

動的線形システムに基づく動作生成の入力成分推定

寺門 修平^{1,a)} 栗山 繁^{1,b)}

概要: モーションキャプチャデータを用いた動作の生成や変換に、動的線形システムを用いた手法が提案されている。既存の手法では、モーションデータを用いてこの動的線形システムを同定する際に、入力信号は指定しないか三角関数などの簡単な信号を用いている。これらの手法は歩行動作などの周期的な動きに関しては有効であるが、非周期的な動作に関する有効性は不明である。そこで本研究では、同類の動きを有するモーションデータ群から基準となる入力信号を「動きの本質的成分」として抽出して入力信号とすることにより、その動きの生成や変換に最適な動的線形システムを構築する手法を提案する。

1. はじめに

3DCGの動きを表現するデータはモーションキャプチャ (Motion Capture; MoCap) や手付けモーションといった様々な方法で作成される。これらのデータはそのままキャラクターモデルへと適用し、動作させる分には問題ないが、よりダイナミックな動きへの変換やこれらのデータから別の動作を生成することを想定した場合、そのままのデータでは扱いにくい。式 (1) のような動的線形システム (Linear Dynamical System; LDS) を導入した手法は状態 x_t と入力 u_t の線形和で出力を決定するモデルであり、既存の手法では、LDSの出力としてそのモーションの姿勢時系列を、入力として任意の信号かあるいは未指定として常に零の信号をそれぞれ与えてシステム同定を行う。システム同定を利用すれば入出力信号や時間間隔、同定するモデルの指定程度のみで容易に任意のLDSモデルを同定、構築することができる。

$$\begin{cases} x_{t+1} = Ax_t + Bu_t \\ y_t = Cx_t + Du_t \end{cases} \quad (1)$$

しかしながら、モーションデータの表現におけるLDSの入力信号は三角関数などの簡単な信号や指定しないことがほとんどであった。これらの場合でも歩行動作などの周期的で単調な動きであれば有効であったが、周期的ではないより複雑な動作の場合、入力として適切な波形が存在するものと考えられる。そこで本研究では、LDSにおける入力信号に着目し、その動きの生成や変換に最適な動的線形シ

ステムが構築されるような入力信号を、同類の動きを有するモーションデータ群から抽出することを目的とする。

2. 関連研究

モーションの生成・変換にLDSを用いている既存手法の一例としてLiらの手法 [1] がある。この研究では、1つのモーションを複数の部分モーションに切り分け、これら部分モーション同士を自由に経路設計し連結することで新たなモーションを構築する手法を提案しており、部分モーションの表現にLDSが用いられている。しかし、この手法におけるLDSは状態のみで出力を決定する、すなわち式 (1) において $B=0$ 、 $D=0$ としたようなモデルを想定しており、実質的に入力を指定していない。また、(1) のような行列ではなく、より高次の概念であるテンソルで定義したLDSを用いた手法としてDingらの手法 [2] がある。こちらは動きを識別するための学習過程でテンソル形式のLDSを使用しているが、テンソルとして内包される情報自体は行列より多いものの、やはり入力を指定しないモデルを採用している。

直接モーションを表現するLDSにおいて入力信号を考慮したものはあまり見られず、より適切な入力を得ることは「動きの本質的成分」を抽出することにつながることから、より柔軟なモーションの生成や変換に有効に働くものと考えられる。

3. 提案手法

3.1 全体のフロー及び概要

本研究で提案するシステムの概略図を図1に示す。これは、入力信号の生成とLDS同定のシステムであり、本研究では特に入力信号の生成に着目する。

¹ 豊橋技術科学大学
TUT, Toyohashi, Aichi 441-8580, Japan

a) terakado@val.cs.tut.ac.jp

b) sk@tut.jp

LDS の同定では計算負荷の軽減を目的として、図 2 のように骨格モデルを大まかなブロックに分割し、それぞれで LDS 同定を行う。その後、同定した各 LDS の係数行列を再結合することで全体の LDS とする。骨格モデルをブロックに分割する手法自体は既に多くの研究 [3], [4], [5] で用いられており、各ブロックの LDS を結合する際には出力が元のモーションの姿勢ベクトルと対応するように結合する。

3.2 入力信号の生成

LDS における入力信号は出力と同一の系列数であれば次元数に関係なく任意に指定できる。本手法では基本的に対象とするモーションの姿勢ベクトルと同一の次元数の信号を与える。これは、入力が本質的にはモーションデータと類似した挙動を示すものと仮定しているためである。

3.2.1 中央値モーション

モーションを表現する LDS の入力信号として、まず中央値モーションを提案する。中央値モーションとは、ある同類の動きを有するモーションデータ群に属する各データとその他のデータとの距離の総和が最小となるモーションデータである。

中央値モーションをそのモーションデータ群における基準モーションとし、これを同一のモーションデータ群に属する LDS 同定対象のモーションに合わせ時間伸縮 (Dynamic Time Warping; DTW) を用いることで系列数条件を満たした入力信号とする。

3.2.2 平均モーション

次に、ある同類の動きを有するモーションデータ群の中の平均を表す、平均モーションを提案する。

平均モーションの生成方法では、まず基準モーションとして中央値モーションを求める。その後、モーションデータ群に属するすべてのモーションを基準モーションに合わせて時間伸縮により基準モーションと同一の系列数に変換

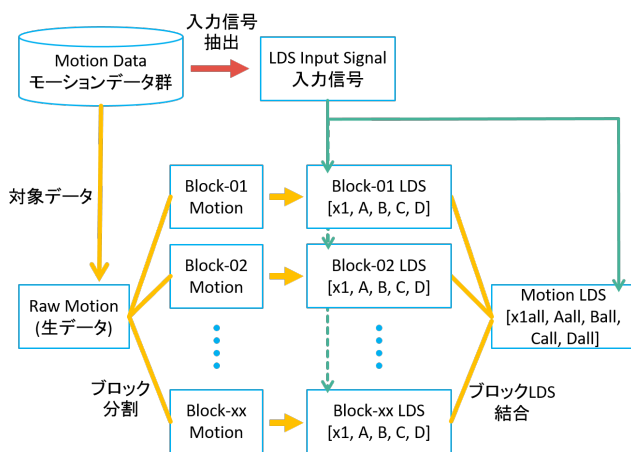


図 1 入力信号の生成及び LDS 同定のフロー

Fig. 1 Flow of generate input & system identification

し、各系列ごとに平均ベクトルを算出することで平均モーションとする。入力信号として使用する際には中央値モーションと同様に、対象のモーションデータに合わせて再度、時間伸縮することで系列数を合わせる。

3.2.3 姿勢間距離

低次元の入力信号として姿勢間距離を用いる手法も提案する。姿勢間距離はルート関節の位置を原点とした各関節の位置を求め、各姿勢の初期姿勢との残差ノルムをとったものであり、モーション間の非類似度を一つの波形として表現する手法の一つである。

本研究ではそれぞれ対象とするモーションデータにおいて、骨格モデルを構成する全関節を対象として姿勢間距離を計算し、最大値が 1.0 になるように正規化した一次元のデータを入力信号として用いる。

3.2.4 数学関数信号

検証における比較対象として、使用されやすい入力信号の一つである正弦波信号を用いる。正弦波信号は 1 次元のデータとするが、周期および波数は対象とするモーションに合わせて可変させる。

4. 検証実験

4.1 実験方法

様々な入力信号を仮定した LDS の評価実験として、実際にモーションデータを用いて LDS の同定、及び復元を行う。実験において利用するモーションは、本研究室で過去に作成した単純動作モーションデータより作成した歩行動作、壁押し動作、荷移動動作のデータ群 3 種で、各動作は 4 人の演者のデータファイルから構成される。また、これらモーションデータは、1 フレーム当たりルート関節の 3 次元位置及びルート関節を含む 30 関節の 3 次元オイラー回転角、計 93 次元のベクトルデータとして表現される。

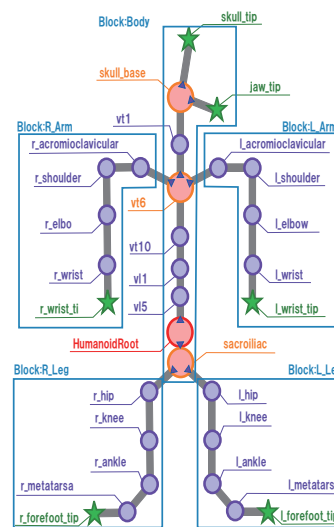


図 2 骨格モデルのブロック分割例

Fig. 2 Block subdivision of skeleton model

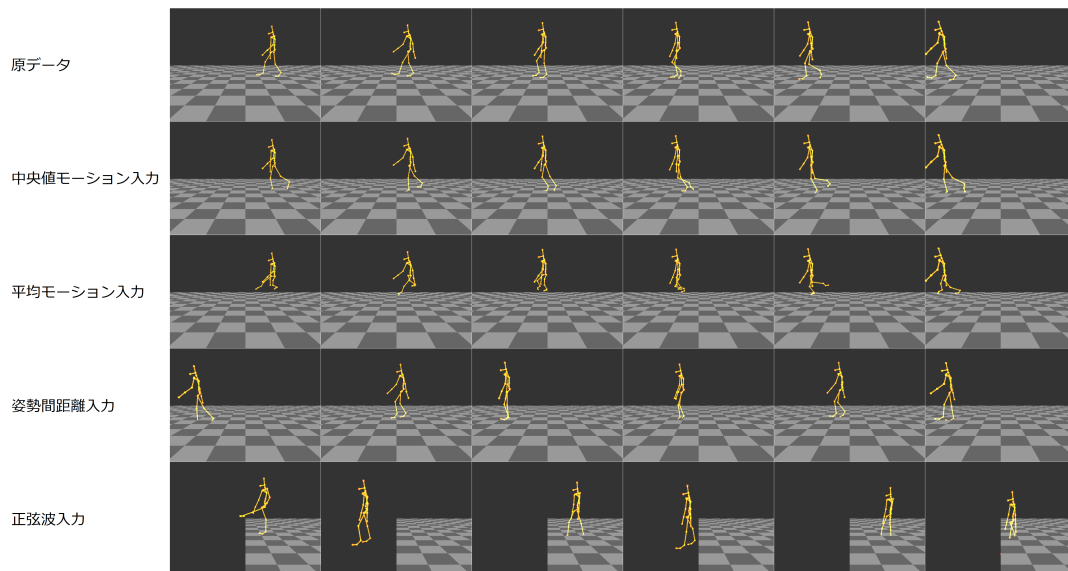


図 3 歩行動作の動画比較結果

Fig. 3 Video comparison: "Walk" motion

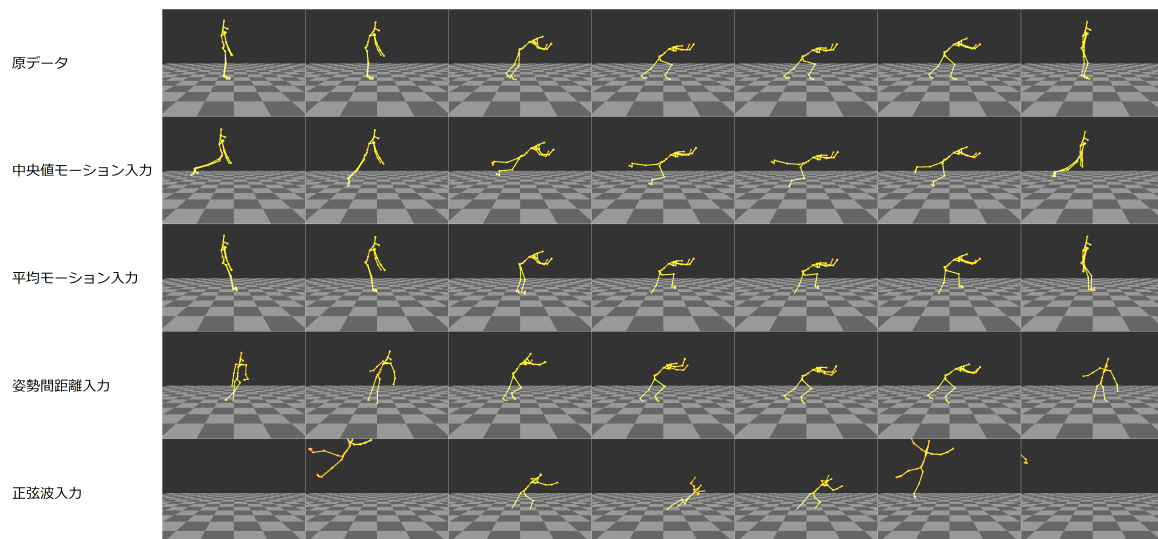


図 4 壁押し動作の動画比較結果

Fig. 4 Video comparison: "Push" motion

それぞれのデータ群において、3.2節で示した各手法で入力信号を生成し、すべてのモーションデータのLDS同定を行う。その後、モーションへ再度復元して動画比較及びフレームごとに全関節の回転角情報の平方根平均二乗誤差(RMSE)を計算し数値評価を行う。

4.2 実験結果

4.2.1 動画比較

動画比較結果の一例として、周期動作である歩行動作における結果と非周期動作である壁押し動作における結果をそれぞれ、図3、図4に示す。図はどちらも上から順に原データ、中央値モーション入力のLDS復元結果、平均モーション入力のLDS復元結果、姿勢間距離を入力としたLDS復元結果、および正弦波入力のLDS復元結果であ

り、横には図3の歩行動作は50フレームごと、図4の壁押し動作は100フレームごとの姿勢を並べている。また、カメラ位置は歩行動作における正弦波入力の結果以外、固定しているため、見え方の違いがそのまま復元差異として判別できる。

これらの結果を見ると、既存手法で用いられる正弦波入力の場合、キャラクターの位置、すなわちルート関節の3次元位置のずれが大きいものの、映っている範囲の姿勢のみで評価すればどちらの場合でも再現できている。一方で、本研究で提案した各手法では位置はどれも比較的一致しているが、中央値モーションを入力とした際には非周期動作である壁押し動作において脚部が大きくずれた姿勢になっており、平均モーションでは周期動作である歩行動作において脚部が画面奥方向に曲がってしまっている。

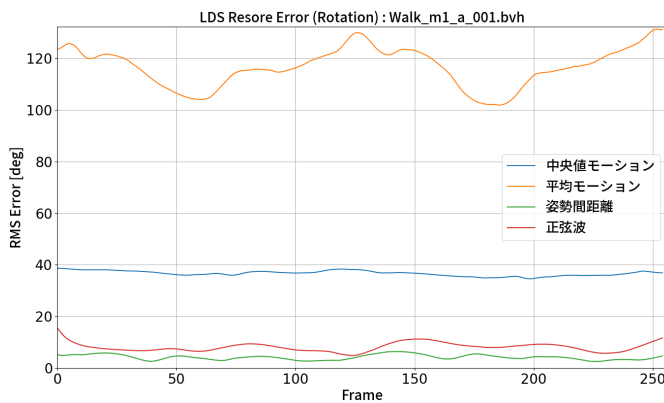


図 5 歩行動作における全関節の回転角の誤差
Fig. 5 Rotation RMS Error of "Walk" motion



図 6 壁押し動作における全関節の回転角の誤差
Fig. 6 Rotation RMS Error of "Push" motion

4.2.2 姿勢誤差による評価

動画では位置による誤差が大きく目立ち、姿勢すなわち関節角に関しての誤差を明確に把握することが困難であるため、時系列として各フレームでの全関節の回転角について原データからの誤差を測定した。それぞれ図 3 の歩行動作と対応したものを図 5、図 4 の壁押し動作と対応したものを図 6 に示す。どちらのグラフも縦軸に RMSE、横軸にフレーム番号をとり、縦軸の単位は度 (deg) とする。

各グラフを見てみると、4.2.1 節でも見られたように、歩行動作において平均モーシオンの平均誤差は 120 度前後と非常に大きく、壁押し動作では中央値モーシオンの平均誤差が 250 度以上となっており、復元精度は非常に悪い結果となった。一方で、どちらの場合でも姿勢間距離の場合には精度が良く、正弦波を入力とするものより高い復元精度であった。

4.3 実験考察

実験結果を総合すると、本研究で提案した中間値モーシオン、平均モーシオンは周期的な動作に対する復元精度が共に低く、中間値モーシオンにいたっては非周期的な動作においても悪い結果であった。非周期的な動作に対する平均モーシオンの復元精度も低くはないものの、正弦波入力と同程度であり、効果的とは言えない。一方で姿勢間距離の場合はどちらの動作に対しても最も精度が良く、本実験で用いた手法の中では最も効果的に動作を表現できる入力信号であったと言える。このような結果となった原因として、いくつかの要因が考えられる。

まず一つは、正規化の有無である。本実験では、正弦波と姿勢間距離の信号は最大が 1.0 となるように正規化したものを採用する一方、中央値モーシオンと平均値モーシオンはそのまま信号として使用した。そのため、後者の場合、LDS の係数行列によって信号値が一律に定数倍されるため、誤差がより強調されてしまったと考えられる。

もう一つの要因として、入力信号の波形が考えられる。

付録 A.1 の各波形を比較し、4.2.2 節で誤差が大きかったものに注目していくと、歩行動作における平均モーシオン入力の場合、各信号が原データよりも全体的に上下に移動しており、LDS の係数行列で吸収されにくい定数的なずれが生じているため、誤差が増大したと考えられる。また、壁押し動作の場合、中央値モーシオン、平均モーシオンの両方において信号が不連続的に変化する部分が多く生じており、特に中央値の場合では信号の概形も一致せず、値も大きく異なっていたため、復元時の誤差が大きかったと考えられる。一方で誤差が小さかった姿勢間距離の場合は、正弦波とは異なり常に正の値をとること、非周期的な動作の場合であってもこれに対応して柔軟に信号の概形が変化することが有効に働いたと考えられる。

5. おわりに

5.1 まとめ

本研究ではモーシオンキャプチャデータを用いた動作の生成や変換に用いられる動的線形システムにおける入力信号の抽出方法を提案した。実験結果を踏まえると、現状で最も有効な入力信号は姿勢間距離を用いたものであり、これは歩行動作などの周期的な動作以外でも安定して再現でき、有効性を示せた。

また、モーシオンデータの表現方法として LDS を用いた際の入力信号の依存性について評価した。LDS の入力信号による差異は関連研究では触れられていなかったが、出力を同じモーシオンデータとしても入力信号によって復元結果が大きく変化していたため、入力信号を適当に選択する必要性が示唆された。

5.2 今後の課題

本研究ではルート関節の位置を含めたモーシオンデータの復元を想定して実験を行った。しかし、ルート関節については姿勢ベクトルの中でも特異的なものであり、これらを除外することで 4 章の実験では精度が低かった中央値

モーションや平均モーションについても精度を改善できる可能性があるため、この条件での実験・検証を行う必要がある。加えて、実験に用いたモーションデータ群の種類は少なく、単純なものであったため、より複雑なモーションデータの場合の優位性についても検証が必要である。

また、本研究では入力信号に主眼を置いていたが、LDSは状態変化をモデル化したものであり、入力信号が「動きの本質的成分」であったとすると、状態変数が何らかの意味を持っていると考えられるため、このことについても検討の余地がある。

参考文献

- [1] Yan Li, Tianshu Wang, and Heung-Yeung Shum. Motion texture: A two-level statistical model for character motion synthesis. *ACM Trans. Graph.*, Vol. 21, No. 3, pp. 465–472, July 2002.
- [2] Wenwen Ding and Kai Liu. Learning linear dynamical systems with high-order tensor data for skeleton based action recognition. *CoRR*, Vol. abs/1701.03869, , January 2017.
- [3] Zhao Wang, Yinfu Feng, Shuang Liu, Jun Xiao, Xiaosong Yang, and Jian J. Zhang. A 3d human motion refinement method based on sparse motion bases selection. In *Proceedings of the 29th International Conference on Computer Animation and Social Agents*, pp. 53–60, May 2016.
- [4] Shu-Juan Peng, Gao-Feng He, Xin Liu, and Hua-Zhen Wang. Hierarchical block-based incomplete human mocap data recovery using adaptive nonnegative matrix factorization. *Comput. Graph.*, Vol. 49, No. C, pp. 10–23, June 2015.
- [5] Jun Xiao, Yinfu Feng, Mingming Ji, Xiaosong Yang, Jian J. Zhang, and Yueting Zhuang. Sparse motion bases selection for human motion denoising. *Signal Process.*, Vol. 110, No. C, pp. 108–122, May 2015.

付 録

A.1 評価実験における各入力信号波形

4章の実験における原データ及び本研究で提案した手法による各モーションへの入力信号波形をそれぞれ示す。歩行動作における原データと中央値モーション及び平均モーション、および姿勢間距離の各波形はそれぞれ図 A.1, 図 A.3, 図 A.5 であり、同様に壁押し動作における原データと、中央値モーション及び平均モーション、姿勢間距離の各波形はそれぞれ図 A.2, 図 A.4, 図 A.6 である。姿勢間距離を除き、各データは93次元の信号であるため、主要関節の信号としてルート関節, 左肩, 左肘, 左股関節, 左膝の5つに対応した3次元回転角の信号のみを抽出している。

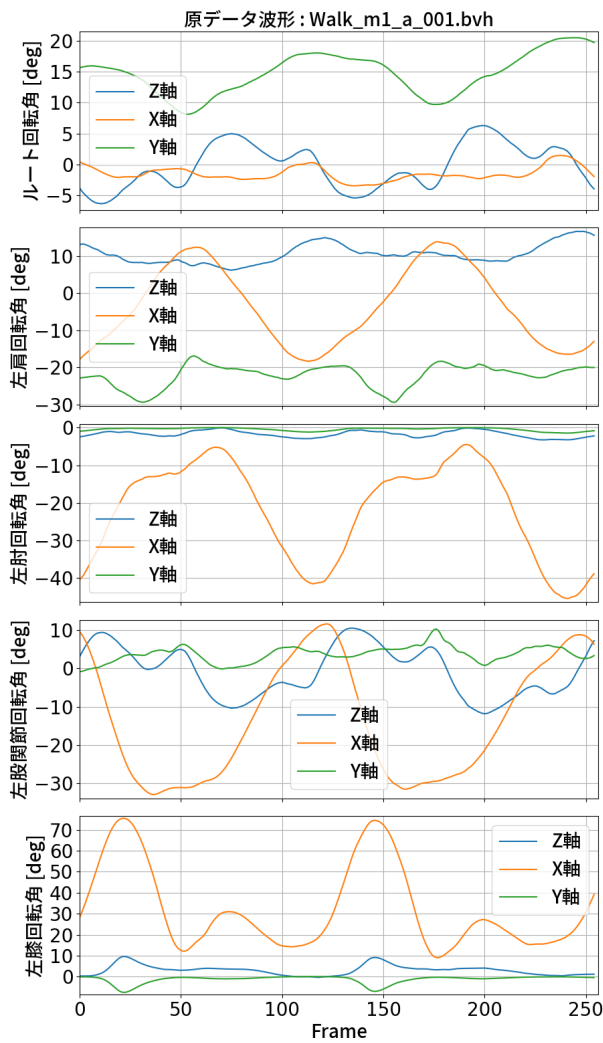


図 A.1 歩行動作における原データの波形
 Fig. A.1 Original motion signals: "Walk" motion

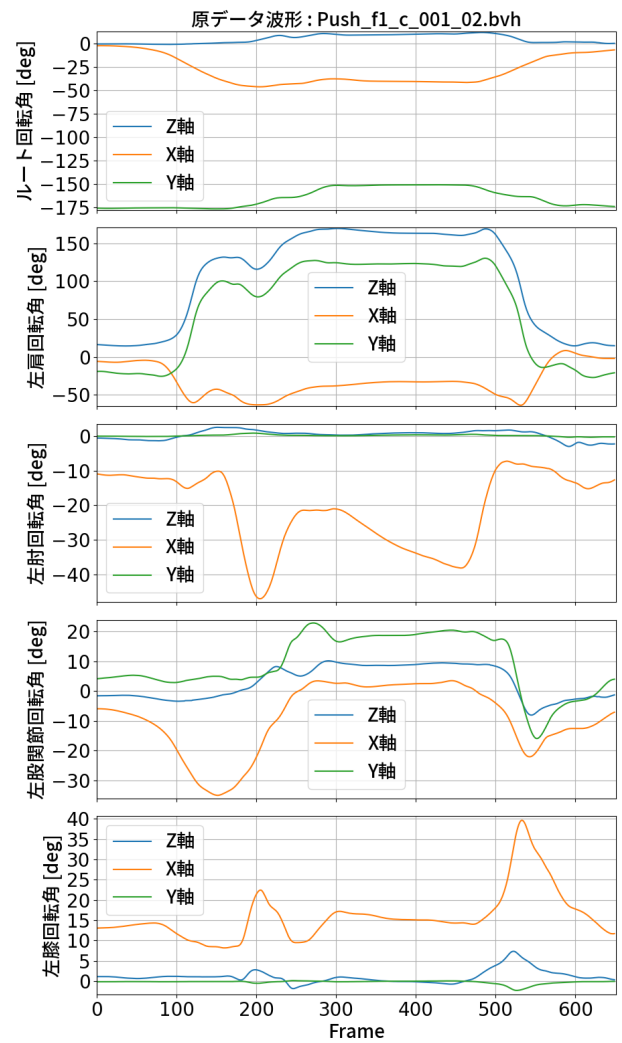


図 A.2 壁押し動作における原データの波形
 Fig. A.2 Original motion signals: "Push" motion

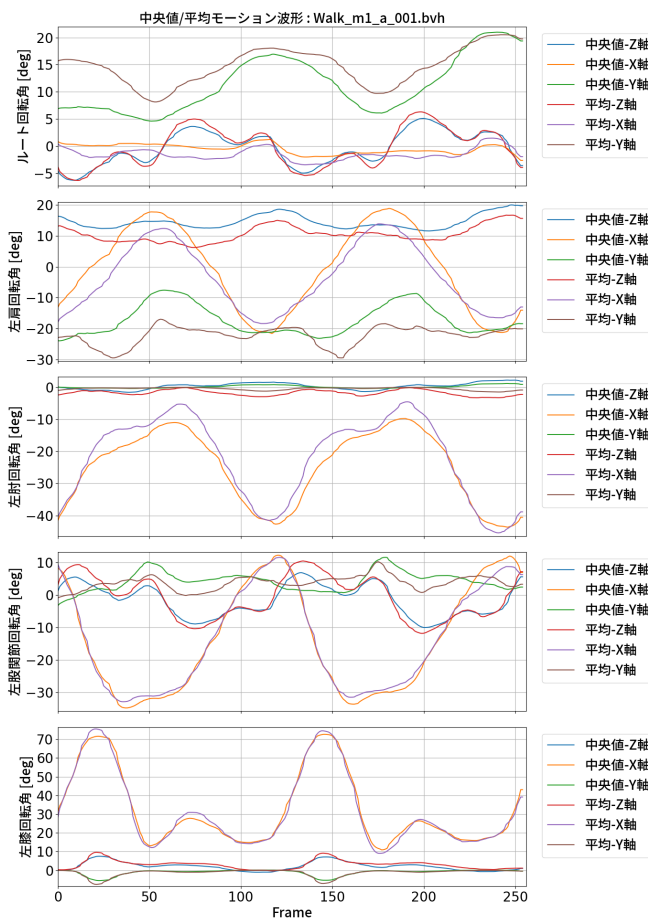


図 A-3 歩行動作での中央値モーショ及び平均モーショの波形
Fig. A-3 Median & Average motion signals for "Walk" motion

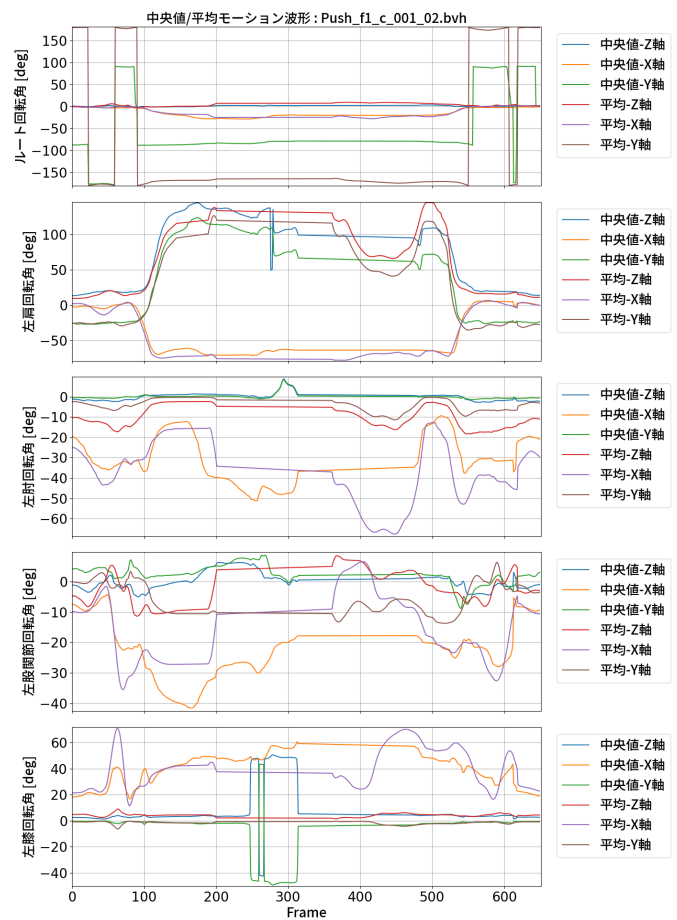


図 A-4 壁押し動作での中央値モーショ及び平均モーショの波形
Fig. A-4 Median & Average motion signals for "Push" motion

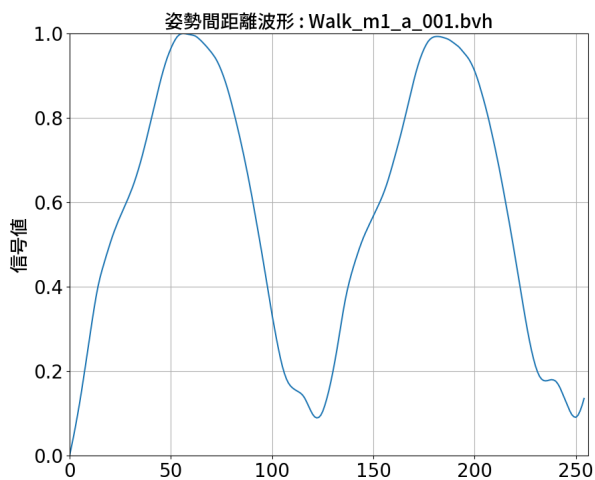


図 A-5 歩行動作での姿勢間距離の波形
Fig. A-5 Pose-Distance signal for "Walk" motion

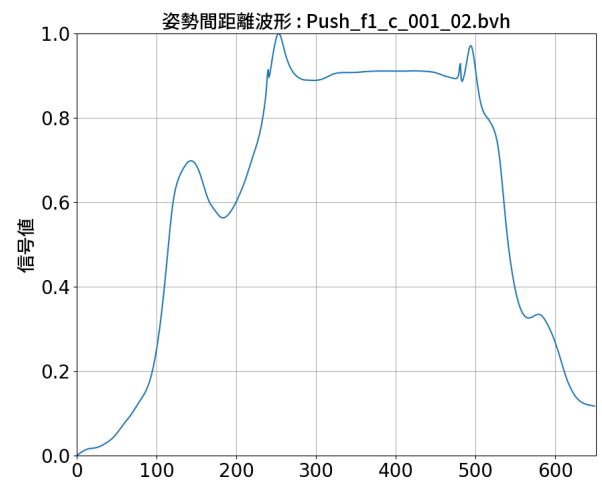


図 A-6 壁押し動作での姿勢間距離の波形
Fig. A-6 Pose-Distance signal for "Push" motion