

LSTM を用いた手指軌跡の予測 Prediction of Finger-tip Trajectories using LSTM

瀬川 修[‡]
Osamu Segawa

1. はじめに

本研究では、身体動作をリアルタイムにモニタリングし、機器等の誤操作につながる不適切な行動を未然に検出する要素技術について検討を行った。一般に機器操作においては、手指の動きが重要と考えられることから、本稿では「手指の軌跡」をトラッキングと予測の対象とする。

動体の軌跡予測については、従来様々な手法[2]-[7]が提案されているが、時系列データの予測手法として近年盛んに研究が行われている「再帰型ニューラルネットワーク」は、時空間における点群系列の予測方式として有望と考えられる。本稿では LSTM (Long Short-Term Memory) [1]による手指軌跡の予測手法を提案し、模擬的な操作ターゲットに対する軌跡予測の有効性を評価する。

2. 手指キーポイントの検出

手指のキーポイント検出には、赤外線測距によるデプスカメラ (Microsoft Kinect v2) を用いた。検出の手順は以下 (i)~(iv)のとおりである。

ここでは、デプスカメラを作業者と操作対象の上方に設置し、真下の空間で手指が動く座標系を設定した。作業者を上から見て床面の左右方向を X 軸、床面の前後方向を Y 軸、カメラから床面方向を Z 軸とする。

- (i) 手指が画角内にない状態での 3 次元点群データを取得し「背景データ」とする。また、画角内に手指などが入った状態で 3 次元点群データを取得し「対象データ」とする。
- (ii) 上記対象データから、Z 軸方向の設定範囲に基づき対象物(身体)の存在する領域外の点群を削除する。
- (iii) 背景データと対象データの点群の差分を取ることで、手指を含む点群だけを抽出する。
- (iv) 上記点群中の Y 値 (作業者から向かって対象物の方向) が最大のものを探索し、手指のキーポイントとする。

提案手法では、キーポイントの 3 次元座標を、作業者上方から見た平面 (X-Y) と、作業者側方から見た平面 (Y-Z) に分割し、それぞれの平面で軌跡の予測を行うものとする。

スタート位置から操作盤を模擬したターゲット(後述)に向かう動作の指先軌跡を実測した例を図 1、図 2 示す。ここで、X 軸と Y 軸の目盛単位は pixel、Z 軸目単位は mm である (1pixel=2mm に相当)。

3. 時系列回帰による軌跡予測

本稿では、従来のコンベンショナルな手法として点群の過去履歴の線形回帰による予測手法についても比較検討を行った。具体的には、2 節の手法により検出された直近の

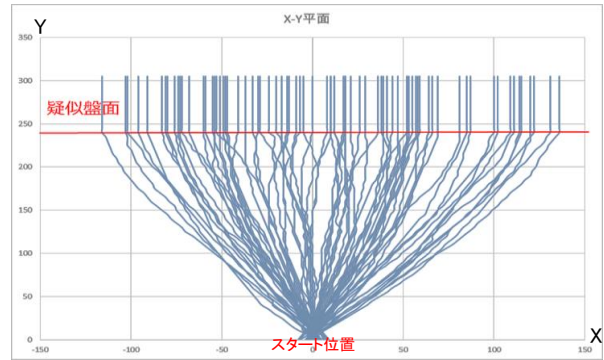


図 1 実測した手指軌跡の例 (X-Y 平面)

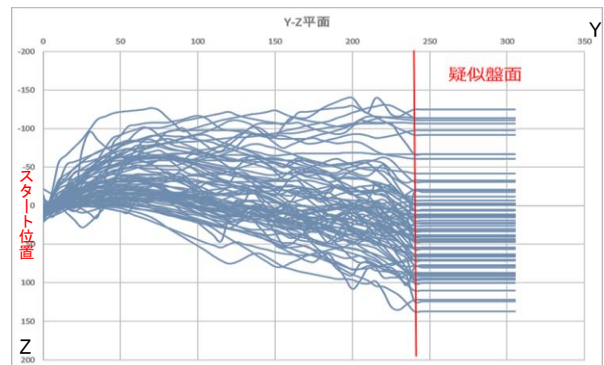


図 2 実測した手指軌跡の例 (Y-Z 平面)

点群系列に基づき最小 2 乗法によって「方向ベクトル」を推定し、ある時点での軌跡予測を逐次的に行う。

4. LSTM による軌跡予測

2 節の手法により手指のキーポイントを抽出し、その軌跡 (スタート位置からターゲットに触れるまで) を一定間隔でサンプリングした点群系列から、LSTM による予測モデルを学習する。ここで、座標系 (X,Y,Z) は前節の手法と同様の設定とする。

今回用いた LSTM のトポロジーは、入力層 1、隠れ層 5、出力層 1 である。学習では実測した 79 個の軌跡をもとに、±0.5 の範囲の乱数で各点を変化させて 10,349 個にデータ拡張を行っている。

5. 評価と考察

5.1 評価方法

今回対象問題として、作業者正面に操作盤を模擬したターゲット (120mm 四方の隣接した矩形 6 個を 3×2 で並べた疑似盤面) を設置し、スタート位置からターゲットに触

[‡] 中部電力(株) エネルギー応用研究所,

Chubu Electric Power Co., Inc.

Energy Applications Research & Development Center

れるまでの指先をトラッキングしながら、軌跡を逐次的に予測するタスクを設定した。作業者とターゲット間の距離は480mmである。

5.2 評価結果

評価では、LSTMの学習用とは異なるテストデータの軌跡12個(3節の時系列回帰の評価と共通)を使用した。時系列回帰による予測誤差を表1に、LSTMによる予測誤差を表2に示す。両手法ともY軸位置が25点目(スタートから約250mm)での予測誤差である。ターゲットの矩形サイズが120mm四方であることから、予測誤差が60mm以内に収まっていれば正しい予測ができていくことになる。ここで「予測開始タイミング」とは最初に正しい予測ができた位置(スタート地点からの距離)である。

表1 評価結果(時系列回帰)

評価指標	手法: 時系列回帰 mm	
	(X-Y)	(Y-Z)
予測誤差(平均)	55.4	41.9
予測誤差(標準偏差)	35.3	28.7
予測誤差(最大値)	145.4	93.1
予測開始タイミング(平均)	233	246

表2 評価結果(LSTM)

評価指標	手法: LSTM mm	
	(X-Y)	(Y-Z)
予測誤差(平均)	11.8	30.0
予測誤差(標準偏差)	11.6	17.1
予測誤差(最大値)	31.7	49.8
予測開始タイミング(平均)	221	233

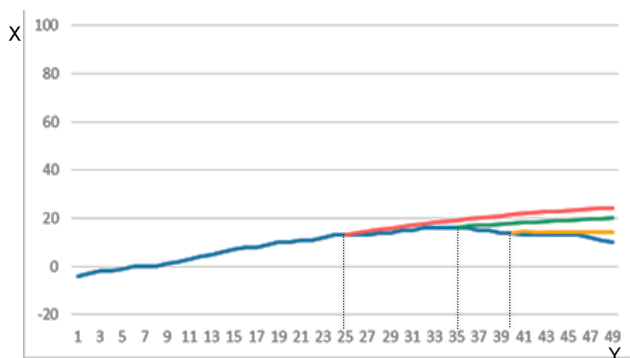


図3 LSTMによる予測結果の例(X-Y平面)

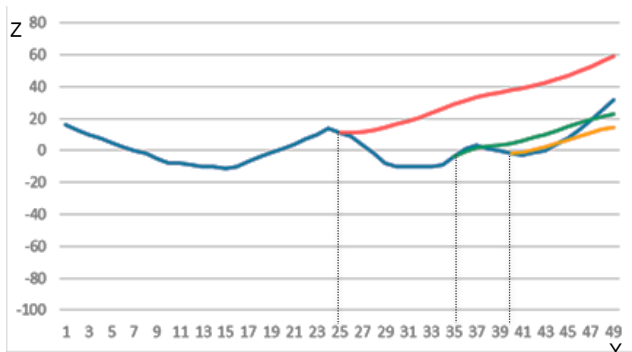


図4 LSTMによる予測結果の例(Y-Z平面)

LSTMによる予測結果の例を図3、図4に示す。ここで、青線が実測値、赤線は25点、緑線は35点、黄線は40点の位置における予測軌跡である(Y軸はスタートからターゲットまでを49分割したスタートからのサンプリング点である)。

5.3 考察

前記5.2の評価結果より、LSTMによる予測手法の有効性が確認された。

- 時系列回帰では、軌跡のサンプリング間隔にもよるが、突発的な変動に対し方向推定を誤ることがある。また、ターゲットに接近しても、予測の分散が大きく方向推定が安定しない。
- LSTMでは、ある時点で予測が正しい場合は、以降安定した予測が行えている。一方、水平方向(X-Y平面)に対し、鉛直方向(Y-Z平面)は予測誤差が大きい傾向にある。これは図1、図2の例からもわかるとおり軌跡変動の性質の違いに起因する(身体構造や認知的要因)と考えられるため、鉛直方向の予測モデルに改良の余地がある。
- 予測開始タイミングについては両手法ともほぼ同等の性能であった。

6. 関連研究

これまでLSTMを用いた「動体の軌跡予測手法」としては、バスケットボールの軌道[2]や、一般道やハイウェイにおける車の軌跡[3][4]などが検討されている。また、[5]では雑踏における各人の動線を予測し、集団の動態の解析を行っている。

最近では、軌跡予測をEncoder-Decoderによりモデル化する試み[6]も行われている。さらに[7]ではEncoder-DecoderにAttentionを併用したモデルにより、人流の予測と異常行動の検出を行う手法を提案している。

7. おわりに

本稿では、再帰型ニューラルネットワーク(LSTM)による手指軌跡の予測手法を提案した。評価実験の結果、LSTMは従来の点群の時系列回帰と比較して予測誤差が小さく、安定した予測が可能であることがわかった。今後の課題としては、LSTMによる予測モデルの改良や、動作の個人差や対象機器の配置に適応するための転移学習の検討、などが挙げられる。

参考文献

- [1] S.Hochreiter et al., "Long short-term memory", Neural Computation, 9 (8), pp.1735-1780 (1997).
- [2] R.C.Shah et al., "Applying Deep Learning to Basketball Trajectories", KDD2016, Large Scale Sports Analytic Workshop (2016).
- [3] A.Khosroshahi et al., "Surround Vehicles Trajectory Analysis with Recurrent Neural Networks", ITSC2016, pp.2267-2272 (2016).
- [4] F.Altche et al., "An LSTM Network for Highway Trajectory Prediction", arXiv:1801.07962v1 (2018).
- [5] A.Alahi et al., "Social LSTM: Human Trajectory Prediction in Crowded Spaces", CVPR2016, pp.961-971 (2016).
- [6] S.H.Park et al., "Sequence-to-Sequence Prediction of Vehicle Trajectory via LSTM Encoder-Decoder Architecture", arXiv:1802.06338v2 (2018).
- [7] T.Fernando et al., "Soft + Hardwired Attention: An LSTM Framework for Human Trajectory Prediction and Abnormal Event Detection", arXiv:1702.05552v1 (2018).