オプティカルフローを用いた複雑背景下における 人物の腕領域の抽出と運動パラメータ推定

非会員 玉 木 徹 (名古屋大学, 理化学研究所)
非会員 山 村 毅 (愛 知 県 立 大 学)
正 員 大 西 昇 (名古屋大学, 理化学研究所)

Extraction of Human Limb Regions and Parameter Estimation using Optical Flow against Complex Background

Toru TAMAKI, Non-member (Nagoya Univ., RIKEN), Tsuyoshi YAMAMURA, Non-member (Aichi Prefectural Univ.), Noboru OHNISHI, Member (Nagoya Univ., RIKEN)

We propose a method for extracting rotating limb regions and estimating their motion parameters, which is based on curl of optical flow. Regarding optical flow under orthographic projection as a vector field, the proposed method computes the curl of the optical flow to extract rotating regions not depending on the position of the limb. We use the EM algorithm which can perform both of dividing the image into multiple moving regions and estimating motion parameters of each region. Since the proposed method doesn't require a simplicity of background, and can discriminate translational movements produced by other moving objects in the background, it is applicable to a practical situation in which conventional methods don't work well. Experimental results with real image sequences demonstrate that the limb extraction can be performed well from noisy optical flow which are obtained from real images.

キーワード：curl, オプティカルフロー, EM アルゴリズム, 姿勢推定, 領域分割, 腕

1. はじめに

動画画像から人物の領域を抽出することは、人物動作の認識システムの重要な一部分となっており、インターフェイスのためのジェスチャ認識やバーチャルリアリティにおける動作の再現、セキュリティ監視などへの応用が期待される。このような動作認識の手法としては、対象人物にマーカー⁽¹⁾⁽²⁾やセンサ⁽³⁾⁽⁴⁾を取り付けて認識を行う研究が行われている。これらは人物の動きを正確に測定できる反面、システムが設置される場所、対象者の動作やその行動範囲を制限しなければならないといった問題点がある。一方で、カメラからの画像に基づいた認識システムも研究され、この方式には、カメラが複数台であっても比較的容易に設置できること、また人物に何も取り付ける必要がなく、行動が制限されないという利点がある。

この画像に基づいた人物動作認識の研究には、人体モデルを用いるものとそうでないものがある。人体モデルを用いない方法^{(5)~(8)}は、どんな動作をしているのかということ进行分类することはできるが、動作を再現することはできない。そこで、実際の体の各部の動きを再現するために、パラ

メータ表現された人体モデルが多く用いられており、この時に、対象となる人物の領域を抽出することが必要となる。これらのモデルは、通常、関節で結合されたリンク^{(9)~(11)}や、2Dの矩形⁽¹²⁾⁽¹³⁾、3Dの円筒や多面体^{(14)~(19)}などで構成されている。こういった人体モデルを用いた認識手法には、大きく分けて以下の二つがある。

一つは、ある姿勢をとったモデルを画像平面に投影し、その投影像と、あらかじめ抽出しておいた対象人物の領域との違いを評価し照合する手法⁽¹⁰⁾⁽¹⁴⁾⁽¹⁵⁾⁽¹⁹⁾である。これにより、抽出された領域にシルエットが最も似ているモデルのパラメータが得られ、対象者の姿勢が認識できる。この時、人物領域の抽出には、背景差分や動物体検出などの手法が用いられる。その多くは、背景差分を簡単にするために背景が既知か少なくとも一様な色であることを仮定しており、また背景に他の移動物体が存在しないことが前提となっている。また照合過程において膨大な探索を必要とすることも問題である。

もう一つは、人体の運動学的モデルを用いて人物の運動を追跡する手法⁽⁹⁾⁽¹²⁾⁽¹⁷⁾である。これらの手法は、動きの情報を用いているため、背景の複雑さには影響されないと

いう利点がある。しかし実画像から得られるオプティカルフローは、エッジや動きの不連続な部分では信頼できるフローが得られず⁽²⁰⁾、それを避けるための過度な平滑化はパラメータの推定に影響を与える。そのため、画素値の時空間勾配を直接用いる手法⁽¹⁷⁾や、ステレオによって動領域を分割する⁽²¹⁾などの手法がとられているが、計算コストが高いことや複数のカメラを必要とすることが問題となる。また多くの場合、追跡する初期フレームや蓄積誤差が大きくなったフレームで⁽²²⁾、その時刻での画像内の人物の位置と姿勢を指定しなければならず、モデルを人物に当てはめるための対話的な操作には労力を必要とする。

以上のように、未知の複雑な背景において他の移動物体が存在する、より実際的な状況で、人物を抽出することが必要になってくる。そのためにまず最初に、ジェスチャなどで情報を多く伝える、人物の腕を抽出することを考える⁽²³⁾。

本論文では、そのような条件の下で、回転する人物の腕の領域抽出と、その運動パラメータの推定を行う手法を提案する。本手法は正射影を仮定しており、腕の回転面が x 軸か y 軸のどちらかにのみ傾いている回転運動（図 1(a)(b)）を対象としている。複数の運動を扱うために EM アルゴリズムを導入し、実画像から得られたノイズの多いオプティカルフローから良好な推定を行うことができる。提案手法はオプティカルフローの curl に基づいているので、背景に依存せずに回転している腕領域を抽出することができ、背景の移動物体も除去できるという特長がある。また各時刻で領域とパラメータが得られるため、前述した追跡手法で必要となる初期状態を与える必要がなく、フレーム毎に照合のための探索を行うこともない。

以下 2 章ではオプティカルフローに基づいた領域抽出について説明し、3 章において運動パラメータの推定手法について、4 章では EM アルゴリズムを用いた領域分割と推定のアルゴリズムについて述べ、実画像を用いた評価と実験結果を 5 章で示す。

2. Curl に基づく回転領域抽出

本手法の領域抽出は、ベクトル場の curl に基づいている。この章では、回転物体によって生成されたオプティカルフローの curl の性質について説明する。そして単純な回転と一般的な回転についての curl の例を示し、それによって腕領域の抽出が行えることを示す。

〈2・1〉 回転ベクトル場の Curl の性質 三次元空間の点 $\mathbf{x} = (x, y, z)^T$ の速度を $\dot{\mathbf{p}}(\mathbf{x}) = (u, v, w)^T$ とすると、この速度ベクトルはベクトル場を形成する。このベクトル場の curl[†] は以下の式で定義される。

$$\nabla \times \dot{\mathbf{p}}(\mathbf{x}) \equiv \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right)^T$$

[†] 回転や rot という語が使われることが多い⁽²⁴⁾が、実際に物体が回転運動することを指す場合との混乱を避けるために、ここでは curl という語を用いることにする。

図 1 スティックモデル：(a) 画像平面と平行な面上での回転 (b) 回転面が画像平面 ($O-xy$) に対して y 軸回りに (c) x 軸回りに傾いている場合

Fig. 1. The stick models moving on a plane.

回転によって生成されたベクトル場において、この curl は次のような性質を持つ。

$$\nabla \times \{\dot{\mathbf{p}}(\mathbf{x} - \mathbf{c}) + \mathbf{t}\} = \nabla \times \dot{\mathbf{p}}(\mathbf{x}) \quad \dots\dots\dots (1)$$

ここで $\mathbf{c}$ は任意のベクトル、 $\mathbf{t}$ は位置に依存しない定数ベクトル場である。二次元画像平面内のオプティカルフローは、 $z = w = 0$ としたベクトル場と考えることができる。回転物体によって生成されたオプティカルフローについて考えると、 $\mathbf{c}$ は回転中心であり、 $\mathbf{t}$ は物体の並進運動であると見なすことができる。したがって、画像内のどこにいても curl の値は変わらず、また物体が並進運動していても影響がない。 $\mathbf{t}$ はまた、カメラや背景の運動物体によって生じる並進運動とみなすことができるので、curl の性質によってそれらの影響を排除することができる。

〈2・2〉 Curl の例 まず初めに、画像平面に平行な面上で回転している図 1(a) のような単純な場合を考える（以下では正射影を仮定している）。点 $\mathbf{p}$ を、回転中心である原点から r ($0 \leq r \leq r_1$) だけ離れ、角速度 ω で回転し、角度 $\theta (= \omega t)$ の位置にある腕上の点とする。このとき、 $\dot{\mathbf{p}}$ の速度 $\dot{\mathbf{p}}$ （画像上でのオプティカルフロー）とその curl $\nabla \times \dot{\mathbf{p}}$ は次のように表される。

$$\begin{aligned} \mathbf{p} &= (x, y, 0)^T = (r \cos \theta, r \sin \theta, 0)^T \\ \dot{\mathbf{p}} &= (-\omega r \sin \theta, \omega r \cos \theta, 0)^T = (-\omega y, \omega x, 0)^T \\ \nabla \times \dot{\mathbf{p}} &= (0, 0, 2\omega)^T \quad \dots\dots\dots (2) \end{aligned}$$

式 (2) より、curl の z 成分は 2ω であり、 θ や r 、 x, y にも依存しない。すなわち、回転中心付近の小さいフローと腕の先端の大きなフローとが同じ特徴量を持つことになる。

一般に回転中心は原点ではなく、画像中のどこにあるのかは未知であるので、次に腕が点 $\mathbf{c} = (c_x, c_y, c_z)^T$ を中心に回転している場合について考える。このときの点 $\mathbf{p}$ の速度 $\dot{\mathbf{p}}$ と curl は以下ようになる。

$$\begin{aligned} \dot{\mathbf{p}} &= \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = \begin{pmatrix} -\omega(y - c_y) \\ \omega(x - c_x) \\ c_z \end{pmatrix} \\ \nabla \times \dot{\mathbf{p}} &= \left(0, 0, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right)^T = (0, 0, 2\omega)^T \quad \dots (3) \end{aligned}$$

この場合も、式 (3) の z 成分は式 (2) と同じであり、同様

に ω のみに依存すると言える。

以上で、腕が画像平面に平行な面上で回転している場合について述べたが、三次元運動への拡張として、以下では画像平面に対して傾いた面上での回転を考える。図 1(b) に示すように、腕の回転面は、画像平面 ($O-xy$) に対して y 軸回りに角 ϕ ($\neq \frac{\pi}{2}$) だけ傾いているとすると、このときの点 p とその速度 $\dot{p}$ は次ようになる。

$$\begin{aligned} p &= (x, y, z)^T \\ &= (r \cos \phi \cos \theta, r \sin \theta, r \sin \phi \cos \theta)^T \\ \dot{p} &= (-\omega r \cos \phi \sin \theta, \omega r \cos \theta, -\omega r \sin \phi \sin \theta)^T \\ &= (-\omega \cos \phi y, \frac{\omega x}{\cos \phi}, -\omega \sin \phi y)^T \end{aligned}$$

回転中心を $c = (c_x, c_y, c_z)^T$ とすれば、 $\dot{p}$ は

$$\dot{p} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -\omega \cos \phi (y - c_y) \\ \frac{\omega (x - c_x)}{\cos \phi} \\ -\omega \sin \phi (y - c_y) \end{pmatrix} \dots\dots\dots (4)$$

となり、その curl の z 成分は次のようになる。

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \omega \frac{1 + \cos^2 \phi}{\cos \phi} \dots\dots\dots (5)$$

ここで、(フレームレート程度の) 微小時間では ϕ は一定であると仮定しているの、式 (5) はやはり ω のみに依存しているとみなせる[†]。

次に図 1(c) のように、腕の回転面が x 軸回りに角 ψ ($\neq \frac{\pi}{2}$) だけ傾いているとすると、次のように式 (4),(5) とよく似た結果が得られる。

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -\omega (y - c_y) \\ \frac{\omega \cos \psi (x - c_x)}{\cos \psi} \\ \omega \sin \psi (x - c_x) \end{pmatrix} \dots\dots\dots (6)$$

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \omega \frac{1 + \cos^2 \psi}{\cos \psi} \dots\dots\dots (7)$$

一般には、回転面は $\phi \neq 0$ かつ $\psi \neq 0$ である。この場合、速度はやや複雑になるが、curl は単純な式になる。

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \omega \frac{\cos^2 \psi + \cos^2 \phi + \sin \phi \sin \psi}{\cos \phi \cos \psi} \dots\dots (8)$$

3章以降の議論では主に ϕ の場合について考えるが、 ψ についても同様であり、4.3 節で両者の関係について述べる。一般的な場合の議論は 6 章で述べる。

〈2・3〉 Curl に基づいた領域分割 式 (5), (7), (8) で示したように、 ϕ 又は ψ が一定という条件で curl の z 成分は ω のみに依存するが、それは速度 $\dot{p}$ の (u, v) 成分であるオブティカルフローから計算することができる。この z 成分は、ある回転している腕の内部では同じ値をとり、それ以外の領域とは異なるので、これを用いて、腕の領域を

[†] 実際、任意の微小な三次元の回転運動は、ある回転軸まわりの回転と見なすことができる⁽²⁵⁾。その回転軸はここでの回転面の法線である。

抽出することができる⁽²⁶⁾。画像中で複数の腕が回転している、かつ接している場合でも、それらの角速度が異なる値である限り、別々の物体として分離することができる。

〈2・4〉 透視投影下の考察 この章の議論は、すべて正射影を仮定していた。これは簡単化のためでもあるが、実際、物体の奥行きがカメラと物体の距離に比べて非常に小さいことを仮定してもよい場合が多いからである。ただし、透視投影下での curl の z 成分は、回転面が画像平面と平行である限り、正射影下のものと同じである。

3. 運動パラメータの推定

前章で、三次元空間内の回転する腕の領域を抽出する方法について述べた。ここでは、その腕の運動パラメータ、つまり回転中心 (c_x, c_y) 、角速度 ω 、傾き角 ϕ 、腕の角度 θ の推定方法について述べる。

実画像から得られるオブティカルフローはノイズを含んでいるため、抽出された領域内のフローを用いて、最小二乗法によりパラメータを推定する。そのためには、領域内の各点 $p_j = (x_j, y_j)^T$ とそのフロー $\dot{p}_j = (u_j, v_j)^T$ について式 (4) を連立して解けばよい。

式 (4) は非線形であるので、このまま連立方程式を解くことは難しいため、以下のように書きなおす。

$$\begin{pmatrix} u_j \\ v_j \end{pmatrix} = \begin{pmatrix} \alpha y_j + \beta \\ \gamma x_j + \delta \end{pmatrix} \dots\dots\dots (9)$$

ここで係数はそれぞれ、 $\alpha = -\omega \cos \phi$ 、 $\beta = \omega c_y \cos \phi$ 、 $\gamma = \frac{\omega}{\cos \phi}$ 、 $\delta = \frac{-\omega c_x}{\cos \phi}$ である。これを行列形式に変形する。

$$\begin{aligned} \dot{p}_j &= A_j q \dots\dots\dots (10) \\ A_j &= \begin{pmatrix} y_j & 1 & 0 & 0 \\ 0 & 0 & x_j & 1 \end{pmatrix}, \quad q = (\alpha, \beta, \gamma, \delta)^T \end{aligned}$$

すなわち、式 (10) のようにオブティカルフローはパラメータ $\alpha, \beta, \gamma, \delta$ の線形式で表すことができる。よって、式 (10) を以下のように連立する。

$$\begin{pmatrix} \dot{p}_1 \\ \dot{p}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \end{pmatrix} q \dots\dots\dots (11)$$

この優決定の連立一次方程式を、QR 分解を用いて最小二乗推定問題として解く^(27, p.914)。これにより得られた q から、運動パラメータは以下のように求められる。

$$c_x = \frac{-\delta}{\gamma}, c_y = \frac{-\beta}{\alpha}, |\omega| = \sqrt{-\alpha\gamma}, \cos \phi = \sqrt{\frac{-\alpha}{\gamma}} \dots\dots (12)$$

ここで ω が解を持つためには、回転運動において $-\alpha\gamma \geq 0$ が満たされなければならない。 ϕ についても、 $0 \leq \frac{-\alpha}{\gamma} \leq 1$ が成立する必要がある。これらの制約と、 ω の符号の決定については 4.2 節で述べる。

以上に述べた手法によって運動パラメータを推定できるが、式 (11) は腕の角度 θ の情報を含まないため、この推定では θ を求めることはできない。したがって、抽出された腕の領域の主軸を θ の推定値として用いることにする。

4. EM アルゴリズムを用いた領域分割と推定

3章では、腕領域内のオプティカルフローを用いた、腕の運動パラメータの推定手法について述べた。そして、その領域はあらかじめオプティカルフローの curl の z 成分を用いて (2章) 抽出してあることが前提であった。つまり、ある領域が得られれば、パラメータ推定を行うことができる。

一方、領域分割のためには精度のよい curl の計算が必要となるが、curl は微分演算なのでノイズに敏感であり、実画像において各点で正確に計算することは困難である。そこで、式 (9) においてオプティカルフローはパラメータ表現されているので、微分演算 ($\frac{\partial v}{\partial x}$, $\frac{\partial u}{\partial y}$) の代わりに、推定されたオプティカルフローの勾配 (γ , α) を用いることができる。

ここで問題となるのは、推定と領域分割の相互依存である。つまり、パラメータ推定は領域分割を必要とするが、領域分割には推定されたパラメータが要求される。この問題を克服するために、本手法では EM アルゴリズム⁽²⁸⁾⁽²⁹⁾を用いる。EM アルゴリズムは反復法の一つであり、初期値として大まかな領域分割を行えばよく、それらの領域に対してパラメータを推定し、再び領域分割を行うという反復処理をすることで相互依存の問題を解決することができる。

また、EM アルゴリズムは隠れ変数を含む正規混合分布の最尤推定に適用できる⁽³⁰⁾。つまり、画像中の複数物体の領域を分割し、かつそれぞれのパラメータを推定することができるということである。複数の運動物体はそれぞれの運動パラメータを持っていて、ある物体の領域内の点は、その物体のパラメータで表されるフローを中心にして分布しているはずである。しかし、各点がどの物体に属しているのかは未知である。ここで、各点がどの物体 (またはそのパラメータ) に属しているかを表す、陽に現れない変数が隠れ変数となり、この隠れ変数の決定に従うことでそれぞれの運動物体に分割することができる。

ここでは、式 (10) における、求められたフロー $\dot{p}_j$ とパラメータ表現された $A_j q$ との誤差の分布は、二次元の正規分布に従うと仮定する。したがって、パラメータ $q = (\alpha, \beta, \gamma, \delta)^T$ が与えられた時に、点 p_j でのフローが $\dot{p}_j$ である条件付確率は以下ようになる。

$$P(\dot{p}_j | p_j, q, \Sigma) = \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\dot{p}_j - A_j q)^T \Sigma^{-1} (\dot{p}_j - A_j q) \right\}$$

ここで Σ は誤差 $\dot{p}_j - A_j q$ の共分散行列であるが、 $u (= \alpha y + \beta)$ と $v (= \gamma x + \delta)$ は別々に推定されるため、それぞれの誤差は互いに独立であると仮定すると、上式は次のように二つの確率密度関数の積で表すことができる。

$$P(p_j | \dot{p}_j, q, \sigma_x^2, \sigma_y^2) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp \left\{ \frac{-(u_j - \alpha y_j - \beta)^2}{2\sigma_x^2} \right\}$$

$$\cdot \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp \left\{ \frac{-(v_j - \gamma x_j - \delta)^2}{2\sigma_y^2} \right\} \cdots \cdots (13)$$

〈4・1〉 反復アルゴリズム 式 (13) で定義された確率密度関数を用いて、腕領域の抽出とそのパラメータを推定するアルゴリズムを以下に示す。

- (1) 各点 $p_j = (x_j, y_j)^T$ ($j = 1, \dots, N$) についてオプティカルフロー $\dot{p}_j = (u_j, v_j)^T$ を計算する。
- (2) 初期値としての大まかな領域分割としてクラスタリングを行ない、クラスタ R_i ($i = 1, \dots$) を得る。
- (3) 重み w_{ij} を、パラメータ $q_i = (\alpha_i, \beta_i, \gamma_i, \delta_i)$ で表されるクラスタ R_i に点 p_j が属する確率を表すとする。この重みの初期値を以下のように決める。

$$w_{ij} = \begin{cases} 1 & (p_j \in R_i) \\ 0 & (p_j \notin R_i) \end{cases}$$

- (4) w_{ij} を、クラスタ R_i についての和が 1 になるように正規化する。

$$\xi_i = \frac{1}{N} \sum_j w_{ij} \text{ として, } w_{ij} \leftarrow \frac{w_{ij}}{\xi_i}$$

- (5) 各クラスタ R_i について、以下の式で表される重みつき二乗誤差を最小にするパラメータ q_i を推定する。

$$\begin{aligned} & \sum_j w_{ij} \|\dot{p}_j - A_j q_i\|^2 \\ & = \sum_j \|(\sqrt{w_{ij}} \dot{p}_j) - (\sqrt{w_{ij}} A_j) q_i\|^2 \end{aligned}$$

すなわち、次の連立一次方程式を QR 分解を用いて解き、 q_i を求める。

$$\begin{pmatrix} \sqrt{w_{i1}} \dot{p}_1 \\ \sqrt{w_{i2}} \dot{p}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sqrt{w_{i1}} A_1 \\ \sqrt{w_{i2}} A_2 \\ \vdots \end{pmatrix} q_i \cdots \cdots (14)$$

- (6) 各クラスタ R_i の重みつき分散 $\sigma_{x_i}^2, \sigma_{y_i}^2$ を計算する。

$$\sigma_{x_i}^2 = \frac{1}{N} \sum_j w_{ij} (u_j - \alpha_i y - \beta_i)^2 \cdots \cdots (15)$$

$$\sigma_{y_i}^2 = \frac{1}{N} \sum_j w_{ij} (v_j - \gamma_i y - \delta_i)^2 \cdots \cdots (16)$$

- (7) 式 (14)~(16) で得られた $q_i, \sigma_{x_i}^2, \sigma_{y_i}^2$ を用いて、次式の事後確率を計算し、重み w_{ij} を更新する。

$$w_{ij} = \frac{\xi_i P(\dot{p}_j | p_j, q_i, \sigma_{x_i}^2, \sigma_{y_i}^2)}{\sum_k \xi_k P(\dot{p}_j | p_j, q_k, \sigma_{x_k}^2, \sigma_{y_k}^2)} \cdots \cdots (17)$$

- (8) もし $\sum_i \|q_i^{new} - q_i^{old}\|^2 > \varepsilon$ ならば、つまり前回の推定値 q_i^{old} と今回の推定値 q_i^{new} の差が大きい値 ε よりも大きければ、ステップ (4) へ。そうでなければ、反復処理を停止する[†]。ステップ (9) へ。

- (9) 領域分割を行う。反復処理を停止した時点で、点 p_j は、各クラスタに対してそれぞれ重み w_{ij} をもっている。その中で最大の重みを求め、点 p_j をそのクラスタ R_{i^*} ($i^* = \arg\max_i w_{ij}$) に分類する。そして画像上で同じクラスタに属する画素を連結し、それぞれの領域に分割する。

[†] EM アルゴリズムの終了条件は本来ならば尤度を用いるべきであるが、非常に収束が遅い⁽²⁸⁾ので、今回はある程度まで推定値が収束した時点で処理を打ち切っている。

〈4・2〉 非回転運動の除去 前節で述べたアルゴリズムによって、各クラスタのパラメータ q_i が求まるので、式 (12) によって c_x, c_y, ω, ϕ が計算できる。3 章で述べたように、 ω は実数であるので、式 (12) から $-\alpha\gamma \geq 0$ が満たされなければならない。しかし、 α と γ はその領域が回転運動しているかどうかにかかわらず別々に推定されるため、条件が満たされない領域も出てくる。そのような領域は、並進や拡大、縮小などの非回転運動に対応するので、これらの運動を分類し、非回転運動の領域を排除することを考える。

表 1 に示すように、 $-\alpha\gamma$ の符号 (つまり α と γ の符号) によって運動を分類することができる。 α と γ が異なる符号を持つとき ($-\alpha\gamma$ が正のとき)、パラメータに対応する運動は回転運動であり、その回転の方向は α ($= -\omega \cos \phi$) からわかる。 α と γ が共に 0 のとき、その領域のオブティカルフローは一定である。すなわち、その領域は並進運動をしているか、または静止している。最後に $-\alpha\gamma \leq 0$ の場合であるが、 α と γ がともに正 (またはどちらか一方が 0 で他方が正) であるときは拡大運動であり、ともに負 (またはどちらか一方が 0 で他方が負) のときは縮小運動である。

〈4・3〉 回転面 (軸) の決定 前節では回転運動の領域を選択する方法を示した。次に、その回転運動が 2.2 節で述べた二つの三次元運動、つまり図 1(b) のように回転面が y 軸回りに角 ϕ だけ傾いている場合と、 x 軸回りに角 ψ だけ傾いている場合 (図 1(c)) のうち、どちらが適切なのかを決定し、選択しなければならない。

3 章で述べたように、式 (12) より、 ϕ について満たされなければならない条件がある。回転運動については $-\alpha\gamma > 0$ なので、 $0 \leq \frac{-\alpha}{\gamma}$ は満たされる。さらに $|\alpha| \leq |\gamma|$ であれば、 $\frac{-\alpha}{\gamma} \leq 1$ となり条件を満たすので、この場合は y 軸回りに ϕ だけ傾いている場合と考えてよい。ここで式 (6) と式 (9) より、 $\cos \psi = \sqrt{\frac{-\gamma}{\alpha}}$ なので、 x 軸回りに ψ だけ傾いている場合、 ϕ とは逆に、 $|\alpha| \geq |\gamma|$ が満たすべき条件となる。

以上をまとめた結果を表 2 に示すが、これを用いることで、どちらの三次元運動を選択するかを決定することができる。すなわち、もし $|\alpha| \leq |\gamma|$ ならば、回転面は画像平面に対して ϕ だけ傾いており、回転面の法線は $(\sin \phi, 0, \cos \phi) = \left(\sqrt{1 - \frac{-\alpha}{\gamma}}, 0, \sqrt{\frac{-\alpha}{\gamma}}\right)$ となる。そうでなければ、 $(0, \sin \psi, \cos \psi)$ である。

5. 実画像を用いた実験結果

本手法を計算機上に実装し、実画像を用いて実験を行った。アルゴリズムのステップ (1) におけるオブティカルフローの計算には、公開されているプログラム⁽³¹⁾⁽³²⁾を用いた。ステップ (2) における初期領域分割には、単純なクラスタリング手法を用いた (フローの方向のヒストグラムを 24 分割し、検出したピークの数だけクラスタを生成

表 1 パラメータの符号による運動の分類

Table 1. Motion classification with the signs of the parameters.

条件	パラメータの符号	対応する運動
$-\alpha\gamma > 0$	$\alpha > 0 \quad \gamma < 0$ $\alpha < 0 \quad \gamma > 0$	回転 ($\omega < 0$) 回転 ($\omega > 0$)
$-\alpha\gamma = 0$	$\alpha = 0 \quad \gamma = 0$	並進
$-\alpha\gamma \leq 0$	$\alpha \geq 0 \quad \gamma \geq 0$ $\alpha \leq 0 \quad \gamma \leq 0$	拡大 縮小

表 2 回転面の決定

Table 2. Decision of rotation plane.

条件	軸	パラメータ
$ \alpha \leq \gamma $	y 軸	$\cos \phi = \sqrt{\frac{-\alpha}{\gamma}}, \psi = 0$
$ \alpha \geq \gamma $	x 軸	$\phi = 0, \cos \psi = \sqrt{\frac{-\gamma}{\alpha}}$

表 3 異なる ϕ に対する推定結果

Table 3. The estimation results with different ϕ .

ϕ	0	10	20	30	[deg]
ϕ の推定値	7.18	7.97	15.8	25.2	[deg]
ω の推定値	0.913	0.939	0.936	0.948	[deg/frame]
$\cos \phi$	1	0.9848	0.9397	0.8660	
$\cos \phi$ の推定値	0.9923	0.9903	0.9624	0.9048	

する)。

〈5・1〉 マニピュレータを用いた推定精度評価 提案手法の推定精度を評価するため、角度を正確に制御できるマニピュレータ (川崎重工 JS005-A, 6 軸, 1 軸のみの回転では $\pm 0.001[\text{deg}]$ の誤差) を用いて実験を行った。角度 ϕ をそれぞれ 0, 10, 20, 30 [deg] に固定し、 θ を 1 度ずらして (つまり角速度を $\omega = 1$ [deg/frame] に設定して) 画像を撮影した。ノイズを抑制するために、同じ角度設定の画像を 10 枚撮影し画素値を平均した画像を処理に用いた (カメラとの距離は約 440cm である)。

パラメータの推定結果を表 3 に示す。 $\omega (=1)$ は 10% 程度の誤差で約 0.9~0.95 と推定された。この誤差の原因には、角速度が小さくなるにつれて最小二乗法による推定精度は悪くなること、また回転中心付近で生じるフローが一画素よりも小さくなるため、うまく計算されないことなどがあげられる。一方 ϕ の推定値は比較的誤差が大きい ($\cos \phi$ は誤差が小さいが) ので、こちらの精度向上は今後の課題である。

〈5・2〉 実画像系列からの腕領域抽出 次に、実際に動作を行っている腕の画像を用いて実験を行った。一般的な屋内環境を想定して、背景には壁などを避けて多数の機材が写るようにしたほかは手を加えず、人物の服装に関しても何も操作をしなかった。動作はカメラから約 3m 離れて行い、毎秒 10 フレームで撮影された。インターレースの影響を排除するために、撮影された画像を間引いて 1/4 のサイズ (340 × 223) に縮小した。

実験で撮影した画像系列の、33 フレーム目の画像と、そのオブティカルフローをそれぞれ図 2(a)(b) に示す。この

図2 実画像に対する実験結果。(a) 第33フレーム原画像。(b) オプティカルフロー。(c)~(h) 反復計算途中の領域分割結果。

Fig. 2. The experimental result using real images.

図3 30~38フレームにおける腕抽出結果。記号 ⊕ は推定された回転中心位置を示す。

Fig. 3. The results of extraction of the arms.

図4 回転中心の推定値の軌跡。原画像（第30フレーム）の肩のあたりに重ねて表示している。

Fig. 4. The trajectory of the estimated rotation center.

図5 角速度と傾き角の推定値の軌跡（太線）。細線は $\alpha\gamma = -\omega^2$ と $\alpha = -\cos^2\phi\gamma$ の曲線。

Fig. 5. The trajectory of the estimated ω and ψ .

図6 (a)(b) 二本の腕が運動している場合 (c)(d) 背後に移動物体が存在する場合

Fig. 6. Other experimental results.

画像系列では、正面を向いた人物が右腕を水平位置から下に振り下ろすという運動を行っており、その際意図的に肘を曲げず肩を動かさないようにした。この運動は途中で停止させることなく、また画像面に対してほぼ平行に行っている。毎秒 10 枚という撮影条件の関係上、動きが速すぎるとフレーム間での対応が取れなくなるため、水平位置から下にきて止まるまでに約 3 秒をかけた。

アルゴリズムの反復によって領域分割結果が変化する様子を見るために、アルゴリズム中の式 (17) によって重みが更新されるたびに、ステップ (9) によって領域分割を行った (図 2(d)~(h))。この図では異なるクラスタ毎に色分けし、輝度は重みに比例するように表示しているため、暗ければ暗いほど、クラスタに属する確率が小さいことを表している。

初期分割 (図 2(c)) では、上腕と下腕が別々の領域に分かれてしまっているが、数回の反復によって、肩から手首まで腕全体が一つの領域に覆われていることがわかる。領域分割は最初の数回の反復で大きく変化し、反復が 10 回を超えたとほとんど変化しなくなり (図 2(g),(h))、パラメータ推定の方もほぼ収束する。

反復 20 回目 (図 2(h)) では、クラスタ数が物体の数よりも多いので、背景が複数のクラスタに分割されている。そのため、背景がすべて黒くなっている反復 1 回目 (図 2(d)) のほうがよい分割結果であるように見えるかもしれない。しかし、前述したように、クラスタに属する確率が低いほど暗く表示しているため、1 回目の結果の背景部分の暗い画素は、背景に (そしてどのクラスタにも) 属する確率が低く、信頼できない結果であることを示している。反復が進むにつれて、それぞれのクラスタに属する確率が高く (したがって明るい表示に) なっていくので、過分割になってはいるが反復 20 回目の結果の方が良い結果であると言える。このように過分割とならないために AIC などを用いてクラスタ数を決定する⁽²⁸⁾ ことについては今後の課題である。

画像中で最も大きい角速度を持つ領域が腕の領域であると仮定して、同じ画像系列の 30~38 フレームにおいて腕領域を抽出した結果を図 3 に示す。この図から、実画像から得られるノイズが含まれたオプティカルフローを用いて、本手法は安定して腕領域を抽出できることがわかる。この図ではそれぞれのフレームで $\oplus$ によって推定された回転中心を表示しているが、それらを一つの画像に重ねて表示した軌跡を図 4 に示す。回転中心は次第に下がっていき、系列の後半ではやや肩の下辺りで留まっているのがわかる。他の推定されたパラメータの軌跡を図 5 に示す。このグラフは、パラメータ ω と $\cos \phi$ が式 (12) において α と γ で表されているので、それらを α - γ 空間にプロットしたものである。腕は振り下ろされるにしたがって遅く (ω が小さく) なっているが、 ψ はそれほど大きく変化していないことがわかる。

他の画像を用いた実験結果を図 6 に示す。複数の腕が同

時に運動している場合の実験 (図 6(a)) は、左側の人物が左腕を振り下ろし、右側の人物は右腕を振り上げている、という画像系列を用いた。実験環境は、背景が異なるほかは図 2 の実験と同じである。ここでは二つの運動が存在するが、表 1 の回転運動であるという制約条件を満たし、かつ角速度がしきい値 (3 deg/s) 以上である領域を抽出することで、それぞれの腕を抽出することができた (図 6(b))。

図 6(c) は、背景に移動物体が存在する場合の実験である。図 2 の実験と同様にある人物が左腕を振り下ろしており、その背後約 1m のところを、別の人物が右から左に移動しているという画像系列を使用した。後ろの人物が背後を通過するのに合わせて手前の人物が腕を動かしており、両者の領域は接しているが、歩いている人物の領域は回転運動を含まないため、4.2 節で述べた条件によって除去され、腕の領域だけが抽出された (図 6(d))。

6. おわりに

本論文では、オプティカルフローの curl に基づいた、回転する腕の領域抽出と、その運動パラメータを推定する手法を提案した。提案手法は、オプティカルフローの勾配による領域分割と運動パラメータの最小二乗法による推定からなり、EM アルゴリズムを用いて推定と領域分割を同時に行う。そして推定されたパラメータに基づいて、画像内に存在する並進運動と回転運動を区別することができ、背景で移動する物体を除去することができる。提案手法を実装し実画像に対して実験を行い、パラメータ推定には多少の誤差を含むが、実画像から得られたノイズを含むオプティカルフローからでも良好に腕の領域を抽出することができることを確認した。また複数の腕が同時に動作している場合や、背後に移動物体が存在する場合についても実験を行い、実環境における本手法の有効性を示した。

本手法はオプティカルフローが正確に計算されることが必要であるので、対象が画像中で小さくなりすぎるとそのフローがノイズに埋もれてしまい、抽出が困難になる。また本手法は正射影を仮定しているが、これは透視投影の近似であり、ノイズのある実画像のフローに対してはこの近似はそれほど影響はないと思われる。

本手法は ϕ と ψ の一方が 0 である場合のみを扱っているが、一般的な三次元運動 ($\phi \neq 0, \psi \neq 0$) については、式 (9) の代わりに、 u と v を $\alpha x + \beta y + \gamma$ という平面を用いて近似することで、3 章の手法が適用できるように拡張することができる。また、本手法はカメラが移動したときや肘などの関節がある場合を考慮していない。これは式 (1) で述べた定数ベクトル t が存在する場合であり、回転中心などの推定される運動パラメータに影響を与える⁽³³⁾。これについては、 t の時間的な変化に (線形であるなどの) 仮定を設けることで、パラメータを補正することができる。現在これらについての検討を行っている。

謝辞 本研究の一部は (財) 西秩奨学会の助成による。

(平成 12 年 3 月 24 日受付, 同 12 年 7 月 5 日再受付)

文 献

- (1) L. Goncalves, E. D. Bernardo, P. Perona : "Reach out and touch space (Motion learning)", *Proc. of FG'98*, pp.234-239 (1998).
- (2) 木本伊彦, 梶尾昭彦, 安田靖彦:「スティックモデルに基づく単眼動画からの人体歩行運動の解析の一手法」, 信学論 D-II, Vol.J74-D-II, No.3, pp.376-387 (1991).
- (3) 戸川達男:「運動計測技術の展望」, 計測と制御, Vol.36, No.9, pp.605-608 (1997).
- (4) 岩田洋夫:「バーチャルリアリティへの応用」, 計測と制御, Vol.36, No.9, pp.639-642 (1997).
- (5) J. Yamato, J. Ohta, K. Ishii : "Recognizing human action in time-sequential images using hidden markov model", *Proc. of CVPR'92*, pp.379-385 (1992).
- (6) 大和教司, 倉掛正治, 伴野明, 石井健一郎:「カテゴリ別 VQ を用いた HMM による動作認識法」, 信学論 D-II, Vol.J77-D-II, No.7, pp.1333-1318 (1994).
- (7) 高橋勝彦, 岡進, 小島浩, 岡隆一:「ジェスチャー動画のスポットティング認識」, 信学論 D-II, Vol.J-77-D-II, No.8, pp.1552-1561 (1994).
- (8) T. Nishimura, T. Mukai, R. Oka : "Non-monotonic continuous dynamic programming for spotting recognition of hesitated gestures from time-varying images", *Proc. of ACCV'98*, pp.734-739 (1998).
- (9) D. D. Morris, J. M. Rehg : "Singularity analysis for articulated object tracking", *Proc. of CVPR'98*, pp.289-296 (1998).
- (10) Y. Guo, G. Xu, S. Tsuji : "Understanding human motion patterns", *Proc. of ICPR'94*, pp.325-329 (1994).
- (11) 石井浩史, 望月研二, 岸野文郎:「人物像合成のためのステレオ画像からの動作認識法」, 信学論 D-II, Vol.J76-D-II, No.8, pp.1805-1812 (1993).
- (12) S. X. Ju, M. J. Black, Y. Yacoob : "Cardboard people: A parameterized model of articulated image motion", *Proc. of FG'96*, pp.38-44 (1996).
- (13) I. Haritaoglu, D. Harwood, L. S. Davis : "W⁴: Who? When? Where? What? A real time system for detecting and tracking people", *Proc. of FG'98*, pp.222-227 (1998).
- (14) D. M. Gavrila, L. S. Davis : "3-D model-based tracking of humans in action: a multi-view approach", *Proc. of ICPR'96*, pp.73-80 (1996).
- (15) J. Ohya, F. Kishino : "Human posture estimation from multiple images using genetic algorithm", *Proc. of ICPR'94*, pp.750-753 (1994).
- (16) 大谷淳, 岸野文郎:「遺伝的アルゴリズムを用いたマルチ画像からの人物の姿勢検出の検討 —— 上半身の関節角度検出 ——」, 信学技報 PRU93-12, pp.47-54 (1994).
- (17) 山本正信, 川田聡, 近藤拓也, 川越和忠:「ロボットモデルに基づく人物動作の3次元動画追跡」, 信学論 D-II, Vol.J79-D-II, No.1, pp.71-83 (1996).
- (18) 亀田能成, 美濃秀彦, 池田克夫:「シルエット画像からの関節物体の姿勢推定法」, 信学論 D-II, Vol.J79-D-II, No.1, pp.26-35 (1996).
- (19) M. Minoh : "3D model centered framework for CV and VR", *Proc. of ACCV'98*, pp.332-339 (1998).
- (20) 六日市正弘, 浜田望:「エッジにおける速度拘束式とエッジ連結を考慮したオプティカルフロー推定」, 信学論 D-II, Vol.J82-D-II, No.3, pp.407-414 (1999).
- (21) 岡田隆三, 白井良明, 三浦純, 久野義徳:「オプティカルフローと距離情報の統合による3次元運動する人間の追跡」, 信学論 D-II, Vol.J82-D-II, No.8, pp.1252-1261 (1999).
- (22) 大田佳人, 山際貴志, 山本正信:「キーフレーム拘束を利用した単眼動画からの人間動作の追跡」, 信学論 D-II, Vol.J81-D-II, No.9, pp.2008-2018 (1998).
- (23) D. M. Gavrila : "The visual analysis of human movement: A survey", *Computer Vision and Image Understanding*, Vol.73, No.1, pp.82-98 (1999).
- (24) 小西栄一, 深見哲造, 遠藤静男: 線形代数・ベクトル解析 (1978) 培風館.
- (25) John J. Craig : ロボティクス —— 機構・力学・制御 —— (1991) 共立出版.
- (26) 玉木徹, 山村毅, 大西昇:「人物モデル当てはめによる人物の位置決め」, 第57回情報学全大, 第2巻, pp.142-143 (1998).
- (27) ——: 現代数理科学辞典 (1991) 大阪書籍.
- (28) 栗田多喜夫, 赤穂昭太郎:「画像理解における統計的手法」, 松山隆

司 久野義徳 井宮淳 (編), コンピュータビジョン: 技術評論と将来展望, 第13章, pp.195-206 (1998) 新技術コミュニケーションズ.

- (29) A. P. Dempster, N. M. Laird, D. B. Rubin : "Maximum likelihood from incomplete data via the EM Algorithm", *J. Roy. Statist. Soc. B*, Vol. 39, pp.1-38 (1977).
- (30) 大羽成征, 石井信:「正規混合分布の推定によるクラスタリングの種々手法の比較」, 信学技報 NC99-12, pp.109-116 (1999).
- (31) N. Ohta : <http://www.ail.cs.gunma-u.ac.jp/Labo/Program/Flow-R00.tar.gz>.
- (32) N. Ohta : "Image movement detection with reliability indices", *IEICE Transactions*, Vol.E74, No.10, pp.3379-3388 (1991).
- (33) 朴基守, 増田基, 伊東敏夫, 金田悠紀夫:「アフィン変換導入によるオプティカルフロー推定の一手法」, 電学論 C, Vol.119-C, No.12, pp.1468-1475 (1999).

玉 木 徹 (非会員) 平8名大・工・情報卒。平10同大学院工学研究科情報工学専攻博士課程前期修了。現在, 同博士課程後期に在学中。平成11年度理化学研究所ジュニア・リサーチ・アソシエイト。平成11年度電子情報通信学会東海支部学生研究奨励賞。画像認識, コンピュータビジョンに興味を持つ。電子情報通信学会, 映像情報メディア学会各学生会員。

山 村 毅 (非会員) 昭62名大・工・電気卒。平1同大学院情報工学専攻博士(前期)課程修了。平4同博士(後期)課程単位取得退学。同年名古屋大学工学部助手。平成7年同講師。平成10年愛知県立大学情報科学部助教授。博士(工学)。自然言語処理, 視覚情報処理の研究に従事。電子情報通信学会, 情報処理学会, IEEE 各会員。

大 西 昇 (正員) 昭48名大・工・電気卒。昭50同大学院電気系専攻修士課程修了。同年労働福祉事業団労働リハビリテーション工学センター研究員。昭61名古屋大学工学部講師。平1同助教授。平5年10月より理化学研究所バイオメトリックコントロール研究センターのチームリーダー兼任。平6同教授で情報工学科に所属。平12名古屋大学情報メディア教育センター教授。生体工学, 福祉工学, 人工知能などの研究・教育に従事。工学博士。1996年度計測自動制御学会賞(技術賞)。電子情報通信学会, 情報処理学会, 計測自動制御学会, ロボット学会, 日本神経回路学会, 映像情報メディア学会, IEEE 等各会員。