

# 身体動作の解読：力学的な不変量による分節化

## Decoding Bodily Actions: Segmentation Based on Dynamical Invariances

日高 昇平  
Shohei Hidaka

北陸先端科学技術大学院大学  
Japan Advanced Institute of Science and Technology  
shhidaka@jaist.ac.jp

### Abstract

We can often receive information, such as physical attributes, emotions, intentions, and so on of actors, by reading others' bodily actions. A fundamental problem is decoding and symbolization of bodily movements, which are free-flowing high-dimensional temporal patterns. In this paper, I propose a formulation of this problem, called *meta-decoding problem*, by viewing it a decoding of how to decode the patterns. As a remedy of a meta-decoding problem, I have developed a statistical methodology which allows us to identify multiple underlying subsystems from multidimensional time series data. Applying the method, I show that a set of bodily actions were segmented successfully without prior knowledge on the physical body. I also discuss the possibility of this approach as a more general symbolization mechanism of bodily actions.

**Keywords** — **Embodiment, Symbolization, Fractal dimension, Nonlinear dynamical systems**

### 1. はじめに

身体動作はほとんど全ての人の活動の基礎である。動作によって我々は物理的世界に作用し、また自身の知覚のありようを変化させる。我々の得る多くの情報は、特定の身体動作と連動しており、それゆえに、逆に身体動作から動作主に関する多様な情報を得ることができる。例えば、複数の光点の動きとして提示した身体(バイオリジカルモーション)から、動作の種類、方向(Chang & Troje, 2009)などの基礎的な情報から、動作主(Troje, Westhoff, & Lavrov, 2005)、その性別(Cutting & Kozlowski, 1977; Troje, 2002)、情動(Pollick et al., 2001; Atkinson, 2009; Hobson & Lee, 1999)、など認知的な性質まで推定することが可能である。

本稿では身体動作が潜在的に有する情報(すなわち動作の“意味”)の解読の可能性について、計算論的な立場から論じる。身体は大きな自由度を

もち、複雑な相互作用を行う。従って、その結果として発生する身体運動データの意味、あるいは潜在するメカニズムを調べるのは通常容易ではない。こうした複雑なデータに潜在するある種の“同一性”は、運動に潜在するメカニズムの次元あるいは自由度によって与えられる。ある種の次元が同一である2つの運動は、ある滑らかな写像が存在し、1対1に対応付けることができる。本研究では、この次元の性質を利用して、時系列データを特性づける次元クラスタリング法(Hidaka & Kashyap, 2013)を用い、身体運動に反映される認知過程の分析を行った。

ある現象を記号として解釈するとはいかなることだろうか。本稿では、意味あるいは記号化について根本から考え直すために、まず「情報」の基礎に戻って考え直す。身体動作を含むあらゆる現象は、何らかの意味と関連付けることが可能な符号である。しかし、その意味付けは陽に与えられておらず、意味付けの方式自体を解読する必要がある。従って、この問題は情報伝達の中心に位置する情報理論の範疇を超えた議論が必要になる。

以下、まず情報理論を基礎に据える論考から、その限界を指摘し、私が呼ぶ「メタ復号問題」と呼ぶ定式化を与える。次にメタ復号問題の中心問題の解として、次元推定法を提案する。そして、身体動作からの意味推定を一種のメタ復号問題として捉え、身体動作の基礎単位として、特定の動作区間を分節化できるか検討した。最後にこうした次元推定法の分析結果の含意に触れ、身体の意味論に向けた理論構築について述べる。

### 2. 情報の3水準

Shannon & Weaver (1948)は情報の計算論において、伝達可能性、意味理解、因果性の3水準があると論じている。情報はまず正確に伝わる必要がある。誤りのない、あるいは一定の誤りを許容して、単位時間あたりの伝達可能な情報量につい

ての計算論を与えるのが Shannon の情報理論である。

一方、Shannon の情報理論は情報の意味や因果性について何も語らない。伝達された情報の意味は、狭義にはそれが符号化する何らかの対象による。情報理論では、ある技術者が符号器と復号器を同時に設計可能であるとの仮定の下、符号と復号の一貫性の問題に立ち入らない立場をとる。これは、Shannon が情報理論を定式化した情報通信の分野において、符号器・復号器は人工物であるため確かに成り立つ。

しかし、この仮定は一般には成り立たない。つまり、我々など「非人工物」の“通信”において、情報の送り手(符号器を持つ側)が、情報の受け手(復号器を持つ側)がいかに復号するか、あるいはその逆を事前に十全に知りえることはありえない。いや、日常我々は、慣習や自然言語など、様々な事前に取り決められた“通信ルール”を共有している、と考えられるかもしれない。しかし、我々はこうした慣習あるいは言語をどのように確立してきたのであろうか。あるいは、こうした事前知識を持たないと想定される乳・幼児はどのように“通信ルール”を推測すればよいのだろうか。こうした問題を考えるとき、やはり事前の取り決めを前提とせず、しかし未知なる“通信のルール”自体を、通信そのものから推定する、という問題に行き着く。

### 3. メタ復号問題：復号器の復号

こうした未知なる“通信のルール”自体を、通信そのものから推定する問題をメタ復号問題と呼ぼう。メタ復号問題とは、情報の受け手は、与えられた復号方式によって符号を意味づけるのではなく、符号列から、その符号列クラスの復号方式を推定する問題である。

メタ復号問題の第一の問いは、まずそれが可能であるか、あるいはどのような復号器の性質であれば復号可能であるか、である。

まず、ある種の同型性を超えて、復号器を同定するのは困難であり、また必要がない。例えば、ある送り手が知覚した“赤”と“青”を、「アカ」と「アオ」という符号に置き換えて、受け手に送ったとする。受け手は符号「アカ」を、受け手の知覚する‘青’に、符号「アオ」を知覚‘赤’に割り当てて復号したとする。このとき、もし知覚現象“赤”と“青”、‘赤’と‘青’が対称で交換可能である場合、この二つと符号には識別可能な2つの組み合わせが存在する。こうした冗長性を

受け手が解消するには、送り手の知覚“赤”・“青”を直接参照する以外に、実質的に不可能である。しかし、二つの符号化・復号化には本質的な違いがなく、同型である。従って、メタ復号問題において、符号のアルファベットがあり、そのうち交換可能な文字列に関する冗長性を超えて復号器を同定する必要はない。つまり、メタ復号の対象は符号交換を同一視した商位相空間である。

この議論から、メタ復号問題では、具体的な復号方式そのものではなく、特定の復号器のクラスが、復号可能な符号列の集合において「どのような同型性」が保たれているか、と言う問題が本質であることが分かる。

同型性をどのように分類したらよいだろうか。これを端的に表す一つの性質は、符号列のクラスを表現するための「アルファベットサイズ」である。前の例を使うと、知覚の集合{“赤”, “青”}を区別するためには最小で2文字、もし虹の7色{“赤”, “青”, “緑”, ...}を区別するとすると、3文字必要である。あるいは、光学的なスペクトルを表現する場合、(どの程度の誤差を許容するか、あるいはスペクトル上の色の出現頻度によるが、)これに対応する1次元を表現する文字群が必要となる。明らかに、最小サイズより小さいアルファベットで対象を網羅的には識別できない。つまり、符号のアルファベットサイズは、それが表現する対象クラスの本質的な制約を反映している。

### 4. メタ復号としての次元推定

こうした議論から、このアルファベットサイズとは、対象を十分に表現するための符号の「自由度」であり、あるいは表現すべきデータの集合が張る空間の「次元」と言える。この「次元」の有効な候補として、数学的には点次元(Young, 1982)と呼ばれているものが挙げられる。本稿では技術的な詳細に立ち入らず、その骨子のみを以下に述べる。

Hidaka & Kashyap (2013)は、2つの可測な距離空間の同値性を、なめらかな写像で対応付け可能な集合で定義し、この同値性の下で、空間の特性量として点次元が導出できることを示した。つまり、特異的ではない変換で対応付けられる全ての可測距離空間は、同一の点次元を持ち、また逆も真である。ある空間の次元は特定のクラスの写像に対して不変である事は良く知られているが、Hidaka & Kashyap (2013)はその逆を示している。

この議論から、前述のメタ復号の第一問題は、符号列(データ)から点次元を推定することである、

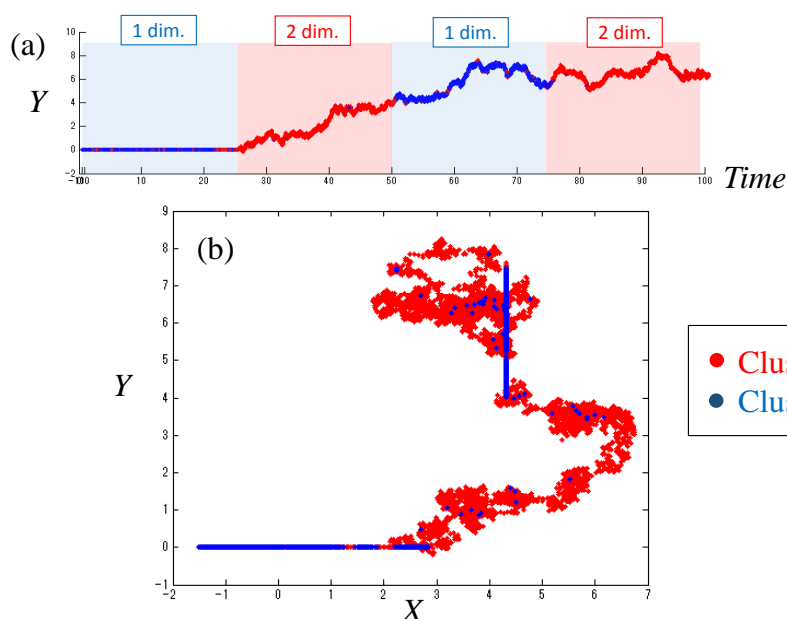


図 1: 1次元/2次元ランダムウォークの混合データに対する次元推定法の分析例. (a) Y 軸上の時系列, (b) 2次元上のランダムウォークの各点の色は2つのクラスターを表す.

と言い換えることができる。Hidaka & Kashyap (2013; 2014)は点次元を推定する新たな手法として、従来の方法論とは異なり、混合分布の統計的推論による方法を提案している。これを以下、点次元推定法と呼ぶ。

点次元  $d$  は、空間上のある点  $x$  の半径  $r$  の近傍の測度  $B(x, r)$  で以下のように定義される次元である：
$$d = \lim_{r \rightarrow 0} \log B(x, r) / \log r$$

この点次元の数学的に厳密な定義は、無限小の極限を要するため、直接計算することはできない。従って、データから何らかの近似・推定をする必要がある。点次元推定法では、データの各点の十分に大きい近傍内において、一定の密度で(観察されていない)点が分布すると仮定する。この仮定から、密度  $\lambda$ 、次元  $d$  の空間上の任意の点に対し、 $n$  番目に近い点までの距離  $r$  は、Weibull-gamma 分布  $P(r | n, d, \lambda)$  (Hidaka, 2013)になる事が示される(Hidaka & Kashyap, 2013)。従って、この分布のパラメータ  $(d, \lambda)$  を  $n$ -近傍距離データから推定することで、推定されたパラメータ  $d$  を次元とみなすのが Hidaka & Kashyap の次元推定法の基本的なアプローチである。この発展形として、混合 Weibull-gamma 分布を用いることで、データに混合する複数の異なる次元成分を検出できる。

## 5. シミュレーション

提案する次元推定法の検証のため、1次元/2

次元の切り替えながらランダムに生成される時系列(ウィーナー過程)であると想定した人工データを分析した(図 1(a))。生成した 10,000 点のデータのうち、1-2500 点は X 軸のみ、2501-5000 点は 2 次元上、5001-7500 点は Y 軸のみ、4 番目の 7501-10,000 点は再度 2 次元上のランダムウォークである。図 1(a)はそのデータの Y 軸上の時系列、図 1(b)は(X, Y)平面を示す。この 2 次元系列(X, Y)に対し、次元クラスタリングを適用した次元推定の結果を、各点の赤/青色で示している。この結果から、次元推定法により潜在する次元の違いを正しく推定できることが示された。

## 6. 身体動作の分析

本研究では、多様な身体動作に対して、動作単位の分節化が可能か検討した。動作の分節化とは、連続的な身体動作データから、意味の異なる動作の区間を同定することである。この問題は、身体に潜在する「行為」という情報の単位を、その動作だけから推定する一種のメタ復号問題と考えられる。

周期性を前提としない一般的な動作で、特定の教師信号に基づく学習を使わず、動作を分節化するには各種の技術的な困難性がある。これは、同

時に動き、互いに時空間的な相関を持つ身体動作のデータから、その多次元状態空間上の動作を簡潔に表現する、より本質的な低次元の特徴空間を発見する問題として捉えられる。

具体的には、Ma ら(2006)の身体運動のデータベースを用い、全身 15 点の特徴点で捉えた 3 種類の右手を主に使った運動を分析した。図 1 は全身の 15 点の特徴点(A)の動き(15×3次元=45次元)を第一・二主成分で表現した(B)ノック、(C)持ち上げ、(D)投げの動作を示す。各軌道は 5 回連続で“同じ”動作を繰り返しているが、試行によってばらつきと、類似性が見られる。従って、動作の認識では、速度や軌道パタンの異なる(冗長な)複数の動作を「ノック動作」と認識する必要がある。

本研究では、このような動作分節化・認識の問題に対し、身体動作の生成系に固有な不変量の推定に基づく分節化方法を提案する。提案手法では、身体動作を生成する系が、動作の単位ごとに変化すると考え、これを検出することで身体動作を分節化する。この系の変化点は、身体動作を一種の非線形系とみなした場合、身体動作の時系列データから、動作単位・動作種に固有な不変量である局所的なアトラクター次元の変化を推定する事に対応する。図 1(A,B,C)の各データ点のうち太い円で強調された動作の区間は、Hidaka & Kashyap (2013)の点次元推定法で検出された次元の変化点である。

身体動作の時系列の各点に次元推定法を適用した場合、典型的には図 2 に示すような結果を得られた。図 2 は図 1(B)の時系列プロットで、全身 15 箇所につけられたマーカーの 3 次元空

間上の座標の時系列(計 45 次元)を示し、次元が同一と推定された点には同じ色づけをした(次元の変化点を下向き矢印で強調)。この例では、上記のデータベースに含まれる特定の被験者がノックの動作を計 5 回繰り返しており、人が直接動画を見て判断をしたノック動作の切れ目が時間軸上の縦線(開始点：太線、終点：細線)で示されている。この身体動作データベースには 30 人の動作者がそれぞれ手の動作 3 種類、歩行、歩行しながらの手の複合的な動作の計 5 種類の動きを複数回繰り返したデータが収録されている。このデータベースのうち、手の動作はノック、投げ、持ち上げ、の 3 種類で、いずれも単純な周期性の動作ではない(ノックに関してはノック中(3,4 回の腕振り)の周期と開始・終わりの 2 重の周期性がある)。こういった複数種類の動作や異なる 30 の動作者に対し、約 85%の割合(全 736 試行中の 632 試行)で、推定された異なる次元成分(クラスター)と動作の分節に統計的に有意な関係があり、動作の開始点と終点を一定の誤差の範囲で検出できることを確認した。一方で、一部のデータでは分節化に失敗する場合があります、こうした問題の解決にはさらなる最適化が必要であると考えられる。

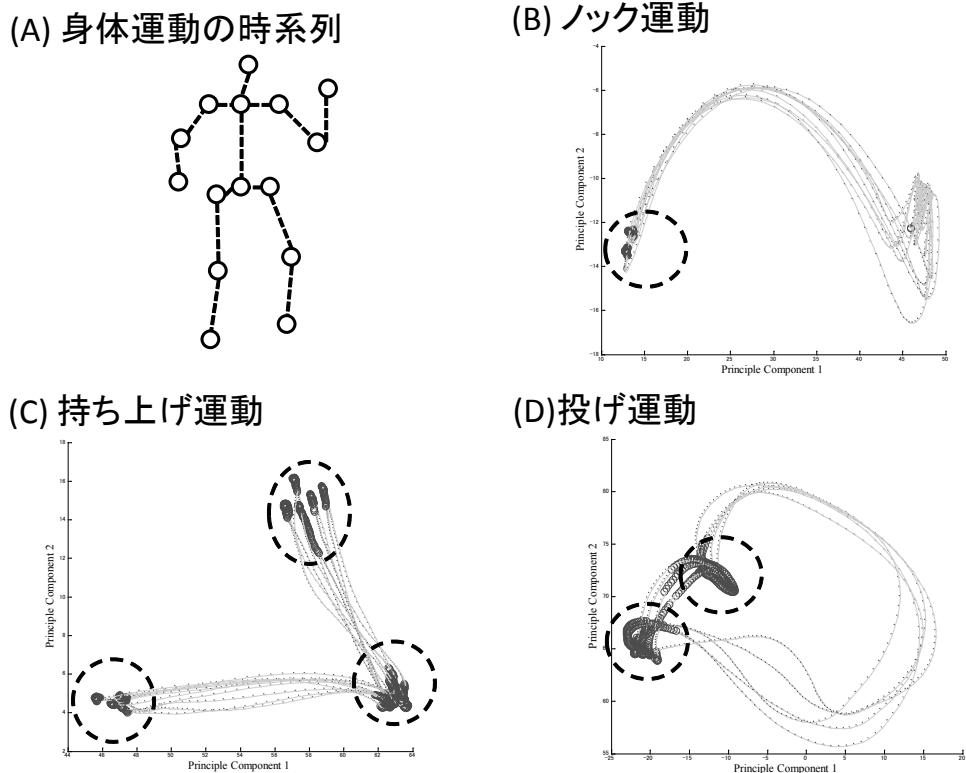


図 1: 動作の不変量による特徴づけ

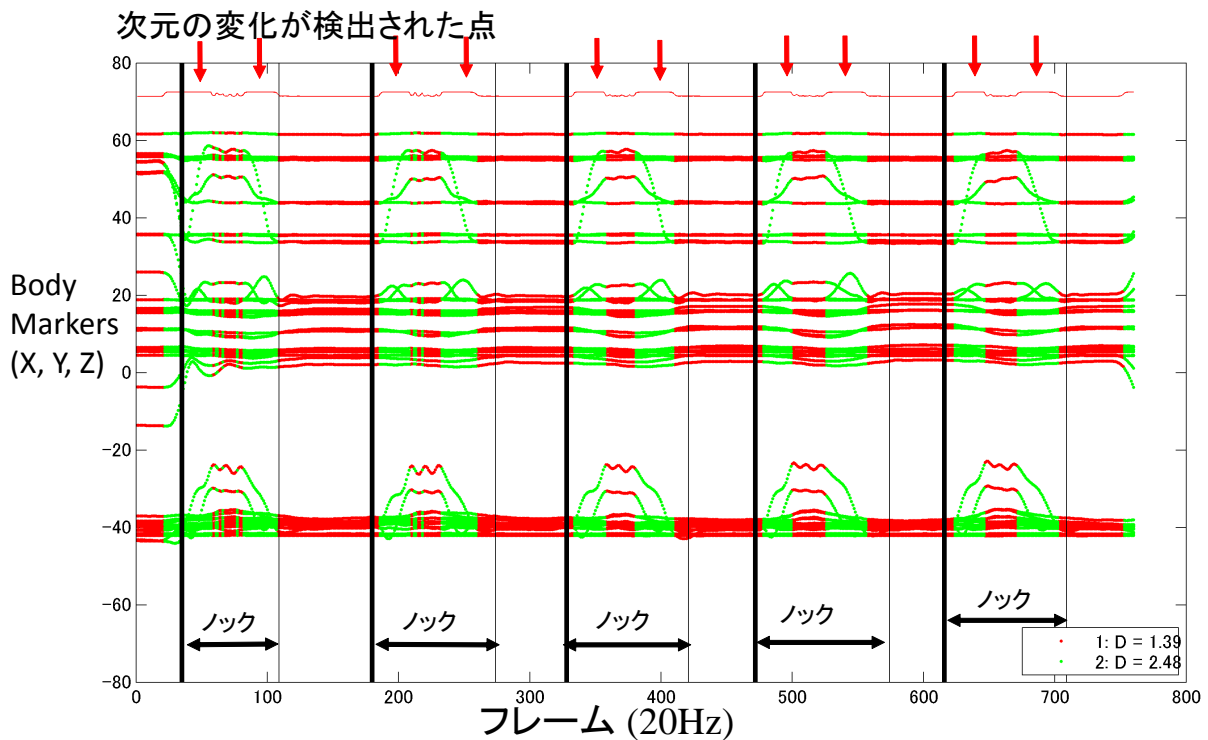


図 2: ノック動作の分節化の一例(図 1(B)も参照)

### 7. 結語：身体動作の記号化

多くの身体動作は連続的におこなわれ、明確な切れ目を持つことは多くはない。しかし、我々はいとも簡単に、動作を分節化、さらには記号化する

ることができる。なめらかな動作に埋め込まれた、特異性はいかなる性質を持つのだろうか。本稿で示した分析結果から、一見連続的でなめらかに見える動作の開始と終わりには、特異的に異なる次

元成分があることが示唆された。

この分析で用いた次元推定法および分析手続きには、動作に関する固有の事前知識が含まれていない。それにも関わらず、多次元上の時系列として表現された身体動作のみから、こうした動作の開始と終わりの符合と解釈可能な成分が検出されたことは、驚くべき結果であると私は考えている。

意味が発生するには、モノ・コトの識別が必要であり、その識別・分節化にはモノ・コトではない「モノ・コト」が必要である。身体動作において、複数の動作レパートリーを連続的にかつなめらかにつなぎ合わせても、そこには「音」、「音節」、「単語」、「文」、あるいは「文章」が自然に構造化される。本研究は身体動作の「単語」の分節化の可能性を示唆する一分析を報告した。身体動作の意味論を論じるための基礎は着実にできつつある。

## 謝辞

本研究は科学研究費補助金 挑戦的萌芽研究 No. 25560297 の補助を受けて行われた。

## 参考文献

- [1] Atkinson, A. P. (2009). "Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders". *Neuropsychologia*, 47, 3023-3029.
- [2] Chang, D. H. F., & Troje, N. F. (2009). "Acceleration carries the local inversion effect in biological motion perception". *Journal of Vision*, 9, (1):19, 1-17
- [3] Cutting, J. E., & Kozlowski, L. T. (1977). "Recognizing friends by their walk: Gait perception without familiarity cues". *Bulletin of the psychonomic society*, 9(5), 353-356.
- [4] Hobson, R. P. & Lee, A. (1999). "Imitation and Identification in Autism", *Journal of Child Psychological Psychiatry*, 40, 4, 649-659.
- [5] Hidaka, S. & Kashyap, N. (2013). On the

"Estimation of Pointwise Dimension" ., eprint arXiv:1312.2298.

- [6] Hidaka, S. & Kashyap, S. (2014). The Generalist Approach to Frame Problems., In *Proceedings of The Third Asian Conference on Information Systems*, 318-325.
- [7] Ma, Y., Paterson, H. M., Pollick, F. E. (2006). "A motion capture library for the study of identity, gender, and emotion perception from biological motion", *Behavior Research Methods*, 38, 1, 134-141.
- [8] Pollick, F. E., Paterson, H., Bruderlin, A. & Sanford, A. J. (2001) "Perceiving affect from arm movement". *Cognition*, 82, B51-B61.
- [9] Shannon, C. E. & Weaver., W. (1949). *The Mathematical Theory of Communication*. Univ of Illinois Press.
- [10] Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2, 371-387.
- [11] Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: effects of structural and kinematic cues. *Perception & Psychophysics*, 67 (4), 667-675.
- [12] Young, L. S. (1982). Dimension, entropy and Lyapunov exponents. *Ergodic theory and dynamical systems*, 2(01), 109-124.