

運動密度を用いた繰り返し現れる動作の抽出

田邊 康史(九州大学) 小川原 光一(九州大学)
倉爪 亮(九州大学) 長谷川 勉(九州大学)

Detecting Frequent Motion Patterns using Motion Density

*Yasufumi TANABE (Kyushu University), Koichi OGAWARA (Kyushu University),
Ryo KURAZUME(Kyushu University), Tsutomu HASEGAWA(Kyushu University)

Abstract— In this paper, we propose a method that detects frequent motion patterns from observation of a set of motion sequences efficiently. Frequent motion patterns are the structured information that can be obtained without knowledge about the context of motions. The major contribution of the proposed method is that the problem is formulated as a combinatorial optimization problem in which the motion label with large motion density is assigned to each time frame, and is solved via Dynamic Programming(DP) in polynomial time $O(N \log N)$, where N is the total amount of data. The proposed method was evaluated by detecting frequent interactions between objects in everyday manipulation tasks and outperformed the previous method in terms of both detectability and computational time.

Key Words: machine learning, recognition, frequent pattern detection, intelligent robot

1. はじめに

生活支援ロボットに人間の日常生活を支援させるためには、Fig.1 に示すように、ロボットが人間の動作を観測し、何をしているのか理解する仕組みが必要である。この機能の実現に向けた代表的な手法として、人間が動作の認識器を設計し、動作を学習・認識する方法がある [1] [2]。しかし、動作に関する人間の事前知識に依存したこのような手法では、システムを扱う人間の負担が大きくなる。そのため本研究では、タスクに対する事前処理なしに自動的に観測データから意味を持った動作の抽出を行う手法の開発を目的とする。

物体や動作に関する事前知識がない場合に観察によって得ることのできる情報は、物体や人間の手の運動軌跡などであり、これから意味をもった情報を取り出す必要がある。そこで本研究では、何度も繰り返し現れる動作には意味があると仮定し、観測された長時間の運動軌跡データから互いに類似する部分運動軌跡を自動的に抽出する手法を提案する。また、本研究で扱う動作は物体間の相互作用とし、類似相対運動を抽出対象とする。

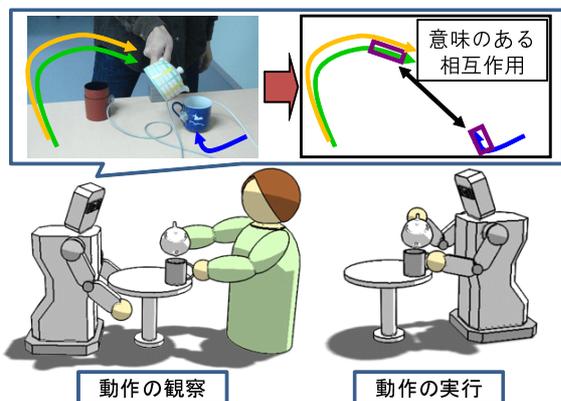


Fig.1 人間の動作を観測し、動作を学習するロボット

本論文では、この問題を組み合わせ最適化問題として定式化し、動的計画法 (DP) により平均して $O(T \log T)$ (T は観測データ数) の多項式時間で解く手法を提案する。以下の章では、まず 2 章で関連研究と提案手法の概要を述べ、3 章で提案手法の詳細を説明する。さらに 4 章で繰り返し動作の抽出実験とその結果について述べ、5 章でまとめを行う。

2. 関連研究と提案手法の概要

2-1 関連研究

本研究では、繰り返し現れる未知のパターンを時系列データから抽出する問題を扱う。この問題は、これまで生物情報学 [3]、データマイニング [4] [7]、運動解析 [8] などの分野で取り組まれてきた。

Staden はすべてのパターンの組み合わせについて投票する方法で、固定長のパターンを計算量 $O(T)$ で抽出した [3]。

Lin らは時系列データを符号列に離散化し、ハッシュ関数に投票する方法で、固定長で最も頻出するパターンを計算量 $O(T)$ で抽出した [4]。

Zhao らは認識器を構成する手法を提案した [5]。しかし、この手法ではタスクを細かく分割してしまうため長期の複雑な動作に対応できない。

森らは DP を用いる手法 [6] を提案したが、この手法では複数の物体の運動軌跡を同時に扱うことができないため、同時刻において複数の異なる物体組の相互作用が検出されてしまう可能性がある。また、これらの手法は T 個の時系列データに対して計算量が $O(T^2)$ であり、長時間のデータの処理には適さない。

計算量を削減する手法として、Yankov らは [4] の手法を拡張し、計算量 $O(T^{(1+1/a)})$ で頻出する固定長の運動軌跡を検出する手法を提案した。また、計算量 $O(T^{(1+1/a)})$ で長さが未知の運動軌跡にも対応するため、Meng らは 3 段階から構成される手法を提案した [8]。しかし、この手法では大域的な最適性を保証する

ことができない。

また、我々の以前の研究 [9] では、物体が運動していることを抽出の条件としたため、「急須のお湯をコップに注ぐ」動作などのように、動作の本質部分で速度が下がる動作が抽出できない問題があった。

そこで本論文では、新たに類似度項と把持項に基づく定式化を提案する。類似度項により、同一物体組の異なる動作を判別可能とし、把持項により、動作の本質部分で速度が下がるような動作の抽出を可能とした。

2.2 提案手法の概要

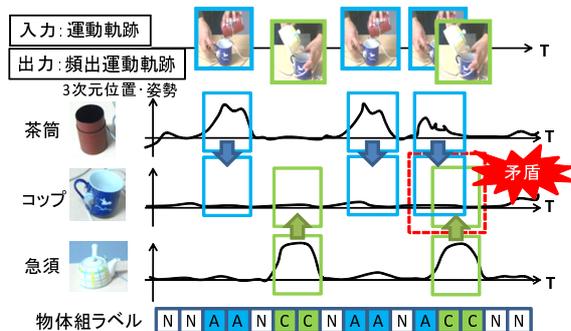


Fig.2 システムの入力と出力

複数の物体の運動軌跡に対して、独立に類似パターンの抽出を行うと、Fig.2 に示すように、同時刻において複数の異なる物体組の相互作用を抽出される。Fig.2 に示すような動作は排他的なものであり、これらの動作が同時に抽出されることは本来ありえない。

ここでは Fig.2 下に示すように、動作を行っている物体組を表す物体組ラベル (例えば茶筒とコップの組の場合は A) を時刻ごとに割り当てることにより、ある時刻において相互作用を行っている物体組を一意に推定する。さらに、このラベル割り当て問題を DP を使って解くことによって計算量を削減する。

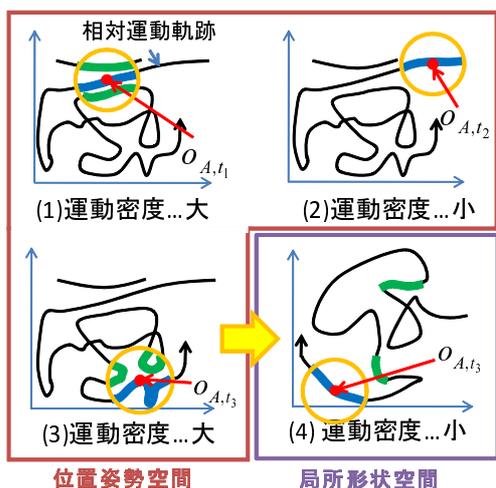


Fig.3 相対運動軌跡の運動密度

また、類似運動軌跡を探索する計算量を削減するため、Fig.3 に示す運動密度を定義する。Fig.3 中の軌跡はある物体組の相対運動軌跡を 2 次元平面に投影したものである。運動密度はある時刻の運動軌跡に対する近傍の運動軌跡の集中度を表す値であり、値が大きい時刻の運動軌跡は繰り返し動作であるとみなすことが

できる。例えば、(1) のように中心付近に軌跡が集中すると運動密度の値が高くなる。逆に (2) のように周囲に軌跡が存在しない場合や、中心付近に軌跡が集中しない場合は運動密度の値が小さくなる。

一方、同一物体組によってなされる異なる相互作用は、(3) のように位置姿勢空間中に軌跡が集中することがある。そこで、異なる動作の繰り返しによって運動密度が上昇することを防ぐため、(4) のように局所的な軌跡の形状を考慮した空間における運動密度を定義する。同じ動作であれば、局所形状空間での運動密度も大きくなるため、この 2 つの空間両方で運動密度の高い区間を繰り返し動作であるとみなす。さらに、これら 2 つの空間で見つかった軌跡同士の対応を取ることにより、同一物体組によってなされる異なる相互作用を判別する。

この運動密度の値は Kd-tree を用いて高速に求めることが可能であり、動的計画法と組み合わせることによって、全体の計算量を $O(T \log T)$ に抑えることができる。

3. 提案手法

3.1 定式化

複数物体の運動軌跡情報から繰り返し現れる相互作用を検出する問題は、Fig.2 中の物体組ラベル $X = \{\cup x_t | x_t \in \{A, B, \dots, N\}\}$ のラベル割り当て問題として定式化できる。物体組ラベルは繰り返し動作を行っている物体組を表すラベル $\{A, B, \dots\}$ とどの物体組も繰り返し動作を行っていないことを表すラベル $\{N\}$ で構成される。観測 O が与えられた時、物体組ラベル X は事後確率 $P(X|O)$ を最大化することで求めることができる。 $P(X|O)$ はベイズの定理により Eq.(1) のように変換できる。

$$\arg \max_X P(X|O) \propto \arg \max_X P(O|X)P(X) \quad (1)$$

各時刻における観測値は互いに独立であると仮定すると、 $P(O|X)$ は以下の Eq.(2) のように書き直すことができる。

$$P(O|X) = \prod P_{\text{grasp}}(o_{x_t, t} | x_t) P_{\text{density}}(O_{x_t} | x_t) \quad (2)$$

$o_{x_t, t}$ は物体組ラベル x_t と一致する物体組の時刻 t における相対運動軌跡を表す。 $P_{\text{grasp}}(o_{x_t, t} | x_t)$ は手と物体との距離が離れている状態にペナルティを与える把持項であり、Eq.(3) のように定義する。

$$P_{\text{grasp}}(o_{x_t, t} | x_t) = -\exp(-\alpha \cdot \text{dist1}(x_t) \cdot \text{dist2}(x_t)) \quad (3)$$

$\text{dist1}(x_t), \text{dist2}(x_t)$ は時刻 t における物体組ラベル x_t に一致する物体組それぞれと手との距離である。

$P_{\text{density}}(O|x_t)$ は類似運動軌跡が少ない区間にペナルティを与える類似度項であり、Eq.(4) のように定義する。

$$P_{\text{density}}(O_{x_t} | x_t) = md_p(x_t, t) \cdot md_s(x_t, t) \quad (4)$$

$md_p(x_t, t)$ と $md_s(x_t, t)$ は 3.2 節で定義する運動密度である。

$P(X)$ は物体組ラベルが変化する部分にペナルティを与える平滑化項であり、Eq.(5) のように定義する。

$$P(X) = \prod P(x_t, x_{t+1}) \quad (5)$$

$P(x_t, x_{t+1})$ は以下の Eq.(6) で定義される.

$$P(x_t, x_{t+1}) = T(x_t \neq x_{t+1}) \cdot K \quad (6)$$

K は定数である. $T(s)$ は s が真になった時に 1 となり, 偽の場合は 0 となる関数である.

3.2 運動密度の定義

本節では, 位置・姿勢空間における運動密度 $md_p(x_t, t)$ と局所形状空間における運動密度 $md_s(x_t, t)$ の 2 つの運動密度を定義する.

それぞれの運動密度 $md(x_t, t)$ は Eq.(7) のように定義される.

$$md(x_t, t) = \frac{\#motion\ segments \cdot N(o_{x_t}; o_{x_t}, \sigma)}{\sum_{x_t} \#motion\ segments \cdot N(o_{x_t}; o_{x_t}, \sigma)} \quad (7)$$

$\#motion\ segments$ は $o_{x_t, t}$ から半径 R 以内の近傍に存在する軌跡の本数である. $N(o_{x_t}; o_{x_t}, \sigma)$ は $o_{x_t, t}$ の近傍に存在する軌跡の集中度合いを $o_{x_t, t}$ を中心とした正規分布で表現したものである.

$$N(o_{x_t}; o_{x_t}, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp -\frac{(o - f(o_{x_t, t}))^2}{2\sigma^2} \quad (8)$$

局所形状空間における運動密度 $md_s(x_t, t)$ を求めるには, 運動軌跡を局所形状空間にマッピングする必要がある. このマッピングは以下の 4 つのステップに従って行う.

1. 正規化

入力データは 3 次元の位置と 3 次元の姿勢の計 6 次元からなる. これら 6 パラメータを独立に正規化する. この段階での次元数は 6 次元のままである. 位置・姿勢空間における運動密度は, この正規化したデータ $o_{norm}(t)$ を用いて求める.

2. 局所形状情報の取得

局所形状の情報を得るために, $o_{norm}(t)$ の前後 33 フレームのデータを用いて軌跡形状ベクトル $o_{shape}(t)_i$ をすべてのパラメータ i に対して求める. この段階で, 次元数は 33×6 次元に増加する.

3. ガウス関数と離散コサイン変換 (DCT) の適用

各パラメータ毎の軌跡形状ベクトル $o_{shape}(t)_i$ にガウス関数と DCT を適用することで次元数を削減する. これにより次元数を 4×6 次元に削減する.

4. 主成分分析 (PCA)

ステップ 3 で次元削減したデータの 6 パラメータ全てに対して PCA を行う. これにより, 最終的に 6 次元まで削減する.

また, ステップ 3 を実行せず, すべてのパラメータの軌跡形状ベクトル $o_{shape}(t)_i$ に PCA を行って次元数を削減することはできる. しかし, 先にステップ 3 の処理を行い次元数を削減することで, 計算量の多い PCA の処理時間を削減することが可能となる.

3.3 DP を用いたラベル割り当て問題の解法

$P(O_{x_t}|x_t)$ と $P(x_t, x_{t+1})$ が求まると, ラベル割り当て問題は, Fig.4 に示すように DP を用いることで解析的に解くことができる. また, このときの計算量は $O(T)$ となる.

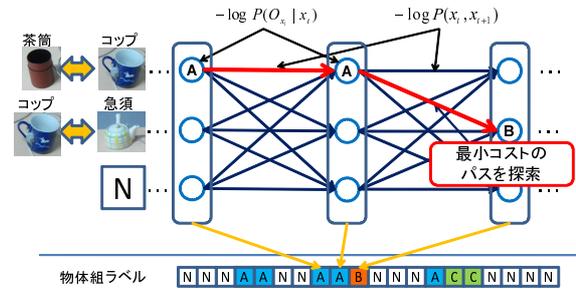


Fig.4 DP によるラベル割り当て

3.4 抽出する動作の数

前節の DP を計算するには, どの物体組も相互作用を行っていないことを示すラベル $\{N\}$ に関する尤度 $P(O|N)$ を定める必要がある. しかし, $P(O|N)$ を定義する明確な基準は存在しない. そこで, ここではあらかじめ人間が抽出する動作の種類数を指定し, 要求された数の動作が抽出されるように $P(O|N)$ の値を自動的に決定する方法を用いる. 具体的には, 要求された動作数と最も合致する $P(O|N)$ を二分探索法で探索する. DP とこの処理の合計計算量は $O(T \log T)$ となる.

4. 類似相対運動の抽出実験

4.1 複数の物体の相対運動抽出実験

本章では, 提案手法の有効性を確認するために, 机上の 3 つの物体と人間の手の運動軌跡から, 頻出する類似相対運動を抽出する実験を行った.

実験では, 米国 POLHEMUS 社製の磁気式位置計測装置 FASTRAK を使用した.

実験は, 人間の右手の中指と机上に置いた物体にレシーバを装着し (Fig.5), 取得した物体の運動軌跡から類似相対運動の抽出を行った. なお, 実験に使用した机上物体は, 急須, コップ, スプーンである.



Fig.5 実験環境

また, 動作としては, Fig.5 に示す 3 種類 (A1 ~ B2) の動作を定義し, この組み合わせで構成される以下の 3 つのデータセットに対して, 提案手法を用いて実験を行った.

- 動作 A1, B1, B2 を 6 回ずつ行ったデータセット.
- 動作 A1, B1, B2 を 6 回ずつ行ったデータセット. さらにスプーンでコップを叩く, スプーンを急須の上で回転させる, など 7 種類計 7 回のノイズデータを加えた.
- 動作 A1 を 15 回, B1 を 13 回, B2 を 12 回行ったデータセット. さらに急須をコップの上に乗せる, スプーンで急須の中のゴミを掏う, など 9 種類計 10 回のノイズデータを加えた.

今回実験を行った際の各パラメータの値は Eq.(3) の

α を 0.15, Eq.(6) の K を 10 とし, Kd-tree を用いて類似運動軌跡を探索する際の半径 R の値を位置・姿勢空間では 0.25, 局所軌跡形状空間では 0.10 に設定した.

4.1.1 実験結果

実験結果を Table 1~3 に示す.

ここでは, 把持項の有用性を確認するために, 把持項の代わりに速度項を実装した我々の以前の手法 [9] との比較を行う. 速度項では, 物体の速度が小さい状態にペナルティを与える.

なお, 表中の TP(TruePositive), FP(FalsePositive), FN(FalseNegative) はそれぞれ, 提示動作の対象物体を正しく抽出した回数, 本来抽出すべきでない動作を抽出した回数, 提示動作の対象物体を正しく抽出できなかった回数を表す. また, 表中の適合率, 再現率は式 (9) で定義する. それぞれのデータセットに対する計算時間は 15sec, 38sec, 231sec であった.

$$\text{適合率} = \frac{TP}{TP + FP} \quad \text{再現率} = \frac{TP}{TP + FN} \quad (9)$$

Table 1 理想的なデータセット 1 [4320 frames]

動作	A1	B1	B2	TP	FN	FP	適合率	再現率
提示回数	6	6	6					
把持項	6	6	6	18	0	0	1.00	1.00
速度項	6	5	6	17	1	0	0.94	1.00

Table 2 ノイズを含むデータセット 1 [7012 frames]

動作	A1	B1	B2	TP	FN	FP	適合率	再現率
提示回数	6	6	6					
把持項	4	6	6	16	2	0	0.89	1.00
速度項	3	3	6	12	6	0	0.67	1.00

Table 3 ノイズを含むデータセット 2 [16553 frames]

動作	A1	B1	B2	TP	FN	FP	適合率	再現率
提示回数	15	13	12					
把持項	15	13	0	28	12	4	0.70	0.88
速度項	1	13	0	14	26	7	0.35	0.67

4.1.2 考察

理想的なデータセット 1 に関しては, 高い精度で動作を抽出できた.

ノイズを含むデータセット 1 に関しては, 動作 A1 に関する抽出精度が低下している. これは急須のお湯をコップに注ぐ区間において, 近くに配置されていたスプーンとコップの軌跡の類似度が一時的に大きくなったことが原因と考えられる. これは, 比較的短時間のデータセットであったため起こった失敗と考えられ, 長時間のデータセットにおいては発生しないと推測される.

また, 動作 B1 については, 提案手法ではほぼ正しい抽出結果が得られたが, 以前の手法 [9] を用いた場合において一部の抽出に失敗している. これは, 動作 B1 が極めて短時間の動作であるため, 速度による影響を受けやすいことが原因と考えられる.

ノイズを含むデータセット 2 に関しては, 動作 B1 と動作 B2 の判別失敗した. これは, 動作 B2 のスプーンをコップに近づける区間が動作 B1 と類似していたため, 動作 B2 が動作 B1 と混同されてしまったことが原因だと考えられる. さらに, 動作 B1 と動作 B2 が一つの動作として検出されたため, 3 つ目の繰り返し動作として, ノイズデータの一部が誤検出された.

また, 動作 A1 については, 以前の手法 [9] ではほぼすべての区間で動作の抽出に失敗したが, 提案手法では高い精度で動作を抽出できた. これは, 「急須のお湯をコップに注ぐ」動作の本質部分である, お湯を注いでいる最中の動作がゆっくり行われたため, 速度項よりも把持項が有効に作用したことを表している.

5. まとめ

繰り返し現れる動作には意味があると仮定し, 人間の事前知識を必要とせず繰り返し運動軌跡を多項式時間 $O(T \log T)$ で抽出する手法を提案した.

本手法は, 複数の物体の運動軌跡から頻出する類似相対運動を抽出する問題を, 相互作用を行う物体組ラベル X の組み合わせ最適化問題として解くものである. このとき, 類似運動軌跡の探索を Kd-tree を用いて行い, 相互作用を行っている物体組を動的計画法で高速に求めることで, 全体の平均計算量を $O(T \log T)$ に抑える.

また, 新たに類似度項と把持項を定義することで, 同一物体組によってなされる異なる相互作用を判別可能とし, さらに動作の本質部分で速度が下がるような動作の抽出を可能とした.

実際に机上に置かれた物体を操作し, 提案した手法を適用してその 3 次元位置と姿勢情報から繰り返し相対運動を抽出する実験を行った. その結果, 頻出動作中に多くのノイズを含まない場合は, 頻出動作が高い精度で抽出可能であることを確認した.

- [1] K. Ikeuchi, T. Suehiro: "Toward an Assembly Plan from Observation Part I: Task Recognition With Polyhedral Objects", IEEE Trans. Robotics and Automation, 10, 3, pp.368-384, 1994.
- [2] Y. Kuniyoshi, M. Inaba, H. Inoue: "Learning by watching", IEEE Trans. Robotics and Automation, 10, 6, pp.799-822, 1994.
- [3] R. Staden: "Methods for discovering novel motifs in nucleic acid sequences", Computer Applications in the Biosciences, 5, 5, pp.293-298, 1989.
- [4] J. Lin, E. Keogh, S. Lomardi, P. Patel: "Finding motifs in time series", The 2nd Workshop on Temporal Data Mining, pp.53-68, 2002.
- [5] T. Zhao, T. Wang, H. Shum: "Learning A Highly Structured Motion Model for 3D Human Tracking", The 5th Asian Conference on Computer Vision (ACCV2002), 2002.
- [6] 森 明慧, 内田 誠一, 倉爪 亮, 谷口 倫一郎, 長谷川 勉, 迫江 博昭: "論理判定型 DP マッチングによる類似区間検出とその基本動作抽出への応用", 画像の認識・理解シンポジウム (MIRU2006), 2006.
- [7] D. Yakov, E. Keogh, J. Medina, B. Chiu, V. Zordan: "Detecting time series motifs under uniform scaling", The 13th Knowledge Discovery and Data Mining (ACMKDD2007), pp.844-853, 2007.
- [8] J. Meng, J. Yuan, M. Hans, Y. Wu: "Mining motifs from human motion", EUROGRAPHICS'08, 2008.
- [9] K. Ogawara, Y. Tanabe, R. Kurazume, T. Hasegawa: "Detecting Repeated Motion Patterns via Dynamic Programming using Motion Density", International Conference on Robotics and Automation (ICRA2009), pp.1743-1749, 2009.