

## 時系列視体積からの多関節剛体の形状と運動の同時獲得手法 Automatic Shape and Motion Acquisition for Articulated Objects

飯山 将晃<sup>†</sup>  
Masaaki Iiyama

美濃 導彦<sup>†</sup>  
Michihiko Minoh

### 1. はじめに

デジタル博物館のような多種多様な物体の形状や運動をデジタルデータとして必要とする分野への応用を目的として、昆虫などの多関節剛体の形状と運動を同時に獲得する手法を提案する。

これまで、多関節剛体の運動を獲得する手法として、体節の幾何形状に対して何らかのモデル（一般化円筒など）を与えて体節への切り分けを行い、切り分けられた各体節に対して剛体運動を推定するアプローチがとられてきた。しかしながら、デジタル博物館のように多種多様な物体を獲得対象とする場合、獲得対象それぞれに対して体節の幾何形状のモデルを用意するのは非現実的である。この一方で、近年さまざまな形状計測方法が提案され、物体の形状をボリュウムデータとして獲得することが可能となった。しかしながら、物体の形状を計測したからといってその物体がどのように運動するのかを知ったことにはならず、物体の運動を知るためには、計測された形状をさらに体節ごとに切り分け、そこから各体節の剛体運動を抽出する処理が必要となる。

我々は体節の運動推定と全身形状の体節への切り分けを陽に行うのではなく、ボリュウムデータとして獲得された物体形状を入力データとし、多関節剛体がいくつの体節で構成されているかを前知識として与えることにより、物体形状のどの部分がどの体節に切り分けられ得るかについての確率分布を推定する。そしてその確率分布が与えられた条件下で、最も尤もらしい体節の運動成分を推定する。この2つの処理を反復的に行うことにより物体の運動と形状とを同時に推定する。物体の運動が小さい場合には、体節への切り分けに対して十分な情報が得られないため、推定された確率分布はどの体節にも属するような分布になるが、物体の運動が大きい場合には、体節への切り分けを行うのに十分なほど先鋭化された分布になり、体節への切り分けと運動推定を安定して行うことができる。

### 2. 問題設定

撮影環境として、多関節物体を取り囲むように複数台のカメラを設置した環境を想定する。各カメラで多関節物体を時系列で撮影し、各時刻における多関節物体の形状を視体積交差法 [1] で獲得する。この時系列で得られた視体積を入力データとし、この入力データより体節の形状と運動を推定する。なお、体節の数  $R$  は既知であるとする。

各時刻  $t_n (n = 0, \dots, N)$  における視体積  $\mathcal{V}_n$  を  $N_{x_x} \times N_{x_y} \times N_{x_z}$  個のボクセルからなる立方格子（ボクセル空間）で表現する。各ボクセルには2種類の値を持たせる。ひとつめの値は、視体積表面-ボクセル間の距離で

ある。座標  $\mathbf{x} = \{x_x, x_y, x_z\}$  に位置するボクセルから視体積表面までの最短距離を  $\mathcal{V}_n(\mathbf{x})$  とし、 $\mathcal{V}_n(\mathbf{x})$  の集合  $\mathcal{V}_n = \{\mathcal{V}_n(\mathbf{x}) | x_x = 0, \dots, N_{x_x} - 1; x_y = 0, \dots, N_{x_y} - 1; x_z = 0, \dots, N_{x_z} - 1\}$  を定義する。なお、 $\mathbf{x}$  に位置するボクセルが視体積に含まれない場合  $\mathcal{V}_n(\mathbf{x}) = 0$  とする。

ふたつめの値は、そのボクセルがどの体節に切り分けられるべきかを判別するラベル  $L = 0, \dots, r$  である。なお、 $\mathbf{x}$  に位置するボクセルがどの体節にも属さない場合や  $\mathbf{x}$  が視体積に含まれない場合にはラベル 0 を割り付ける。 $\mathbf{x}$  に位置するボクセルに割り付けられたラベルが  $i$  のとき 1、そうでないとき 0 を返す関数を  $z_i(\mathbf{x})$  とし、 $z$  を  $z_i(\mathbf{x})$  の集合 ( $z = \{z_i(\mathbf{x}) | i = 0, \dots, r; \mathbf{x} \in \mathcal{V}_n\}$ ) とすると、体節への分割は  $z$  を決定することによって実現される。

また、時刻  $t_n$  から  $t_0$  における各体節の剛体運動を  $R_{in} (i = 1, \dots, r)$  で表現する。

これより、多関節剛体の形状と運動の同時獲得問題は、 $\mathcal{V}_n$  より  $z$  と  $R_{in} (i = 0, \dots, r; n = 0, \dots, N)$  を推定する問題となる。

提案手法では2フレームの視体積を用いて  $z$  の確率分布と  $R_{in}$  を推定し、各フレーム組で得られた  $z$  の確率分布の推定結果を統合することで体節への切り分けを行う。

### 3. EM アルゴリズムを用いた視体積の体節への分割手法

体節への分割  $z$  の確率分布の推定と、その確率分布下での体節の運動  $R_{in}$  の推定を行うために EM アルゴリズム [2] を用いる。EM アルゴリズムは不完全データの最尤推定アルゴリズムであり、ある完全な確率変数（完全データ）のうち、その一部のみ観測できるという状況において、観測された不完全データ  $\mathbf{y}$  と、パラメータ  $\Phi$  を持つ確率モデル  $g(\mathbf{y}|\Phi)$  が与えられたときに、 $g(\mathbf{y}|\Phi)$  を最大とする  $\Phi$  を推定するアルゴリズムである。

多関節剛体の形状と運動の同時獲得問題を EM アルゴリズムに照らして考えた場合、観測データは視体積  $\mathcal{V}_n$ 、完全データは  $\{\mathcal{V}_n, z\}$  となり、 $R_{in}$  は確率モデルのパラメータ  $\Phi$  に含まれると考えることができる。

EM アルゴリズムはパラメータに適切な初期値  $\Phi^{(0)}$  を与え、E(Expectation) ステップと M(Maximization) ステップを繰り返すことによって  $\Phi$  を更新する。

#### 3.1 E ステップ

EM アルゴリズムの E ステップでは、与えられた  $\Phi^{(p)}$  に対して、完全データの対数尤度の条件付期待値  $Q$  を計算する。ここで  $Q$  を計算するために、入力データに関して以下の2つの性質を仮定する。

隣接するボクセル間の相関 隣接するボクセルに割り当てられたラベルは同一である可能性が大きい。この性質

<sup>†</sup>京都大学 学術情報メディアセンター

をマルコフ確率場モデルでモデル化する。  $\mathbf{x}$  に位置するボクセルの近傍のラベルが既知であるとき、  $\mathbf{x}$  に割り当てられたラベルが  $i$  である確率  $\pi_i(\mathbf{x})$  は、

$$\pi_i(\mathbf{x}) = P(z_i(\mathbf{x}) = 1 | z_\eta(\mathbf{x})) = \frac{\exp(\alpha_i + \beta \kappa_i(\mathbf{x}))}{\sum_k \exp(\alpha_k + \beta \kappa_k(\mathbf{x}))} \quad (1)$$

となる。ここで  $zeta_{\mathbf{x}}$  は  $\mathbf{x}$  の 26 近傍にあるボクセルに割り当てられたラベルの集合、  $\alpha_i (i = 0, \dots, r)$ ,  $\beta$  はマルコフ確率場モデルの内部パラメータであり、確率モデルのパラメータ  $\Phi$  に含まれる変数である。また、  $\kappa_i(\mathbf{x})$  は  $\mathbf{x}$  の 26 近傍にあるボクセルのうち、ラベル  $i$  が割り当てられているものの個数の期待値である。

視体積表面-ボクセル間の距離値変化 時刻  $t_n$  における  $\mathcal{V}_n(\mathbf{x})$ ,  $t_n$  と  $t_0$  の間の運動  $R_{in}$ , そして、時刻  $t_0$  において  $\mathbf{x}$  を  $R_{in}$  移動させた位置に存在するボクセルが持つ距離値  $\mathcal{V}_0(R_{in}\mathbf{x})$  について、  $\mathcal{V}_n(\mathbf{x})$  と  $\mathcal{V}_0(R_{in}\mathbf{x})$  とはほぼ等しいことが予想される。そこで、この関係を以下のように定式化する。

$$P(\mathcal{V}_n(\mathbf{x}) | z_i(\mathbf{x}) = 1, R_{in}) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{[\mathcal{V}_n(\mathbf{x}) - \mathcal{V}_0(R_{in}\mathbf{x})]^2}{2\sigma_i^2}\right\} \quad (2)$$

ここで、分散  $\sigma_i$  と剛体運動  $R_{in}$  が確率モデルパラメータ  $\Phi$  に含まれる。

完全データの対数尤度の条件付期待値  $Q$  は完全データ  $(\mathcal{V}_n, z)$  の対数尤度  $l_c$  を用いて  $Q = E[l_c | \mathcal{V}_n, \Phi^{(p)}]$  となる。ここで、  $\Phi$  は確率モデルのパラメータであり、  $\alpha_i, \beta, R_{i,t_n}, \sigma_i$  で構成される。

$l_c$  は  $P(\mathcal{V}_n) = \prod P(\mathcal{V}_n(\mathbf{x}))$  であることを利用して以下の式で表される

$$\begin{aligned} l_c &= \log P(\mathcal{V}_n, z | \Phi) \\ &= \sum_{\mathbf{x}} \log P(\mathcal{V}_n(\mathbf{x}) | z, \Phi) + \log P(z | \Phi) \\ &= \sum_{\mathbf{x}, i} z_i(\mathbf{x}) \log [P(\mathcal{V}_n(\mathbf{x}) | z_i(\mathbf{x}) = 1, \Phi)] \\ &\quad + \log P(z | \Phi) \end{aligned} \quad (3)$$

$\{\mathcal{V}_n, \Phi^{(p)}\}$  が与えられたときの  $z_i(\mathbf{x})$  の期待値  $\lambda_i(\mathbf{x})$  は、

$$\begin{aligned} \lambda_i(\mathbf{x}) &= E[z_i(\mathbf{x}) | \mathcal{V}_n, \Phi^{(p)}] = P(z_i(\mathbf{x}) = 1 | \mathcal{V}_n, \Phi^{(p)}) \\ &\approx \frac{P(\mathcal{V}_n(\mathbf{x}) | z_i(\mathbf{x}) = 1, \Phi^{(p)}) \pi_i^{(p)}(\mathbf{x})}{\sum_k P(\mathcal{V}_n(\mathbf{x}) | z_k(\mathbf{x}) = 1, \Phi^{(p)}) \pi_k^{(p)}(\mathbf{x})} \end{aligned} \quad (4)$$

となる。また、  $P(z | \Phi) \approx \prod_{\mathbf{x}, i} P(z_i(\mathbf{x}) | z_\eta(\mathbf{x}), \Phi)$  と近似すると、  $\log P(z | \Phi)$  の条件付期待値は、

$$\begin{aligned} E[\log P(z | \Phi) | \mathcal{V}_n, \Phi^{(p)}] &= \sum_{\mathbf{x}, i} E[z_i(\mathbf{x}) | \mathcal{V}_n, \Phi^{(p)}] \log P(z_i(\mathbf{x}) = 1 | z_\eta(\mathbf{x}), \Phi) \\ &= \sum_{\mathbf{x}, i} \lambda_i(\mathbf{x}) \log \pi_i(\mathbf{x}) \end{aligned} \quad (5)$$

となる。式 (2)(3)(4)(5) より、  $l_c$  の条件付期待値  $Q$  は、

$$\begin{aligned} Q &= E[l_c | \mathcal{V}_n, \Phi^{(p)}] \\ &= \sum_{\mathbf{x}, i} \lambda_i(\mathbf{x}) \left[ -\frac{1}{2} \log(2\pi\sigma^2) - \frac{\{\mathcal{V}_n(\mathbf{x}) - \mathcal{V}_0(R_{i,t_n}\mathbf{x})\}^2}{2\sigma^2} + \log \pi_i(\mathbf{x}) \right] \end{aligned} \quad (6)$$

で得られる。

### 3.2 Mステップ

Mステップでは  $Q$  を  $\Phi$  について最大化し、得られた  $\Phi$  を  $\Phi^{(p+1)}$  とする。

## 4. シミュレーション実験

上腕・前腕・手の3つの体節を持つ物体に対して、提案手法を適用した結果を図1に示す。図1(a)から図1(b)へと運動する200フレームの時系列視体積より図1(a)を3つの体節に分割した結果を図1(c)に示す。実験結果より、ほぼ正確に体節への分割が行われていることが確認された。

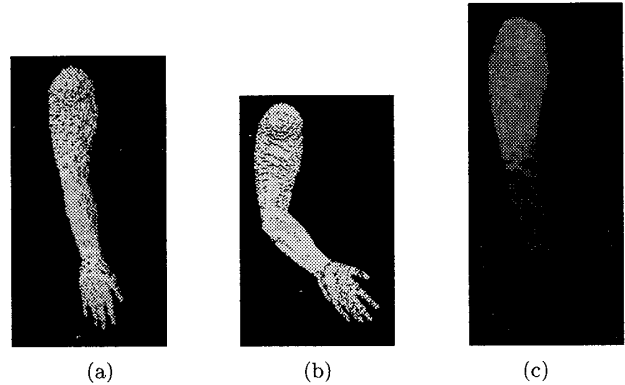


図1: 体節への分割結果

## 5. まとめ

本稿では、時系列で獲得される視体積を入力として、視体積を体節ごとに切り分け、各体節の運動を推定する手法を提案した。提案手法の利点は物体の各体節の幾何形状モデルを必要としない点であるが、物体が観測している間に多関節運動しなければ、正しく体節ごとに切り分けることができないという問題が存在している。今後の課題として、観測中に多関節運動しているとみなせる部分だけを入力データとして用いるような前処理を行い、体節の切り分け精度を向上させることが挙げられる。

## 参考文献

- [1] W.N. Martin and J.K. Aggarwal, "Volumetric Descriptions of Objects from Multiple Views," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 5, No. 2, pp.150-158, 1983.
- [2] N. Vasconcelos and A.Lippman, "Empirical Bayesian motion segmentation," PAMI, vol.23, no.2, pp.216-221, 2001.