

動画像からの3Dコンピューティング

金谷 健一

群馬大学工学部情報工学科

画像から3次元運動を計算するアルゴリズムに関する最新の成果を述べる。まず2画像間に特徴点の対応関係が与えられている場合の特徴点までの距離とカメラの運動を計算するロバストな算法を記述し、基本行列の分解可能性、解の一意性、直接的最適化法、臨界曲面について述べる。平面運動は別個に扱う。次にオブティカルフローの3次元解釈のロバストな算法を記述し、基本パラメータの分解可能性、解の一意性、直接的最適化法、臨界曲面、および有限運動のアルゴリズムとの関係について述べる。平面オブティカルフローは別個に扱う。

3-D Computing from Motion

Kenichi Kanatani

Department of Computer Science
Gunma University, Kiryu, Gunma 376, Japan

State-of-the-art algorithms for computing 3-D motion from images are reviewed. First, a noise robust algorithm is described computing the depths of the feature points and the camera motion from correspondences of feature points between two images. Decomposability of the essential matrix, uniqueness of the solution, direct optimization, and the “critical surface” are also discussed. Planar surface motion is treated separately. Then, a noise robust algorithm is described for 3-D interpretation of optical flow. Decomposability of the essential parameters, uniqueness of the solution, direct optimization, the “critical surface”, and relationships to the algorithm for finite motion are also discussed. Planar surface optical flow is treated separately.

1. 序論

物体またはカメラが3次元剛体運動をするとき、得られた動画像から物体の形状および(物体またはカメラの)3次元運動を計算する問題は shape from motion または structure from motion と呼ばれ、Ullman[44]によって基本的な定式化がなされて以来、コンピュータビジョンの重要な問題となっている[21]。これは入力された動画像を単に解析するだけでなく、解析が容易になるようにカメラを意図的に移動すること(アクティブビジョン[1])により、自律ロボットの環境理解の手段としても利用できるからである。

有限運動の場合、2画像間で8個の特徴点の対応が与えられていれば解が解析的に計算できることは Longuet-Higgins[23] および Tsai and Huang[42]によって独立に発見された。Longuet-Higginsの方法は初等的なベクトル解析のみを用いているが Tsai and Huangの方法は特異値分解を含んだ複雑なものであった。その後これらはいろいろ変形され、総称して8点アルゴリズムと呼ばれている。しかし、これらはすべてデータに誤差がないときに成立する理論的な関係に立脚しているため、それを満たしていない不正確なデータを用いると矛盾が生じる。矛盾の生じない“ロバストな”解法は Wengら[46]によって提案された。第2節では Wengら[46]の解法を洗練した算法を示す。物体が平面の場合は第3節で述べる。

物体またはカメラの運動が微小であれば、画像上での対応関係は連続した流れを定義する。これをオプティカルフローと呼ぶ。これは理論的には瞬間的運動の速度場として取り扱える。運動が小さいほど画像の誤差の影響が相対的に大きくなるので、解の信頼性が有限運動に比べて低下する。しかし、有限運動に対して特徴点を区別して追跡することは高度な推論を要する困難な問題であるのに対して、オプティカルフローは全画素に対する一様な画像処理によって計算できるという長所がある。

ただしフローの精度は画像の濃淡パターンに大きく依存する(濃淡変化が大きいほど検出が容易であり、濃淡がほぼ一定な領域では困難)。また何らかの拘束条件(フローの滑らかさ等)が必要で、精度は拘束条件にも依存する。しかし、検出したフローの信頼性を適切に評価し、その信頼性を重みとして画像全体に対する最適化手法を適用すれば信頼できる解が得られると期待できる。数年前まではオプティカルフローの検出は時間を要する困難な問題と思われていたが、今日ではワークステーションの普及によって計算能力が高まり、オプティカルフロー検出はコンピュータビジョンの最も基本的な処理の一つとみなされつつある。

オプティカルフローの3次元解釈は古くから知覚心理学の問題であった。実際、“オプティカルフロー”という言葉は心理学者 Gibson[6]によって使われ始めた。このため、オプティカルフローの初期の研究はほとんど知覚心理学の立場からのもので、工学的観点からのものは少なかった。工学的観点からまず試みられたの

はカメラの並進成分と回転成分の分離である。フローの導関数を用いれば解が得られるが、導関数の計算は誤差に弱いので実際的ではない。カメラが回転すればシーンの遠近に関係しない滑らかなフローが得られるのに対し、並進すればフローの大きさはシーンの遠近に依存する(運動視差; motion parallax)。したがってフローに不連続があれば、それは遠近の不連続と解釈でき、これから並進成分が推定できる。あるいはカメラの回転速度を、それによる回転成分を差し引いたフローがある一点(出現点; focus of expansion)から発散しているように探索することも考えられた。また、並進速度がわかれば遠近が計算できるから、回転速度と並進速度を理論的に導かれるフローと検出したフローとが一致するように探索することも試みられた。いずれにしても、何らかの探索が必要であった[8, 14]。探索を要しない解析的解法は田川ら[38]によって提案された。第4節ではこれを洗練した算法を示す。物体が平面の場合は第5節で述べる。

2. 有限運動の3次元解釈

空間中の物体の特徴点を $\{P_\alpha\}$, $\alpha=1, \dots, N$ とする。カメラのレンズの中心(視点)を始点とし、それらを指す単位ベクトルを $\{m_\alpha\}$ とする。これら特徴点の N ベクトル[19, 20]と呼ぶ。 N ベクトルはカメラの構造が既知であれば画像から計算できる。以下、カメラが物体に対して運動する場合を考える。カメラの運動を視点の周りの回転 R と並進 h とで指定し、 $\{R, h\}$ を運動パラメータと呼ぶ。視点から各特徴点までの距離を運動前後についてそれぞれ r_α, r'_α とする。

カメラの運動によって各 N ベクトル m_α が m'_α に変化するとき(図1(a)), これを最初のカメラ系からみれば第2のカメラ系は R だけ回転しているから Rm'_α に等しい。したがって図1(b)の関係から、問題は「式

$$r_\alpha m_\alpha - r'_\alpha Rm'_\alpha = h \quad (1)$$

の関係を満たす運動パラメータ $\{R, h\}$ と距離 r_α, r'_α を求めよ」と述べるができる。

これから、並進 h と距離 r_α, r'_α には“スケール不定性”があることがわかる。すなわち $\{R, h\}$ と r_α, r'_α が解なら任意の零でない定数 k に対して $\{R, kh\}$ と kr_α, kr'_α も解である。これは遠くの大きな物体に対する大きな運動と近くの小さな物体に対する小さな運動とが区別できないことを意味する。観測データが N ベクトル(=視線の方向)という無次元量のみである以上、長さの次元を含む量を導けないのは当然である。

並進が $h=0$ であるかどうかの判定は容易であるから[22], $h \neq 0$ と仮定し、 $\|h\|=1$ なるスケールを採用する($h=0$ なら何らの3次元情報も得られない[18])。図1(b)の関係から解が存在する必要十分条件はベクトル h, m_α, Rm'_α が“共面”であること、すなわちそれらのスカラー三重積が零となることである。

$$|h, m_\alpha, Rm'_\alpha| = 0 \quad (2)$$

ここで行列

$$G = h \times R \quad (3)$$

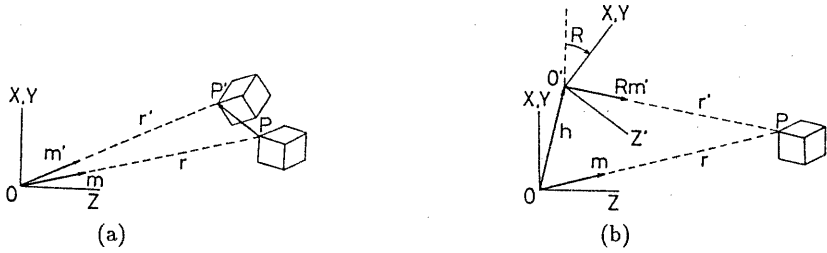


図 1: (a) N ベクトルの対応 (カメラ系からみた記述). (b) 運動パラメータ $\{R, h\}$ (物体からみた記述).

を定義する (右辺は h と R の各列とのベクトル積). すると式 (2) は

$$(m_\alpha, Gm'_\alpha) = 0 \quad (4)$$

と書ける ($(\cdot, \cdot)$ はベクトルの内積). G を基本行列, 式 (2) および (4) をエビ極線方程式と呼ぶ. G の定義より $\|G\| = \sqrt{2}$ であることが直ちにわかる ($\|G\|$ は各要素の二乗和の平方根). これから, 解は次の方針で求まる.

1. 観測した N ベクトル $\{m_\alpha\}, \{m'_\alpha\}, \alpha=1, \dots, N$ に対してエビ極線方程式 (4) が満たされるように基本行列 $G, \|G\| = \sqrt{2}$ を定める.
2. 得られた基本行列 G を式 (3) のように単位ベクトル h と回転行列 R に分解する.
3. 各特徴点までの距離を次のように定める.

$$r_\alpha = \frac{(h, m_\alpha) - (m_\alpha, Rm'_\alpha)(h, Rm'_\alpha)}{1 - (m_\alpha, Rm'_\alpha)^2},$$

$$r'_\alpha = \frac{(m_\alpha, Rm'_\alpha)(h, m_\alpha) - (h, Rm'_\alpha)}{1 - (m_\alpha, Rm'_\alpha)^2} \quad (5)$$

基本行列 G には 9 個の要素があるが, 制約 $\|G\| = \sqrt{2}$ があるから, ステップ 1 では少なくとも 8 個の特徴点が観測されれば一般には G が一意に定まる. これが “8 点アルゴリズム” の名称の由来である. 特徴点の個数が 8 個以上の場合には最小二乗法により固有値問題に帰着させて解析的にかつロバストに計算できる (付録 A). ステップ 2 は “特異値分解”, “極分解”, “四元数” などを用いて解析的にかつロバストに計算できる (付録 B). 式 (5) は式 (1) から最小二乗法によって導かれたものであり, これを運動差方程式と呼ぶ.

運動パラメータ $\{R, h\}$ に含まれる未知数は R (回転行列) に 3 個 [18], h (単位ベクトル) に 2 個の合計 5 個であるから, 本来は 5 個の特徴点で十分なはずである (ただし複数の解が得られる [5]). この食い違いはすべての行列が式 (3) のように分解できるとは限らないことに起因する. そのように分解できる行列を分解可能行列 [11] と呼ぶ. 行列 G が分解可能である必要十分条件はその特異値が 1, 1, 0 となることである [11]. これは次式と同値である.

$$\det G = 0, \quad \|G\| = \|GG^T\| = \sqrt{2} \quad (6)$$

ステップ 1 では厳密にはこの 3 個の分解可能条件のもとで計算すべきであるが, これを無視しているのは計算の簡化化のためである. ただし, データ $\{m_\alpha\}, \{m'_\alpha\}$ が厳密であればそれを無視しても, 得られる解はそれを自動的に満たす. 分解可能条件 (6) を課せば解はもはや解析的には求まらず, 数値的探索 [35, 36] あるいは数式処理 [5, 13] が必要になる. これを直接的探索解法と呼ぶ. これにより解が定まる特徴点の個数の最小値は 5 個である. ただし, 数値的探索で極小値に落ち込まずに最適解に到達するようにするのは複雑な技術を要する. 画像に誤差があれば上の方法による解析解にせよ直接的探索解法にせよ, 得られた解には統計的偏差が存在する [31]. しかし, この偏差はくりこみ法と呼ぶ手法によって取り除くことができる [39].

ステップ 1 で基本行列 G の符号は不定であるから, ステップ 2 では 2 組の解が得られる. しかし, 並進 h の符号が互いに反転しているから, 式 (5) の運動視差方程式から正の距離 r_α が得られるものを選べばよい. したがって基本行列 G が符号を除いて一意に定まりさえすれば解は一意的である.

基本行列 G が定まらない唯一の例外はすべての特徴点が特殊な配置をしていて, ある $\tilde{G} (\neq G)$ に対して “偽のエビ極線方程式” $(m_\alpha, Gm'_\alpha) = 0$ が恒等的に成立することである. そのような配置はカメラの移動前後の二つの視点を通るある 2 次曲面であることが容易に導ける [24]. これを弱臨界面という. このとき上の解法では基本行列を定めることができないが, 偽の基本行列 $\tilde{G}$ が分解可能でなければ, 分解可能条件 (6) を課すことによって G を定めることができる. あるいは直接的探索解法によっても解が求まる. しかし $\tilde{G}$ が分解可能であれば真の解と区別することは原理的に不可能である. そのような曲面を (強) 臨界面という. これは 2 次線織面であり, 連結双曲面かその縮退であることが知られている [10, 27, 29, 30, 33, 34].

3. 平面運動の 3 次元解釈

平面は縮退した双曲面でもあり, 任意の平面は (強) 臨界面でもある [10, 27, 29, 30, 33, 34]. したがって, すべての特徴点が同一平面上にあればその 3 次元解釈には必然的に多義性がある. これは知覚心理学では古くから知られていた事実でもある [7]. 解析的な解法も

一般の運動とは別個に導かれている [40, 41, 43, 26].

視点を通らない平面の(外向き)単位法線ベクトルを n とし, 視点から平面までの距離を d とするとき, ベクトル $p=n/d$ を P ベクトルと呼ぶ. これは向きが法線方向, 長さが距離の逆数であるベクトルである. P ベクトルを用いれば平面の方程式は $(p, r)=1$ と書け, 平面の向きと位置が一意的に定まる.

物体が平面であればその運動を画像面上に投影したものは画像面に射影変換(相称変換)を引き起こす. 2次元射影変換は N ベクトル(=同次座標)を用いればある正則行列 A により

$$m = \pm N[A m] \quad (7)$$

の形に表せる [19] ($N[\cdot]$ は単位ベクトルへの正規化). A は定数倍の不定性があるから, $\det A=1$ と正規化する. この射影変換の行列 A は運動パラメータ $\{R, h\}$ と P ベクトル p によって

$$A = \frac{1}{k}(I - ph^T)R, \quad k = \sqrt{1 - (p, h)} \quad (8)$$

と表せる [19]. これから, 解は次の方針で求まる.

1. 観測した N ベクトル $\{m_\alpha\}, \{m'_\alpha\}, \alpha=1, \dots, N$ に対して式 (7) を満たす射影変換の行列 $A, \det A=1$ を求める.
2. 得られた行列 A を式 (8) のように運動パラメータ $\{R, h\}$ と P ベクトル p に分解する.

行列 A には 9 個の要素があるが, 正規化条件 $\det A=1$ があるから, ステップ 1 では少なくとも 4 個の特徴点の変位の x, y 成分が与えられれば A が一意的に定まる. 特徴点の個数が 4 個以上の場合是最小二乗法により固有値問題に帰着させて解析的にかつロバストに計算できる(付録 C). ステップ 2 は行列 A の“特異値分解”を用いて解析的に計算できる(付録 D).

式 (8) から直ちにわかるように, $\{h, p\}$ が解であれば $\{-h, -p\}$ も解である. この二つの解に対応する平面は視点に関して互いに鏡像の位置にある. そして, それぞれに対して運動後にそれを表, 裏の両方からながめている二つの解釈が求まり, 解は合計 4 個ある. しかし, もとの対応点の位置関係から面の表裏は定まるので, 結局解は 2 個となる. 並進方向が平面に直交するときは解は一意的である [26].

4. オプティカルフローの 3 次元解釈

レンズの中心(視点)が並進速度 v で移動し, かつカメラがその周りに回転速度 ω で回転する瞬間的運動を考える. ある特徴点の N ベクトル m はカメラ系からみると速度 $\dot{m}$ で変化する. これを N 速度と呼ぶ. N 速度 $\dot{m}$ を N ベクトル m の関数とみなしたものをオプティカルフローと呼ぶ [19]. 視点 O からオプティカルフローの定義されている領域 S を見込む立体角 $\Omega(S)$ を視野と呼ぶ(図 2(a)).

有限運動の場合と同様に $\{\omega, v\}$ を運動パラメータと呼び, 視点 O から物体面までの距離を N ベクトル m の関数とみなした $r(m)$ を距離写像と呼ぶ(図 2(b)).

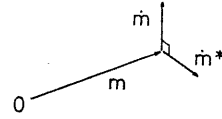


図 3: ねじれフロー

有限運動と同様にオプティカルフローの 3 次元解釈にはスケール不定性があり, $\{\omega, v\}, r(m)$ が解なら任意の零でない定数 k に対して $\{\omega, kv\}, kr(m)$ も解である. これはオプティカルフロー(=視線の変化速度)の単位が $1/\text{sec}$ であり, 長さの次元を含んでいないことから当然である. 並進速度が $v=0$ であるかどうかの判定は容易であるから [22], $v \neq 0$ と仮定し, $\|\omega\|=1$ なるスケールを採用する ($v=0$ なら何らの 3 次元情報も得られない [18]).

解析解を得るためには N 速度 $\dot{m}$ 自身でなく, これを N ベクトル m の周りに右ネジ方向に 90 度回転した $\dot{m}^*$ を考えると都合がよい. これをねじれフローと呼ぶ. さらに, 運動パラメータ $\{\omega, v\}$ 自身でなく $\{K, v\}$ を考えると都合がよい. ただし K は次のように定義した対称行列である.

$$K = (\omega, v)I - \frac{1}{2}(\omega v^T + v \omega^T) \quad (9)$$

$\{K, v\}$ を基本パラメータと呼ぶ.

ねじれフロー $\dot{m}^*$ が与えられたとき, それを生じるような距離写像 $r(m)$ が存在する必要十分条件はすべての $m \in \Omega(S)$ に対して次の式を満たす基本パラメータが存在することである [22].

$$(\dot{m}^*, v) + (m, Km) = 0 \quad (10)$$

これをエビ極線方程式と呼ぶ. これから, 解は次の方針で求まる.

1. 観測したオプティカルフロー $\dot{m}$ からねじれフロー $\dot{m}^*$ を計算し, 式 (10) のエビ極線方程式が満たされるように基本パラメータ $\{K, v\}$ を定める.
2. 得られた基本パラメータ $\{K, v\}$ を分解して, 式 (9) を満たす運動パラメータ $\{\omega, v\}$ を計算する.
3. 距離写像 $r(m)$ を次式によって計算する.

$$r(m) = \frac{1 - (m, v)^2}{|m, \omega, v| - (\dot{m}, v)} \quad (11)$$

ステップ 1, 2 は最小二乗法により固有値問題に帰着させて解析的にかつロバストに計算できる(付録 E). 式 (11) も最小二乗法によって得られるものであり [19], 有限運動の場合と同様に運動視差方程式と呼ぶ.

任意の対称行列 K と単位ベクトル v が与えられたとき式 (9) を満たすベクトル ω が存在するとは限らない. 存在するとき $\{K, v\}$ は分解可能であるという. $\{K, v\}$ が分解可能である必要十分条件は次のように表せる [22].

$$(v, Kv) = 0 \quad (12)$$

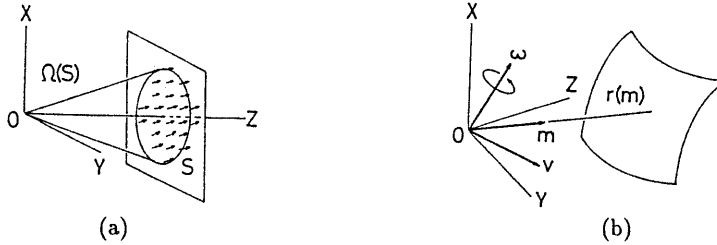


図 2: (a) オプティカルフロー (カメラ系からみた記述). (b) 運動パラメータ $\{\omega, v\}$ (物体からみた記述).

$$K = \frac{1}{2} \text{tr} K (I - vv^T) + Kvv^T + vv^T K \quad (13)$$

ステップ 1 ではこの分解可能条件のもとで計算すべきであるが、これを無視しているのは計算の簡単化のためである。フロー m が厳密であればそれを無視しても、得られる解はそれを自動的に満たす。分解可能性を課せば解はもはや解析的には求まらず、数値的な直接的探索解法が必要となるが、極小値に落ち込まずに最適解に到達させるには複雑な技術を要する [14]。

画像に誤差があれば解析解にせよ直接的探索解法にせよ、得られた解には統計的偏差が存在する [38]。しかし、この偏差はくりこみ法と呼ぶ手法によって取り除くことができる [12]。

対称行列 K には独立な要素が 6 個あり、 v は単位ベクトルであるから基本パラメータ $\{K, v\}$ は 8 個の未知数を含んでいる。したがって、上の解法は 8 画素だけのフローにも適用できる (“8 点アルゴリズム”)。分解可能条件 (12), (13) を課せば 5 画素でもよい。もちろんこれは理論上の話で、フローの精度を考えると非現実的である。信頼性の高い解を得るには視野 $\Omega(S)$ を広くとり、検出したフローの信頼性を適切に評価した重みつき総和を用いることが必要である (付録 B)。

エビ極線方程式 (10) からわかるように、基本パラメータ $\{K, v\}$ の符号は不定である。このため回転速度 ω は一意的に定まるが並進速度 v の符号が不定となる。しかし、運動視差方程式 (11) からわかるように v の符号を反転すると距離写像 $r(m)$ の符号も反転するので、これが正になるように正しい符号を選ぶことができる。したがって、基本パラメータ $\{K, v\}$ が符号を除いて一意的に定まりさえすれば解は一意的である。

基本パラメータ $\{K, v\}$ が定まらない唯一の例外は物体面が特殊な形状をしていて、ある $\{\tilde{K}, \tilde{v}\} (\neq \{K, v\})$ に対して “偽のエビ極線方程式” $(\tilde{m}^*, \tilde{v}) + (m, \tilde{K}m) = 0$ が恒等的に成立することである。そのような形状はカメラの視点を通るある 2 次曲面であることが容易に導ける。これを弱臨界面面という。このとき上の解法では基本パラメータを定めることができないが、偽の基本パラメータ $\{\tilde{K}, \tilde{v}\}$ が分解可能でなければ、分解可能条件 (12), (13) を課すことによって $\{K, v\}$ を定めることができる。あるいは直接的探索解法によっても解が求まる。しかし $\{\tilde{K}, \tilde{v}\}$ が分解可能であれば真の解と区別することは原理的に不可能である。その

ような曲面を (強) 臨界面面という。これは 2 次線幾面であり、連結双曲面かその縮退であることが知られている [9, 28, 32]。

本節の記述の流れは第 2 節の記述とほとんど同じである。これはオプティカルフローが有限運動の瞬間的極限であることから当然である。実際に確かめてみると、オプティカルフローのエビ極線方程式 (10) は有限運動のエビ極線方程式 (4) の微小運動の極限であることがわかる [22]。また、オプティカルフローの基本パラメータ $\{K, v\}$ は有限運動の基本行列 G の微小運動の極限であり、式 (3) の極限が式 (9) になる [22]。したがってオプティカルフローに関する分解可能条件 (12), (13) は有限運動に関する分解可能条件 (6) に対応している。さらにオプティカルフローに関する弱臨界面面および (強) 臨界面面はそれぞれ有限運動に関する弱臨界面面および (強) 臨界面面の瞬間的極限として得られることもわかる [22]。

5. 平面オプティカルフローの 3 次元解釈

平面は縮退した双曲面であり、有限運動の場合と同様に任意の平面は (強) 臨界面面でもあることが導けるから、物体面が平面であればその 3 次元解釈には必然的に多義性がある。解析的な解法は Longuet-Higgins [25], Subbarao and Waxman [37], Kanatani [17] によって独立に提案された。Longuet-Higgins [25] の解法は行列の固有値の解析に基づくものであり、Subbarao and Waxman [37] の解法は方程式の直接的消去操作によるものであり、Kanatani [17] の解法は画像面の変換の群の既約表現から導かれる不変量を用いるものである。以下では Longuet-Higgins [25] の解法を変形したものを示す [19]。

式 (7) の微小運動の極限をとると次式を得る。

$$\dot{m} = W^T m - (m, W^T m) m \quad (14)$$

ただし $W = A$ である。正規化条件 $\det A = 1$ は $\text{tr} W = 0$ となる。行列 W をフロー行列と呼ぶ。式 (8) の微小運動の極限をとれば、フロー行列 W を瞬間的運動パラメータ $\{\omega, v\}$ と平面の P ベクトル p によって次のように表すことができる。

$$W = \omega \times I + \frac{1}{3} (p, v) I - pv^T \quad (15)$$

これから、解は次の方針で求まる。

1. 観測したフロー m に対して式 (14) を満たすフロー行列 W を求める。
2. 得られたフロー行列 W を式 (15) のように運動パラメータ $\{\omega, v\}$ と P ベクトル p に分解する。

フロー行列 W には 9 個の要素があるが、正規化条件 $\text{tr}W=0$ があるから、ステップ 1 では少なくとも 4 画素でのフローの x, y 成分が与えられれば W が一意的に定まる。しかし、すでに指摘したように現実的には多数の画素について、フロー検出の信頼性によって重みづけした最小二乗法によってロバストに計算すべきである (付録 F)。あるいはフローを検出しなくても、閉曲線の運動がエッジ検出によって追跡できれば、閉曲線の各点での法線方向への速度 (法線フロー) からフロー行列 W が一般には決定できる [4, 45, 15, 16, 47, 48]。ただし、この方法は単純な代数曲線の多く (例えば楕円, 双曲線, 放物線) には適用できないことが Bergholm [2] によって指摘され、そのようなあまい曲線は Bergholm and Carlsson [3] によって詳細に分類されている。

ステップ 2 は行列 W に関する固有値と固有ベクトルの解析によって解析的に計算できる (付録 G)。式 (15) から直ちにわかるように、 $\{v, p\}$ が解であれば $\{-v, -p\}$ も解である。有限運動の場合と同様に、この二つの解に対応する平面は視点に関して互いに鏡像の位置にある。それぞれに対して二つの解が求まるから、解は合計 4 個ある。しかし、その内の二つはカメラから見えない方向にあるので、結局解は 2 個となる [22]。平面の (外向き) 単位法線ベクトルを $n=N[p]$ とするとき、 $\{v, n\}$ が並進速度および平面の単位法線ベクトルと解釈できるなら、他方の解釈は $\{n, v\}$ である [25]。並進方向が平面に直交するときは解は一意的である。

付録

A. 基本行列の計算

与えられた N ベクトル $\{m_\alpha\}, \{m'_\alpha\}, \alpha=1, \dots, N$ から基本行列 G は次のように計算できる。

1. 相関テンソル $M=(M_{ijkl})$ を

$$M_{ijkl} = \sum_{\alpha=1}^N m_{\alpha(i)} m'_{\alpha(j)} m_{\alpha(k)} m'_{\alpha(l)} \quad (16)$$

で定義する。ただし $m_{\alpha(i)}$ は m_α の第 i 成分である。

2. 9 次元行列 $\hat{M}=(\hat{M}_{\kappa\lambda})$ を $i=(\kappa-1)\text{div}3+1, j=(\kappa-1)\text{mod}3+1, k=(\lambda-1)\text{div}3+1, l=(\lambda-1)\text{mod}3+1$ とおき、 $\hat{M}_{\kappa\lambda}=M_{ijkl}$ と定義する。ただし div, mod はそれぞれ整数商, 整数剰余の演算を表す。
3. 行列 $\hat{M}$ の最小固有値に対するノルム $\sqrt{2}$ の 9 次元固有ベクトルを $\hat{G}=(\hat{G}_\kappa)$ とする。
4. 基本行列 $G=(G_{ij})$ は $G_{ij}=\hat{G}_{3(i-1)+j}$ で与えられる。

B. 運動パラメータの計算

N ベクトル $\{m_\alpha\}, \{m'_\alpha\}$ から基本行列 G が求まれば、運動パラメータ $\{R, h\}$ が次のように計算される。

1. 並進 h は行列 GG^T の最小固有値に対する単位固有ベクトルで与えられる。符号は次のように選ぶ。

$$\sum_{\alpha=1}^N |h, m_\alpha, Gm_\alpha| > 0. \quad (17)$$

2. 行列 $-h \times G$ を次のように特異値分解する (V, U は直交行列, Λ は非負の対角要素が大きい順に並んだ対角行列 [19])。

$$-h \times G = V \Lambda U^T \quad (18)$$

3. 回転 R は次のように与えられる。

$$R = V \begin{pmatrix} 1 & & \\ & 1 & \\ & & \det(VU^T) \end{pmatrix} U^T \quad (19)$$

C. 2 次元射影変換の計算

与えられた N ベクトル $\{m_\alpha\}, \{m'_\alpha\}, \alpha=1, \dots, N$ から射影変換の行列 A は次のように計算できる。

1. 式 (16) の相関テンソル $M=(M_{ijkl})$ からテンソル $T=(T_{ijkl})$ を次のように定義する (δ_{ij} は Kronecker のデルタ)。

$$T_{ijkl} = \sum_{m=1}^3 M_{imkm} \delta_{jl} - M_{ijkl} \quad (20)$$

2. 9 次元行列 $\hat{T}=(\hat{T}_{\kappa\lambda})$ を $i=(\kappa-1)\text{div}3+1, j=(\kappa-1)\text{mod}3+1, k=(\lambda-1)\text{div}3+1, l=(\lambda-1)\text{mod}3+1$ とおき、 $\hat{T}_{\kappa\lambda}=T_{ijkl}$ と定義する。
3. 行列 T の最小固有値に対する 9 次元単位固有ベクトルを $\hat{A}=(\hat{A}_\kappa)$ とする。射影変換の行列 $A=(A_{ij})$ は $A_{ij}=\hat{A}_{3(i-1)+j}$ で与えられる。

D. 射影変換の 3 次元解釈

与えられた射影変換の行列 A から運動パラメータ $\{R, h\}$ と平面の P ベクトル p は次のように計算できる。ただし、カメラの運動が平面の同じ側で起こる場合 (場合 1) と平面の反対側に回り込む場合 (場合 2) を区別する。

1. 行列 AA^T の固有値を $\lambda_1 \geq \lambda_2 \geq \lambda_3 (> 0)$ とし、対応する固有ベクトルのつくる正規直交系を $\{u_1, u_2, u_3\}$ とする。 $\sigma_i = \sqrt{\lambda_i} (> 0), i=1, 2, 3$ と置く。
2. 場合 1 では P ベクトル p は次のように与えられる。

$$p = \frac{1}{\sigma_2} \sqrt{\frac{\sigma_1 - \sigma_3}{\sigma_1 + \sigma_3}} (\sqrt{\sigma_1^2 - \sigma_2^2} u_1 \pm \sqrt{\sigma_2^2 - \sigma_3^2} u_3) \quad (21)$$

場合 2 では P ベクトル p は次のように与えられる。

$$p = \frac{1}{\sigma_2} \sqrt{\frac{\sigma_1 + \sigma_3}{\sigma_1 - \sigma_3}} (\sqrt{\sigma_1^2 - \sigma_2^2} u_1 \pm \sqrt{\sigma_2^2 - \sigma_3^2} u_3) \quad (22)$$

3. 場合1では並進 h は次のように与えられる.

$$h = \frac{-\sigma_3 \sqrt{\sigma_1^2 - \sigma_2^2} u_1 \pm \sigma_1 \sqrt{\sigma_2^2 - \sigma_3^2} u_3}{\sigma_2 \sqrt{\sigma_1^2 - \sigma_3^2}} \quad (23)$$

場合2では並進 h は次のように与えられる.

$$h = \frac{\sigma_3 \sqrt{\sigma_1^2 - \sigma_2^2} u_1 \pm \sigma_1 \sqrt{\sigma_2^2 - \sigma_3^2} u_3}{\sigma_2 \sqrt{\sigma_1^2 - \sigma_3^2}} \quad (24)$$

4. 場合1では回転 R は次のように与えられる.

$$R = \frac{1}{\sigma_2} (I + \sigma_2^3 p h^T) A \quad (25)$$

場合2では回転 R は次のように与えられる.

$$R = \frac{1}{\sigma_2} (-I + \sigma_2^3 p h^T) A \quad (26)$$

5. それぞれの解について, p と h の符号を同時に反転させたものも解である.

E. オプティカルフローの3次元解釈

与えられたフロー m から運動パラメータ $\{\omega, v\}$ が次のように計算できる.

1. テンソル $\mathcal{L}=(L_{ij})$, $\mathcal{M}=(M_{ijk})$, $\mathcal{N}=(N_{ijkl})$ を次のように定義する.

$$L_{ij} = \int_{\Omega(S)} m_i^* m_j^* d\Omega(m) \quad (27)$$

$$M_{ijk} = \int_{\Omega(S)} m_i^* m_j m_k d\Omega(m) \quad (28)$$

$$N_{ijkl} = \int_{\Omega(S)} m_i m_j m_k m_l d\Omega(m) \quad (29)$$

積分記号 $\int_{\Omega(S)}$ はフローの定義されているすべての画素にわたる総和の意味である. しかしフローの精度は場所ごとに異なるので, 現実的には信頼性を考慮した重みつき総和を用いるべきである.

2. テンソル $\mathcal{N}$ の逆テンソル $\mathcal{N}^{-1}=(N_{ijkl}^{-1})$ を計算する [22].

3. 行列 $A=(A_{ij})$ を次のように定義する.

$$A_{ij} = L_{ij} - \sum_{k,l,m,n=1}^3 M_{ikl} N_{klmn}^{-1} M_{jmn} \quad (30)$$

4. 並進速度 v は行列 A の最小固有値に対する単位固有ベクトルで与えられる.

5. 対称行列 K を次のように計算する.

$$K_{ij} = - \sum_{k,l,m=1}^3 N_{ijkl}^{-1} M_{mkl} v_m, \quad (31)$$

6. 回転速度 ω は次のように与えられる.

$$\omega = \frac{1}{2} (\text{tr} K + 3(v, Kv)) v - 2Kv \quad (32)$$

式 (29) からわかるように, テンソル $\mathcal{N}$ はフローに無関係に視野 $\Omega(S)$ の幾何学的形状のみから定まり, 誤差を含んでいない. その逆テンソル $\mathcal{N}^{-1}$ を含んだ式はそれを係数とする連立1次方程式の解を記述しており, 誤差の影響はその行列式に大きさに依存する. その程度は“条件数”によって評価できる. このことから, 一般に解の安定のために視野 $\Omega(S)$ が広いことが望ましい.

F. フロー行列の計算

あたえられたフロー m からフロー行列 W が次のように計算できる.

1. テンソル $A=(A_{ijkl})$ と行列 $B=(B_{ij})$ を次のように定義する (δ_{ij} は Kronecker デルタ).

$$A_{ijkl} = \int_{\Omega(S)} (\delta_{il} m_j m_k - m_i m_j m_k m_l) d\Omega(m) \quad (33)$$

$$B_{ij} = \int_{\Omega(S)} m_i m_j d\Omega(m) \quad (34)$$

積分記号 $\int_{\Omega(S)}$ の意味は付録Eと同じである.

2. 9次元行列 $\hat{A}=(\hat{A}_{\kappa\lambda})$ と9次元ベクトル $\hat{b}=(\hat{b}_\kappa)$ を $i=(\kappa-1)\text{div}3+1$, $j=(\kappa-1)\text{mod}3+1$, $k=(\lambda-1)\text{div}3+1$, $l=(\lambda-1)\text{mod}3+1$ とおき, 次のように定義する.

$$\hat{A}_{\kappa\lambda} = A_{ijkl} \quad (\kappa \neq 9), \quad 1 \quad (\kappa = 9) \quad (35)$$

$$\hat{b}_\kappa = B_{ij} \quad (\kappa \neq 9), \quad 0 \quad (\kappa = 9) \quad (36)$$

3. 連立1次方程式 $\hat{A}\hat{w}=\hat{b}$ の解を $\hat{w}=(\hat{w}_\kappa)$ とすると, フロー行列 $W=(W_{ij})$ は $W_{ij}=\hat{w}_{3(i-1)+j}$ で与えられる.

G. 平面オプティカルフローの3次元解釈

与えられたフロー行列 W から運動パラメータ $\{\omega, v\}$ と平面のPベクトル p は次のように計算できる.

1. フロー行列 W の対称部分を $W^{(s)}$, 反対称部分を $W^{(a)}$ とする.

$$W^{(s)} = \frac{1}{2} (W + W^T) \quad (37)$$

$$W^{(a)} = \frac{1}{2} (W - W^T) \quad (38)$$

2. ベクトル $w = (W_{23}^{(a)}, W_{31}^{(a)}, W_{12}^{(a)})^T$ を計算する.

3. もし $W_s=O$ であれば $v=0$, $\omega=w$ であり, p は不定である. そうでなければ $W^{(s)}$ の固有値を $\sigma_1 \geq \sigma_2 \geq \sigma_3$ とし, 対応する固有ベクトルのつくる正規直交系を $\{u_1, u_2, u_3\}$ とする.

4. Pベクトル p は次のように与えられる.

$$p = \sqrt{\sigma_1 - \sigma_3} (\pm \sqrt{\sigma_1 - \sigma_2} u_1 + \sqrt{\sigma_2 - \sigma_3} u_3) \quad (39)$$

5. 並進速度 v は次のように与えられる.

$$v = -\frac{1}{\sqrt{\sigma_1 - \sigma_3}} (\pm\sqrt{\sigma_1 - \sigma_2}u_1 - \sqrt{\sigma_2 - \sigma_3}u_3) \quad (40)$$

6. 回転速度 ω は次のように与えられる.

$$\omega = w - \frac{1}{2}p \times v \quad (41)$$

7. それぞれの解について, p と v の符号を同時に反転させたものも解である.

参考文献

- [1] J.(Y.)Aloimonos, I.Weiss and A.Bandyopadhyay: Active vision, *Int. J. Comput. Vision.*, 1(1988) 333-356.
- [2] F.Bergholm: Motion from flow along contours: A note on robustness and ambiguous cases, *Int. J. Comput. Vision.*, 4(1989) 395-415.
- [3] F.Bergholm and S.Carlsson: A "theory" of optical flow, *CVGIP: Image Understanding*, 53(1991) 171-188.
- [4] L.S.Davis, Z.Wu and H.Sun: Contour-based motion estimation, *Comput. Vision Graphics Image Process.*, 23(1983) 313-326.
- [5] O.D.Faugeras and S.Maybank: Motion from point matches: Multiplicity of solutions, *Int. J. Comput. Vision.*, 4(1990) 225-246.
- [6] J.J.Gibson: *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, MA, 1979 (古崎 敬, 古崎 愛子, 辻 敏一郎, 村瀬 晃訳: 「生態学的視覚論」, サイエンス社, 1985).
- [7] J.C.Hay: Optical motions and space perception—An extension of Gibson's analysis, *Psych. Rev.*, 73(1966) 550-565.
- [8] D.J.Heeger and A.D.Jepson: Subspace methods for recovering rigid motion I: Algorithm and implementation, *Int. J. Comput. Vision.*, 7(1992) 95-117.
- [9] B.K.P.Horn: Motion fields are hardly ever ambiguous, *Int. J. Comput. Vision.*, 1(1987) 259-274.
- [10] B.K.P.Horn: Relative orientation, *Int. J. Comput. Vision.*, 4(1990) 59-78.
- [11] T.S.Huang and O.D.Faugeras: Some properties of the E matrix in two-view motion estimation, *IEEE Trans. Patt. Anal. Machine Intell.*, 11(1989) 1310-1312.
- [12] 岩崎 夫, 金谷 健一: くりこみ法によるオブティカルフローの3次元復元, 情報処理学会研究報告 92-CV-80(1992-11).
- [13] C.Jerian and R.Jain: Polynomial methods for structure from motion, *IEEE Trans. Patt. Anal. Machine Intell.*, 12(1990) 1150-1165.
- [14] 梶浦義弘, 大西 昇, 杉江 昇: 回転探索法による運動物体の3次元運動パラメータ推定法, 電子情報通信学会論文集 D-II, J75-D-II-5(1992) 873-882.
- [15] K.Kanatani: Tracing planar surface motion from a projection without knowing correspondence, *Comput. Vision Graphics Image Process.*, 29(1985) 1-12.
- [16] K.Kanatani: Detecting the motion of a planar surface by line and surface integrals, *Comput. Vision Graphics Image Process.*, 29(1985) 13-22.
- [17] K.Kanatani: Structure and motion from optical flow under perspective projection, *Comput. Vision Graphics Image Process.*, 38(1987) 328-354.
- [18] K.Kanatani: *Group-Theoretical Methods in Image Understanding*, Springer, Berlin, 1990.
- [19] 金谷 健一: 「画像理解/3次元認識の数理」 森北出版, 1990.
- [20] K.Kanatani, Computational projective geometry, *CVGIP: Image Understanding*, 54(1991) 333-348.
- [21] 金谷 健一, 3次元運動のコンピュータシミュレーション, 光学, 21-10(1992) 650-656.
- [22] K.Kanatani, *Geometric Computation for Machine Vision*, Oxford Univ. Press, Oxford, 1993 (to appear).
- [23] H.C.Longuet-Higgins: A computer algorithm for reconstructing a scene from two projections, *Nature*, 293-10(1981) 133-135.
- [24] H.C.Longuet-Higgins: The reconstruction of a scene from two projections—Configurations that defeat the 8-point algorithm, *Proc. 1st IEEE Conf. Artif. Intell. Appl.*, Denver, CO, December 1984, pp. 395-397.
- [25] H.C.Longuet-Higgins: The visual ambiguity of a moving plane, *Proc. Roy. Soc. Lond.*, B-223(1984) 165-175.
- [26] H.C.Longuet-Higgins: The reconstruction of a plane surface from two perspective projections, *Proc. Roy. Soc. Lond.*, B-227(1986) 399-410.
- [27] H.C.Longuet-Higgins: Multiple interpretations of a pair of images of a surface, *Proc. Roy. Soc. Lond.*, A-418(1988) 1-15.
- [28] S.Maybank: The angular velocity associated with the optical flow field arising from motion through a rigid environment, *Proc. Roy. Soc. Lond.*, A-401(1985) 317-326.
- [29] S.J.Maybank: The projective geometry of ambiguous surfaces, *Phil. Trans. Roy. Soc. Lond.*, A-332(1990) 1-47.
- [30] S.J.Maybank: Ambiguity in reconstruction from image correspondences, *Image Vision Comput.*, 9(1991) 93-99.
- [31] 森尻智昭, 金谷 健一: 3次元運動の不偏推定法と統計的誤差解析, 情報処理学会研究報告 92-CV-77(1992-03).
- [32] S.Negahdaripour: Critical surface pairs and triplets, *Int. J. Comput. Vision.*, 3(1989) 293-312.
- [33] S.Negahdaripour: Closed-form relationship between the two interpretations of a moving plane, *J. Opt. Soc. Am.*, A-7(1990) 279-285.
- [34] S.Negahdaripour: Multiple interpretations of the shape and motion of objects from two perspective images, *IEEE Trans. Patt. Anal. Machine Intell.*, 12(1990) 1025-1030.
- [35] A.N.Netravali, T.S.Huang, A.S.Krishnakumar and R.J.Holt: Algebraic methods in 3-D motion estimation from two-view point correspondences, *Int. J. Imaging Sys. Tech.*, 1(1989) 78-99.
- [36] M.E.Spetsakis and J.(Y.)Aloimonos: Optimal computing of structure from motion using point correspondences in two frames, *Proc. 2nd Int. Conf. Computer Vision*, Tampa, FL, December 1988, pp. 449-453.
- [37] M.Subbarao and A.M.Waxman: Closed form solution to image flow equations for planar surfaces in motion, *Comput. Vision Graphics Image Process.*, 36(1986) 208-228.
- [38] 田川憲男, 島生隆, 遠藤利生: オブティカルフローを用いた閉形式の3次元剛体運動復元方式, 電子情報通信学会技術研究報告 IE90-108/PRU90-139(1991-03).
- [39] 武田佐千男, 金谷 健一: くりこみ法による3次元運動の解析, 情報処理学会研究報告 93-CV-82(1993-03).
- [40] R.Y.Tsai and T.S.Huang: Estimating three-dimensional motion parameters of a rigid planar patch, *IEEE Trans. Acoustics Speech Signal Process.*, 29(1981) 1147-1152.
- [41] T.Y.Tsai and T.S.Huang: Estimating 3-D motion parameters of a rigid planar patch III. Finite point correspondences and the three-view problem, *IEEE Trans. Acoustics Speech Signal Process.*, 32(1984) 213-220.
- [42] R.Y.Tsai and T.S.Huang: Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces, *IEEE Trans. Patt. Anal. Machine Intell.*, 6(1984) 13-27.
- [43] R.Y.Tsai, T.S.Huang and W.-L.Zhu: Estimating three-dimensional motion parameters of a rigid planar patch, II: Singular value decomposition, *IEEE Trans. Acoustics Speech Signal Process.*, 30(1982) 525-534.
- [44] S.Ullman: *The Interpretation of Visual Motion*, MIT Press, Cambridge, MA, 1979.
- [45] A.M.Waxman and K.Wohn: Contour evolution, neighborhood deformation, and global image flow: Planar surfaces in motion, *Int. J. Robotics Res.*, 4(1985) 95-108.
- [46] J.Weng, T.S.Huang and N.Ahuja: Motion and structure from two perspective views: Algorithms, error analysis, and error estimation, *IEEE Trans. Patt. Anal. Machine Intell.*, 11(1989) 451-467.
- [47] K.Wohn and A.M.Waxman: The analytic structure of image flows: Deformation and segmentation, *Comput. Vision Graphics Image Process.*, 49(1990) 127-151.
- [48] K.Y.Wohn, J.Wu, and R.W.Brockett: A contour-based recovery of image flow: Iterative transformation method, *IEEE Trans. Patt. Anal. Machine Intell.*, 13(1991) 746-760.