Recent Development in Human Motion and Gait Prediction

Junwu Zhang
junwuz@stanford.edu

Monroe Kennedy
monroek@stanford.edu

Stanford University

Assistive Robotics and Manipulation Laboratory

Department of Mechanical Engineering
Stanford University

July 02, 2020

Motivation

- Human intent is often hard to model and predict, especially in scenarios where the underlying human objective might not be clear

Table 1: Taxonomy of human intent prediction

<table>
<tr><td colspan="2" rowspan="2">Human</td><td colspan="2">Execution Strategy
(Action)</td></tr>
<tr><td>Observer
Knows</td><td>Observer
Unknown</td></tr>
<tr><td rowspan="2">Objective
Function</td><td>Observer
Knows</td><td>All is Known (e.g. Ping Pong)
where both objective and actions are clear</td><td>Human Action Model is unclear
or suboptimal (e.g. chess)</td></tr>
<tr><td>Observer
Unknown</td><td>Human action model is well known,
but objective is not (e.g. joy-riding in car
or free running, where destination
or direction is unclear)</td><td>Poor action model and objective
function (e.g. Poor / good cook,
no idea of final dish)</td></tr>
</table>

- Bipedal locomotion can adapt to complex environments

Introduction (cont.)

- Combination of gait dynamics and human intent prediction has significant potential in various robotics applications (fall prevention, active prosthesis, etc.)

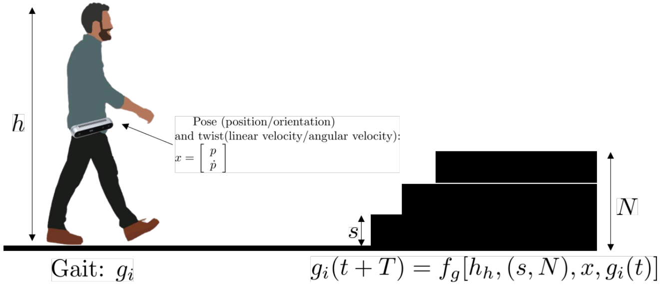


Figure 2: Possible Problem Scenario: Staircases

Literature Review

To predict bipedal gait dynamics and the related human motion, two main steps are required:

- Data Collection: collect various types of gait related data
- Data Analysis: modeling, learning and prediction of gait or motion

Literature Review - Data Collection

Microsoft Kinect

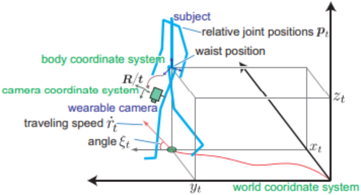
- Objective: Gait prediction from visual data
- Approaches: Use single [1] or multiple [2] Kinects, extract skeletal points directly from Kinect software development kit (SDK) or the underlying raw data streams [3]
- Pros: Easy to set up and use (with the help of Kinect SDK), multiple devices could improve accuracy
- Cons: Cameras are usually fixed in one spot, so it is hard to capture the entirety of an environment, or collect data outdoors

Literature Review - Data Collection (cont.)

Egocentric Visual Data [4]

- Objective: Human gait estimation from visual data
- Approaches: Used a camera mounted on a person's thigh, close to the waist area
- Pros: Thigh-mounted camera is mobile, can be used outdoors and can capture environmental information around a person
- Cons: Could use more testing to find the best point and/or orientation to mount the camera, also could adjust the type of camera mounted

Example using body-mounted camera



(a) Parameter definitions

(b) Actual experiment setup

Figure 3: Thigh-mounted camera for gait prediction [4]

Literature Review - Data Collection (cont.)

Electromyography (EMG)-based methods [5]

- Objective: Gait initiation estimation and detection for a person wearing lower-limb prosthesis
- Approaches: EMG sensors for gait initiation estimation, and inertial measurement units (IMUs) for gait initial movement detection
- Pros: Demonstrated that EMG could be used to predict initial movements in gait motion
- Cons: EMG prediction only seems useful when prosthetic leg leads in the gait motion, EMG sensors are harder to work with and requires a much higher sampling frequency

Literature Review - Data Collection (cont.)

Other wearable sensors

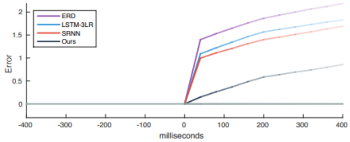
- Objective: Gait data collection for biomechanics analysis
- Approaches: Strategically placed IMUs on thigh, calf and foot [6]; or Vicon markers for 3D gait measurement [7]
- Pros: Some IMU-based methods can be agnostic to sensor placements and initial human postures; marker systems are accurate
- Cons: Data transmission for wearable sensors are not easy enough yet for more mobile application scenarios

Literature Review - Data Analysis

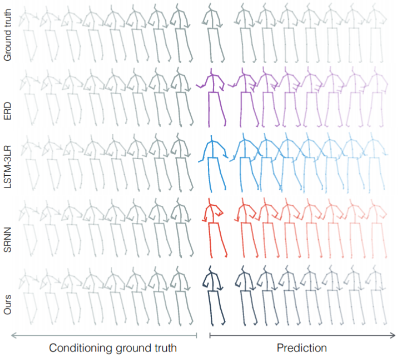
Recurrent Neural Network (RNN) [8]

- Objective: Human motion prediction (short term)
- Approaches: sequence-to-sequence (seq2seq) with sampling-based loss function
- Pros: Reduces the discontinuity at the start of the prediction, lightweight, action-agnostic
- Cons: Best prediction results are dependent on high-level supervision in the form of action labels

Literature Review - Data Analysis (cont.)



(a) RNN yields lower error, especially close to start of prediction



(b) Simulation results shows smoother simulation close to $t = 0$

Figure 4: Motion prediction using Long Short-Term Memory (LSTM) and RNN [8]

Literature Review - Data Analysis (cont.)

Bio-LSTM [9]

- Objective: 3D full-body meshes prediction given history of 3D poses
- Approaches: Biomechanics-based loss function with LSTM
- Pros: Novel loss function is a first step in biomechanical constraints on gait prediction, robust to noise
- Cons: Independence between pedestrians is assumed, genders are not differentiated though it could be [10]

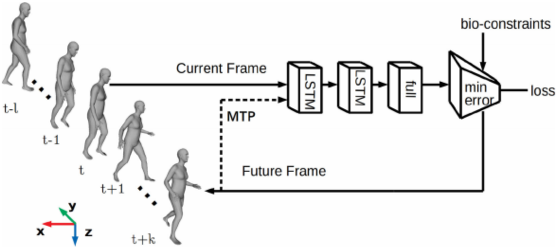


Figure 5: Neural networks (NN) model architecture of Bio-LSTM [9]

Literature Review - Data Analysis (cont.)

Conditional variational autoencoder (CVAE)

- Objective: Multi-modal multi-agent trajectory prediction (no incorporation of human pose)
- Approaches: Trajectron, a framework that combines CVAE, LSTM and dynamic spatiotemporal graphs [11]
- Pros: Able to generate distribution of trajectory prediction for multiple agents simultaneously, in a multimodal environment
- Cons: How robots might incorporate this for lower-level planning is not yet explored

Literature Review - Data Analysis (cont.)

Additional methods:

- Feature Learning [12]
  - Pros: First to model human pose by different body components in representation learning network, lightweight
  - Cons: Feature learning on human pose prediction is not new
- Context Learning [13]
  - Pros: Predicts both human and object motion well in human-environment interaction dataset
  - Cons: Prediction error increases with noise
- Convoluted neural networks (CNN) [14]
  - Pros: Very lightweight model, use CNN to leverage the continuous and temporal nature of trajectory
  - Cons: Model is basic with no social context or physical modeling; number of layers might be reduced

Overall findings:

- Research on the data collection step is relatively more mature. Recent work on this step focuses on getting refined results, instead of having new fundamental approaches.
- Data analysis has seen lots of recent research. Multiple types of machine learning methods for analysis and prediction have been proposed. Traditional filtering or optimization-based methods appear less in recent work.

Discussions - Summaries (cont.)

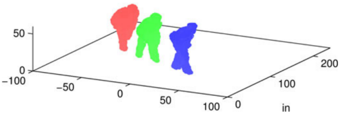
Common approaches:

- For data collection in home environments, Kinect-based system has proved to be an adequate substitute for complex motion capture system like Vicon [15], even in healthcare-related scenarios [3].
- For data analysis, RNN-based methods are very popular, with various models based on LSTM, gated recurrent unit (GRU) and attention are proposed for specific applications.
- We can also see that more research start to look into the multimodal and uncertain nature of prediction, with more variational autoencoder (VAE) and CVAE methods being proposed and predictions generated with a distribution.

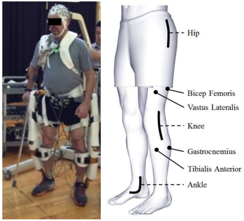
Discussions - Opportunities

High level:

- What are the use case objectives of gait prediction, and what advances are still required to achieve these objectives?



(a) Prediction in fall prevention [3]



(b) Lower-limb exoskeleton with EMG sensors [16] can also benefit from gait predictions

Figure 6: Potential use cases of gait prediction

Discussions - Opportunities (cont.)

Opportunities in data collection:

- Wearable cameras and sensors could be used as tools for gait-related data collection.

(a) Wearable sensors system

(b) Baseline measurement system

Figure 7: Wearable sensor suite in [17]

- For those wearable devices, a metric for sensor limitations and invasiveness has not been proposed

Discussions - Opportunities (cont.)

Opportunities in data analysis:

- How might we predict for longer time horizons?
- How can "gait" be efficiently and more completely characterized?
  - What might be the minimum number of states that are required to describe the gait, and what might those parameters be?
  - Furthermore, what could we say about the minimum degree-of-freedom (DOF) that we need to describe gait motion?

Conclusions

- Various approaches on the data collection and analysis for human walking gait and motion is presented.
- Check out the paper to view more detailed analysis and comparison of the methods.
- Future opportunities are discussed with potential questions raised.
- We hope this meta-analysis could be a useful review of this field, and inspire more future research.

Acknowledgement

The authors thank Luciana Frazao and Shivani Guptasarma for their helpful discussions and contributions.