシーンフローを図 10 に示す。Vedula らの手法では、シーンフローの存在位置を、各入力カメラから算出されるベクトル m_k の一致性から判定する。これにより、図 10 (a) では、シーン内の唯一の動物体である腕部のシーンフローのみを復元することが可能であることが分かる。次に、平行移動と回転運動でのモーション推定結果を表 1 に示す。表 1 より、平行移動においては、回転速度ベクトル $\omega_x, \omega_y, \omega_z$ の値はすべてが 1 度以下であることから、回転運動が観測されていないことが分かる。また、最も数値の高いパラメータは y 軸方向の移動であり、1 frame (1/30 s) で約 4.5 mm 移動したことが分かる。これは、1 秒間で 13.5 cm の平行移動となり、目測による 15 cm に対し、高い精度で推定ができていることが分かる。次に回転運動では、 x 軸周りの回転が 1 frame で約 5.5 度と推定された。実験に用いたシーケンスでは、目測で 1 秒間に約 180 度回転しているため、1 frame では

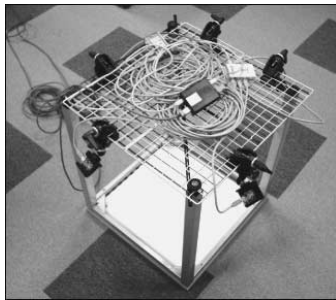


図 8 実験環境
Fig. 8 Experimental environment.

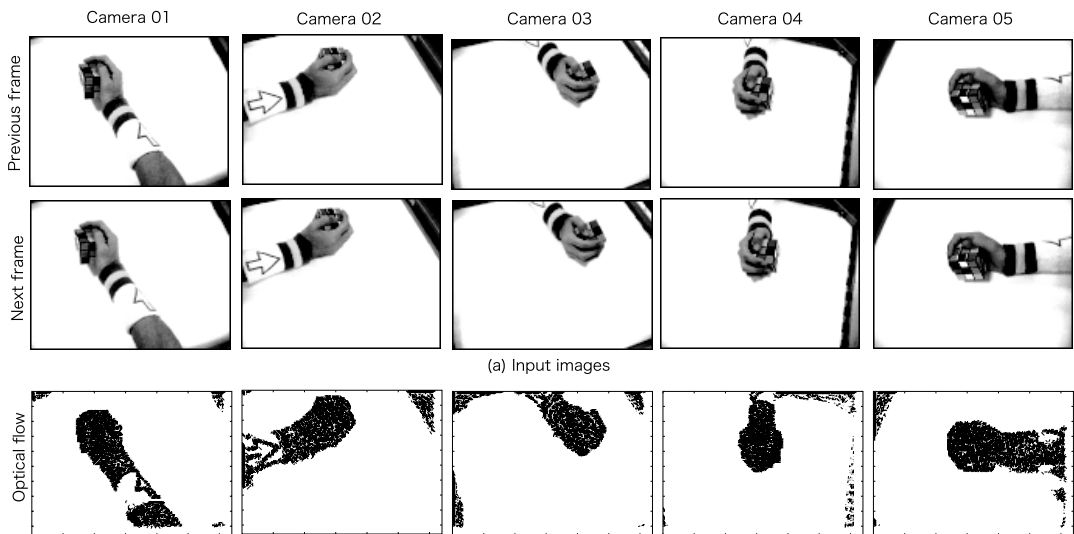


図 9 実画像における入力画像とオプティカルフロー (平行移動)
Fig. 9 Input real image and optical flow (translation).

表 1 実画像を用いた実験におけるモーション推定結果
Table 1 Estimated results of motion for real images.

	ω_x	ω_y	ω_z	t_x	t_y	t_z
	[deg/frame]	[deg/frame]	[deg/frame]	[mm/frame]	[mm/frame]	[mm/frame]
平行移動	-0.494	0.380	-0.244	0.818	-4.523	-0.078
回転運動	-5.490	-0.114	0.978	-0.246	5.901	2.197

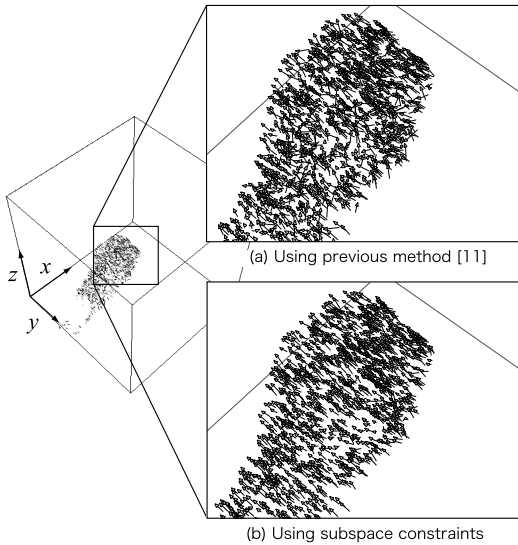


図 10 平行移動における三次元シーンフローの例
Fig. 10 Example of reconstructed 3-D scene flow (translation).

約 6 度となり、高い精度でモーションが推定できていることが分かる。しかし、平行移動していないにもかかわらず、 t_y 成分が高い値を示している。これは、上方からこの運動を観測した場合、その表面の運動は t_y 成分の運動として観測されることが原因である。この問題はカメラの台数を増やすことで解消できると考えられる。

6. 考 察

本手法は、複数の画像中のオプティカルフローにより復元された三次元シーンフローに対し、Subspace 拘束を用いることで、精度を向上させる。その後、Subspace 拘束により修正された三次元シーンフローに対し、RANSAC を用いてアウトライアの影響を除去し、モーション推定を行う。ここで、アウトライアを除去したシーンフロー集合に対し再度ランク制限を行えば、より良い結果が得られることが考えられる。しかし、一度ランク制限を行ったシーンフロー集合は、はずれ値を除去した後、新たに観測行列を作成しても、既に

ランクは 3 である。したがって、Subspace 拘束を用いても観測行列は変化しないため、このような処理により精度を向上させることは不可能である。

次に、観測されるすべてのシーンフローが同一の剛体であるという仮定に対する考察を示す。本論文における実験環境では、すべてのカメラは固定されており、単一物体が運動している状況を撮影している。この環境下では、本手法は十分な有効性を発揮する。しかし、複数の剛体運動が観測される状況下ではランク制限によるシーンフローの修正や、モーション推定ができないという問題が発生する。このような複数の剛体が同時に観測された場合の対策としては、モーションセグメンテーションの利用が有効であると考えられる。二次元のオプティカルフローを用いたモーションセグメンテーションの手法は、既に Liu らにより提案されており [10]、これを用いて画像中のオプティカルフローをモーションごとにラベル付けることで、複数の剛体の分離が可能であると考えられる。分離後のシーンフローに本手法を適用することで、剛体ごとにモーション推定が可能となる。

7. む す び

本論文では、Subspace 拘束を用いた三次元シーンフローの修正手法と、モーション推定法を提案した。Subspace 拘束による三次元シーンフローの修正では、対象物の運動が短時間では剛体であると仮定することで、精度を向上することが可能であることを示した。本手法の特徴として、修正の際に必要な複数フレームのスタック数を少なくすることが挙げられ、実験により、基準フレームから前後 1 フレームのフローから精度を向上させることが可能であることを実証した。

また、修正されたフローから対象物のモーションパラメータを推定し、高い精度でモーションを推定することが可能であることを確認した。従来手法により復元される三次元シーンフローは、フローの大きさ、方向に対するノイズと、フローの存在位置に対するアウトライアが存在する。本手法では、フローの大きさ、

方向に対するノイズには、Subspace 拘束による修正で対応し、フローの存在位置に対するアウトライアには、RANSAC により対応し、高い精度で推定することができた。

今後は、複数物体への対応として、モーションセグメンテーションを検討する予定である。

謝辞 本研究は、文部科学省ハイテク・リサーチ・センター整備事業（平成 16 年度～平成 20 年度）による私学助成を得て行われている。また、様々な意見を頂いたカーネギーメロン大学金出武雄教授、株式会社日立製作所数井誠氏、中部大学奥村吉孝教授に心から厚く御礼申し上げます。

文 献

- [1] B.D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," Imaging Understanding Workshop, pp.121–130, 1981.
- [2] B.K.P. Horn and B.G. Schunck, "Determining optical flow," Artif. Intell., vol.17, pp.185–203, 1981.
- [3] M. Fischer and R. Bolles, "Randoam sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Commun. ACM, vol.24, no.6, pp.381–385, 1981.
- [4] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," Int. J. Comput. Vis., vol.9, no.2, pp.137–154, 1992.
- [5] J.L. Barron, D.J. Fleet, and S.S. Beauchemin, "Performance of optical flow techniques," Int. J. Comput. Vis., vol.12, no.1, pp.43–77, 1994.
- [6] M. Irani, "Multi-frame optical flow estimation using subspace constraints," Proc. IEEE International Conference on Computer Vision, vol.1, pp.626–633, 1999.
- [7] S.M. Seitz and C.R. Dyer, "Photorealistic scene reconstruction by voxel coloring," Int. J. Comput. Vis., vol.35, no.2, pp.151–173, 1999.
- [8] J.Y. Bouguet, "Pyramidal implementation of the Lucas Kanade feature tracker," OpenCV Documentation, Microprocessor Research Labs, Intel Corp., 2000.
- [9] S.M. Seitz and K.N. Kutulakos, "Plenoptic image editing," Int. J. Comput. Vis., vol.48, no.2, pp.115–129, 2002.
- [10] C. Liu, A. Torralba, W.T. Freeman, F. Durand, and E.H. Adelson, "Motion magnification," International Conference on Computer Graphics and Interactive Techniques ACM SIGGRAPH, pp.519–526, 2005.
- [11] S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade, "Three-dimensional scene flow," IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.3, pp.475–480, 2005.
- [12] 青木公也, 輿水大和, "時系列距離画像を用いた 3D フロー検出の検討," 動的画像処理実利用化ワークショップ 2005, pp.140–145, 2005.
(平成 18 年 10 月 6 日受付, 19 年 2 月 9 日再受付)



水野 宏基

2005 中部大・工・情報卒。同年より同修士課程に在籍。コンピュータビジョンに関する研究に従事。情報処理学会会員。



藤吉 弘亘 (正員)

1997 中部大学大学院博士後期課程了。1997～2000 米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow。2000 中部大学講師を経て 2004 より同大助教授。2006 米カーネギーメロン大学ロボット工学研究所客員研究員。工博。計算機視覚、動画像処理、パターン認識・理解の研究に従事。2005 年度ロボカップ研究賞。情報処理学会、電気学会、IEEE 各会員。



岩堀 祐之 (正員)

1983 名工大・情報卒, 1985 東工大大学院修士課程・電気電子工学専攻了。1988 同大学院博士課程了。工博。同年名工大・情報処理教育センター・助手。1992 同大・助教授, 2002 名工大情報メディア教育センター教授, 2004 中部大・工・情報工学科教授, 現在に至る。この間, カナダブリティッシュコロロンビア大学客員研究員。コンピュータビジョン, ニューラルネットワーク, 複合現実感に関する研究に従事。情報処理学会, IEEE 各会員。