

剛体・非剛体物体の3次元構造・運動推定 - 運動学的アプローチ -

6E-6

森川 博之 原 島 博
東京大学・工学部

1. はじめに

一般に、画像は3次元世界の2次元投影像として定義される。従って、このような画像の生成モデルともいえるべき3次元構造・運動情報を利用することにより、柔軟な画像の操作が可能となり、画像の符号化および処理において新たな可能性が生じよう⁽¹⁾。

このような観点から、筆者らは、動画像からの3次元構造・運動推定法、抽出した3次元情報を用いる動画像符号化方式について基礎検討を行っている⁽²⁾⁻⁽⁴⁾。

本稿では、剛体仮定を用いずに、“運動の滑らかさ”条件を用いて、逐次的に3次元構造・運動を推定する手法を示す。以下、ロバストでかつ非剛体物体にも適用できる柔軟な3次元構造・運動推定手法を、運動学的な観点から述べる。

2. 運動の滑らかさ

画像は、3次元世界の2次元投影像であるため、動画像から3次元構造・運動情報を抽出する際には、拘束条件が必要となる。このような拘束条件として、対象物が剛体であるという剛体仮定を用いて、2, 3フレームの画像から3次元構造・運動を推定することが多い。しかし、剛体仮定は、ノイズに対して弱く、また、適用環境も制限されるという欠点を有する。このように、剛体仮定は、3次元構造・運動推定を行うにあたり、かなり厳しい拘束条件であると言える。

一般に、動く物体は、物体の有する慣性のために、滑らかな運動を行う。短いフレーム間隔を考えると、運動パラメータはフレーム間で相関を有し、急激な変化が生じることは少ない。従って、このような“運動の滑らかさ”を拘束条件として用いることは、剛体仮定に比べ、より汎用的であり、強力であると考えられる。“運動の滑らかさ”条件は、動画像系列のもつ時間的冗長性を利用し、構造・運動情報を蓄積する手段を提供する。

実際、人間の視覚システムにおいても、正確な3次元構造・運動の認知には、長い画像系列が必要であることが知られている。また、ロバスト性も画像数に比例することが示されている。

このように、“運動の滑らかさ”を用いることにより、各フレームの3次元情報を蓄積して、逐次的に3次元構造・運動を推定することが可能となる。蓄積された3次元情報を利用することで、ノイズに対してロバストで、かつ剛体からのずれを有する非剛体運動をも許容できるような、柔軟な3次元構造・運動推定が行える。

3. 3次元構造・運動推定手法

3.1 運動モデル

一般に、3次元物体の運動は、空間内の座標点のベクトル値関数 F として表現できる。時刻 t における第 i 点 $\mathbf{x}_i(t) = (x_i(t), y_i(t), z_i(t))^T$ が、次の時刻 $t+1$ において点 $\mathbf{x}_i(t+1) = (x_i(t+1), y_i(t+1), z_i(t+1))^T$ に移動したとすると、運動は次式で表される。

$$\mathbf{x}_i(t+1) = F(\mathbf{x}_i(t)) \quad (1)$$

物体が、剛体運動あるいは一様な変形を行う場合には、 F は $\mathbf{x}(t)$ に関して線形となる。ヘルムホルツの運動の基礎定理によれば、変形物体の小さな部位の運動は、(i)回転、(ii)3つの互いに直角な方向への膨張(あるいは収縮)、(iii)並進、の和として表されることが示される⁽⁵⁾。

ヘルムホルツの定理は、小さな部位における運動という仮定のもとで変形運動を表現したものである。一方、変形運動に対する視点を変えることにより、式(2)に示す表現を用いて、より一般的な非剛体運動を表すこともできる。

$\mathbf{R}_{t+1}$ 、 $\mathbf{D}_{t+1}^i$ 、 $\mathbf{T}(t)$ を、それぞれ時刻 t と $t+1$ 間の剛体としての回転、第 i 点の歪を表すベクトル値関数、3次元並進運動とすると、運動は、

$$\mathbf{x}_i(t+1) = \mathbf{R}_{t+1} \mathbf{D}_{t+1}^i [\mathbf{x}_i(t)] + \mathbf{T}(t+1) \quad (2)$$

で与えられる。

この運動表現は、まず物体を変形させ、その後、剛体としての回転を行うという操作に対応する。式(2)は、式(1)と比べて、剛体運動をより意識した非剛体運動の表現であり、剛体運動からのずれの小さい非剛体運動を考える上で、扱い易い表現法である。 $\mathbf{D}_{t+1}^i$ の与え方により、一様でない変形の表現も可能で、剛体からのずれの小さい多くの非剛体運動を表すことができる。

ところで、非剛体運動は、剛体運動に比べ自由度が増大するため、処理は非常に複雑なものとなる。しかし、現実的に妥当な簡潔化を図ることにより、処理し易い表現が得られる。このような観点から、ここでは、運動が滑らかな、剛体からのずれの小さい非剛体運動を考え、また、動画像系列のフレーム間隔は短いものとする。すなわち、回転 $\mathbf{R}_{t+1}$ 、第 i 点の変形 $\mathbf{D}_{t+1}^i$ 、平行移動 $\mathbf{T}(t)$ を、

$$\mathbf{R}_{t+1} = \mathbf{R}[\mathbf{w}(t+1)] = \begin{bmatrix} 1 & -w_z(t+1) & w_y(t+1) \\ w_z(t+1) & 1 & -w_x(t+1) \\ -w_y(t+1) & w_x(t+1) & 1 \end{bmatrix} \quad (3)$$

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \Delta \mathbf{w}(t) \quad (4)$$

$$\mathbf{D}_{t+1}^i : (x_i(t), y_i(t), z_i(t))^T \rightarrow (x_i(t), y_i(t), z_i(t) + \Delta z_i(t))^T \quad (5)$$

$$\mathbf{T}(t+1) = \mathbf{T}(t) + \Delta \mathbf{T}(t) \quad (6)$$

と表す。ここに、 $\mathbf{w}(t) = (w_x(t), w_y(t), w_z(t))^T$ である。画面上で観察できる運動の関係式は、式(3)~(6)を用

いて得られる。なお、ここでは、イメージモデルとして平行投影 h を仮定する。点 $\mathbf{x}_i(t)$, $\mathbf{x}_i(t+1)$ の画面上で観察される座標を、それぞれ $\mathbf{u}_i(t) = (u_i(t), v_i(t))^T$, $\mathbf{u}_i(t+1) = (u_i(t+1), v_i(t+1))^T$ とすると、画面上での運動の関係式は次式で与えられる。

$$\mathbf{u}_i(t) = h[\mathbf{x}_i(t)] = \{\mathbf{x}_i(t), y_i(t)\}^T \quad (7)$$

$$\mathbf{u}_i(t+1) = h[\mathbf{x}_i(t+1)] \\ = h[\mathbf{R}_{t+1} \cdot \mathbf{D}^i_{t+1}[\mathbf{x}_i(t)]] + \mathbf{T}^*(t) + \Delta \mathbf{T}^*(t) \quad (8)$$

ここに、 $\mathbf{T}^*(t) = h[\mathbf{T}(t)] = (T_x(t), T_y(t))^T$, $\Delta \mathbf{T}^*(t) = (\Delta T_x(t), \Delta T_y(t))^T$ である。

3.2 正則化を用いる3次元構造・運動推定

画面上での非剛体運動の表現として、式(8)(9)を得たが、実際には、モデルと現実の運動とのずれ、動き情報に含まれる誤差のために、式(9)の左辺と右辺は一致しない。そこで、最小2乗的な意味で、左辺と右辺との差を最小にする $\mathbf{w}(t+1)$, $\mathbf{T}^*(t+1)$, $\mathbf{z}(t+1) = (z_1(t+1), z_2(t+1), \dots, z_n(t+1))^T$ を求める。すなわち、 $\mathbf{w}(t)$, $\mathbf{T}^*(t)$, $\mathbf{z}(t)$ を基に、その変化量である $\Delta \mathbf{w}(t)$, $\Delta \mathbf{T}^*(t)$, $\Delta \mathbf{z}(t) = (\Delta z_1(t), \Delta z_2(t), \dots, \Delta z_n(t))^T$ を求める。

ところで、2. で示した“運動の滑らかさ”条件は、上の3.1で述べたモデル化においては、 $\Delta \mathbf{w}(t)$, $\Delta \mathbf{T}^*(t)$, $\Delta \mathbf{z}(t) = (\Delta z_1(t), \Delta z_2(t), \dots, \Delta z_n(t))^T$ の各成分が小さいということを意味する。そこで、このような拘束条件を、正則化の枠組みのもとで定式化する。つまり、正則化における安定化関数 $P(\mathbf{w}(t), \mathbf{T}^*(t), \mathbf{z}(t))$ に、“運動の滑らかさ”条件を反映させる。ここでは、この安定化関数を次のように定義するものとする⁽³⁾。

$$\|P\|^2 = \alpha \|\Delta \mathbf{w}(t)\|^2 + \beta \|\Delta \mathbf{T}^*(t)\|^2 + \gamma \|\Delta \mathbf{z}(t)\|^2 \quad (9)$$

ここに、 α, β, γ は拘束条件の強さを表すパラメータであり、 $\|\cdot\|$ は L_2 ノルムを示す。

このように式(9)に示した“運動の滑らかさ”条件のもとで、式(8)の左辺と右辺を最小にする変化量を求める。従って、最小化する評価関数 $E(\Delta \mathbf{w}(t), \Delta \mathbf{T}^*(t), \Delta \mathbf{z}(t))$ は、

$$E(\Delta \mathbf{w}(t), \Delta \mathbf{T}^*(t), \Delta \mathbf{z}(t)) \\ = \sum_i (\mathbf{u}_i(t+1) - h[\mathbf{x}_i(t+1)])^2 + \|P\|^2 \quad (10)$$

となる。評価関数 E の第1項は、観測値と運動モデルとのずれを示し、式(8)における左辺と右辺との残差平方和である。累積和は同一物体に属している n 個の点に関してとるものとする。第2項は、上で述べた安定化関数である。

4. 特性評価

式(10)に示した最小化は、求めるパラメータ $\Delta \mathbf{w}(t)$, $\Delta \mathbf{T}^*(t)$, $\Delta \mathbf{z}(t)$ に対して非線形となる。そこで、実際の計算においては、変化量が小さいという仮定に基づき、変化量の3次以上の項は微小であると近似する。すると、 E の最小化は線形連立1次方程式を解くことに帰着される。このような近似を用いても、ほぼ正確な構造・運動推定が可能である⁽³⁾。

本推定法を、非剛体運動を行う物体に適用した結果を図1に示す。X-Y平面上での動き情報から、3次元構造・運動の推定を考える。図1は6点のX-Z平面への投影像を示しており、徐々に3次元構造が推定される様子がみられる。

非剛体運動は各回転毎に加わっている。時刻 $t=0$ においては、3次元情報は得られていない。

剛体運動に対しても、数10フレームの繰り返しにより、ほぼ正確な3次元構造・運動の推定が可能であることが示されている⁽³⁾。さらに、“運動の滑らかさ”条件を用いているため、Ullmanの結果⁽⁶⁾でみられた、構造推定の際の振動現象は生じない。

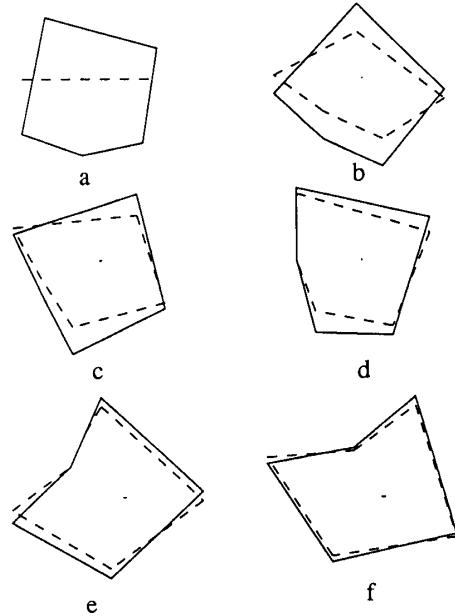


図1 6点の非剛体物体の3次元構造推定(X-Z平面): (a)第1フレーム, (b)第6フレーム, (c)第12フレーム, (d)第18フレーム, (e)第24フレーム, (f)第30フレーム。

5. むすび

剛体仮定に比べ、より一般的な“運動の滑らかさ”条件を用いることにより、ロバストでかつ非剛体運動にも適用できる柔軟な3次元構造・運動推定手法について述べた。“運動の滑らかさ”条件は、剛体仮定より、強力な拘束条件である。

今後、本推定法の収束性に関する数学的理論解析、適用範囲、中心投影への拡張などについて、詳細な検討を加える予定である。

【参考文献】

- (1) 原島, 森川, 青木: “3次元構造モデルを用いる画像の符号化・処理・表示”, 1989年信学春全大, SD-3-12 (1989-03).
- (2) 森川, 原島, 齊藤: “3次元構造・運動情報を用いる動画画像符号化”, PCSJ89, 8.3 (1989-10).
- (3) 森川, 原島: “正則化を用いる剛体・非剛体運動からの3次元構造復元”, 「画像理解の高度化と高速化」シンポジウム, B3.2, pp. 97-102 (1989-04).
- (4) 森川, 青木, 原島: “ステレオ動画画像からの3次元構造復元”, 信学技報, PRU89-57 (1989-09).
- (5) A. Sommerfeld: “変形体の力学”, ゾンマーフェルト理論物理学講座II, 講談社 (1969).
- (6) S. Ullman: “Maximizing rigidity: the incremental recovery of 3-D structure from rigid and nonrigid motion”, Perception, 13, pp. 255-274 (1984).