

Research Article

Motion Prediction of Human Wearing Powered Exoskeleton

Xin Jin ^{1,2}, Jia Guo ², Zhong Li ¹, and Ruihao Wang ²

¹Department of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

²Department of Mechanical Engineering, Virginia Tech, Blacksburg, VA 24061, USA

Correspondence should be addressed to Xin Jin; xinjin@njjust.edu.cn

Received 13 September 2020; Revised 1 December 2020; Accepted 11 December 2020; Published 22 December 2020

Academic Editor: Georgios I. Giannopoulos

Copyright © 2020 Xin Jin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the development of powered exoskeleton in recent years, one important limitation is the capability of collaborating with human. Human-machine interaction requires the exoskeleton to accurately predict the human motion of the upcoming movement. Many recent works implement neural network algorithms such as recurrent neural networks (RNN) in motion prediction. However, they are still insufficient in efficiency and accuracy. In this paper, a Gaussian process latent variable model (GPLVM) is employed to transform the high-dimensional data into low-dimensional data. Combining with the nonlinear autoregressive (NAR) neural network, the GPLVM-NAR method is proposed to predict human motions. Experiments with volunteers wearing powered exoskeleton performing different types of motion are conducted. Results validate that the proposed method can forecast the future human motion with relative error of 2%~5% and average calculation time of 120 s~155 s, depending on the type of different motions.

1. Introduction

Powered exoskeleton is a type of mechanical skeleton that can be worn by human user. This specific kind of machinery can provide external force and torque for the user to accomplish the motion, such as running, jumping, and weight bearing. Therefore, powered exoskeletons are widely used in military, medical rehabilitation, and disaster relief [1–3]. Motion prediction plays a crucial role in the interaction between powered exoskeleton and human. Accurate motion prediction makes exoskeleton react to users' need properly and provide efficient movement. A practical motion prediction algorithm is supposed to predict user's intention of motion based on the sensor information, such as joint angular of lower-limb [4], EMG signals of muscles [5], neural activation signals [6], human motion sequences [7, 8], and multisensor information [9].

Because of the stochasticity and nonlinearity of the captured motions, researchers have involved neural network to improve the prediction results. For example, Ghosh et al. used a Dropout Autoencoder LSTM (long short-term

memory) method to predict natural looking motion sequences over long-time horizons without catastrophic drift or motion degradation [10]. Cheng et al. proposed an adaptable neural network for human motion prediction, which is able to accommodate human's time-varying behaviors and to provide uncertainty bounds of the predictions in real time [11]. Tang et al. employed a motion context modeling to predict long-term motion using the recurrent neural network (RNN) [12]. Other works, such as multilayered LSTM [13] and deep RNNs [14], also achieved great success. The common strategy of these remarkable works is using neural network algorithms as the encoder learning from motion sequence along time domain, with distributed hidden states or quantify features to store information about motion in the past. However, neural network algorithms also have intrinsic problems. One is the dilemma of choosing training data. High-dimensional observed data usually leads to poor model accuracy, low calculation efficiency, and impacted generalization ability. On the other hand, using hidden states or motion features [15] to train the neural network may raise the risk of information loss and inaccurate prediction.

GPLVM is a nonparametric probabilistic algorithm for dimensionality reduction, which has seen remarkable successes in generating natural and smooth human motion [16–19]. Given the dataset $\{x_i\}_{i=1}^n \subset X$, it generates the corresponding low-dimensional representations $\{y_i\} \subset Y$ in the latent space. The GPLVM can be viewed as a nonlinear probabilistic extension to the classical PCA, and it has been extensively used in analyzing the complex human motion. However, the GPLVM only identifies the configuration manifold where the human motion evolves on, it does not identify the vector field defined on the manifold which is useful for future motion prediction. As a natural consideration, we apply the NAR neural network to identify the evolution in the low-dimensional latent space, which is much more sufficient because the computation is implemented to the low-dimensional representation. In this paper, a new prediction method is proposed as the combination of GPLVM and NAR network algorithms. The novel method integrates the advantages of both algorithms and as a result achieves higher prediction accuracy with relatively lower calculation time.

2. Method

2.1. Problem Analysis. In practice, exoskeletons are expected to be compliant with users' body gestures. The conventional multibody dynamical models are mostly developed based on the assumption of ideal joints. However, the hypothesis of ideal joints may neglect many features such as the geometric properties of skeleton and the elasticity of tissues. Modeling all the uncertainties with physical approaches is intractable. Thus, in this paper, we take a detour and apply the data-driven approach.

One natural representation of human poses can be given with the positions of characteristic points, such as joints and tips, of human body. Such representation belongs to a high-dimensional Euclidean space \mathbb{R}^{3N} , where N denotes the number of characteristic points. Due to the actual constraints of human body, the motion happens in a subset of the entire space. In particular, we assume that the representations of all possible poses form an embedded manifold in \mathbb{R}^{3N} , which is called configuration manifold in geometric mechanics. Although the configuration manifold of human body is complicated due to the uncertainties mentioned above, most of human motion can actually be categorized into a number of regimes, each of which evolves on a submanifold of much lower dimensions. With the measurements of poses from motion capture experiments, we may apply manifold learning algorithms to reveal an empirical estimate of the submanifold.

Two ingredients are necessary in predicting the behavior of dynamical systems. One is the configuration manifold on which the motion evolves and another is the vector field defined over the manifold, which determines the evolution of motion. In this paper, the manifold

learning algorithm is applied to generate estimate of the submanifold of regime of motion. In particular, the GPLVM is implemented since it generates smooth empirical estimates, which is necessary for defining smooth evolution law. The vector field is then identified using the NAR neural network.

2.2. Definitions and Fundamental Assumptions of GPLVM Method. The Gaussian process latent variable model (GPLVM) proposed in is a nonparametric probabilistic algorithm for dimensionality reduction, which has seen lots of remarkable successes in generating natural and smooth human motion [20–22]. As a probabilistic algorithm, GPLVM is inherently robust to the measurement error. This feature is especially helpful for handling digitized motion capture data, since the usage of adhesive landmarks may introduce error due to the displacement and deformation of the skin. Compared to other common manifold learning algorithms, such as isometric mapping (ISOMAP) and local linear embedding (LLE), the estimate generated with GPLVM is a smooth map from the latent space to configuration space along with the probability distribution, which are both helpful for estimating the evolution law.

Here is a brief review of GPLVM. Gaussian process (GP) is the essential underlying definition of the algorithm. Roughly speaking, for $x \in R^n$, a stochastic process $S: x \mapsto S(x) \in R^m$ is called GP if any finite collection of images $\{S(x_i)\}_{i=1}^N$ form the joint Gaussian distribution. Any function $f: R^n \rightarrow R^m$ can be viewed as a realization of the GP. Analogous to the finite dimensional case, GP is characterized by a mean function $\mu(x)$ and a kernel (or covariance) function $k(x_i, x_j)$. The kernel function must have positive definite Gram matrix such that the covariance matrix is valid. Such functions $k(x_i, x_j)$ are also called to be of positive type. For any set of images $\{S(x_i)\}_{i=1}^N$, the probability distribution is characterized by the mean and covariance as follows:

$$E[S(x_i)] = \mu(x_i), \quad (1)$$

$$\text{cov}[S(x_i), S(x_j)] = k(x_i, x_j), \quad \forall i, j = 1, \dots, N. \quad (2)$$

One commonly used kernel function is the Gaussian kernel $k(x_i, x_j) = \exp(-x_i - x_j^2/\sigma^2)$. The constant coefficients of kernel functions are called hyperparameters. In statistical inference, the mean function $\mu(x)$ of a GP is usually set to constant zero, and the type of kernel function is selected based on the smoothness. Any given samples of the mapping $\{x_i, y_i\}_{i=1}^N \in R^{n \times m}$ form an event. Given the value of hyperparameters β , the probability can be calculated from the joint Gaussian distribution:

$$P\{(x, y)\} = \mathcal{N}\{y: x, [k(x_i, x_j)]_{N \times N}\}, \quad (3)$$

where $\mathbf{y} = [y_1, \dots, y_N]$ and $\mathbf{x} = [x_1, \dots, x_N]$. If only samples $\{(x_i, x_j)\}$ are given, the estimation of hyperparameters

can be obtained with maximum likelihood estimation (MLE):

$$\beta^* = \arg \min\{\mathbf{P}(x, y); \beta\}. \quad (4)$$

In the problem of dimensionality reduction, the only available information is the samples $\{(y_i)\}$ in the original space. The corresponding latent coordinates $\{(x_i)\}$ and hyperparameters are to be identified. The study [23] gives an algorithm with the assumption that the hyperparameter β follows a certain probability distribution. The conditional probability $P(y|x)$ can be calculated by marginalizing out the hyperparameters from $P(y|x, \beta)$. Then, the coordinates $\{(x_i)\}$ and β are jointly estimated through optimizing the marginalized likelihood. Readers are referred to reference [23, 24] for the details of the algorithm.

2.3. NAR Neural Network. NAR neural network is a dynamic neural network with feedback and memory function [25]. Its output depends not only on the current input, but also on the previous input and output. It uses the dynamic correction method to reduce the calculation time of model updating, and the matrix order remains unchanged when the sample increased, which improves the calculation efficiency [26]. This method has the advantages of less operation time and high prediction accuracy and has the characteristics of strong learning ability and approximation of any nonlinear function. It is more suitable for time series prediction than static neural network.

The algorithm model of the NAR neural network is expressed as follows:

$$y(t) = \mathbf{f}[y(t-1), y(t-2), \dots, y(t-d)], \quad (5)$$

where $y(t)$ is the output value at time t , $y(t-1)$, $y(t-2)$, \dots , $y(t-d)$ are the output values before time t , d is the delay order, and $\mathbf{f}[\]$ is the nonlinear function obtained by learning and training. It is clear that the predicted value of $y(t)$ at this moment is determined by the values of $y(t)$ in the past.

The NAR dynamic neural network is composed of input layer, output layer, hidden layer, and delay variable. It has two network modes, one is parallel (close-loop) network mode, in which the output of the neural network will be feedback to the input layer and continue to learn with other inputs. The other is series-parallel (open-loop) network mode, in which the expected output of the neural network will be feedback to the input layer. In this paper, we choose the series-parallel network model, which can improve the prediction accuracy.

To summarize, our strategy of GPLVM-NAR is shown in Figure 1. By using the GPLVM method, the high-dimensional motion data are mapped into low-dimensional latent space. Then, prediction in the latent space is made using the NAR neural network. At last, future motion can

be forecasted by inverse mapping of the predicted representation.

The proposed procedure consists of five steps: (1) compute sequence of human motion wearing exoskeleton with motion capture devices; (2) calculate the velocities of every mark points between frames; (3) obtain the latent coordinates of velocity data by training the GPLVM; (4) predict future changes of the representations of training data in latent space using NAR neural network; (5) get future motion data in observation space by inverse mapping.

3. Experiments and Results

Experiment is conducted in order to verify the capability of the proposed human motion prediction algorithm. A motor-driven lower-limb prototype exoskeleton with 6 degrees of freedom is used in the experiment. The volunteer wearing this powered exoskeleton is asked to perform different types of motion indoors. The assigned motion includes (1) walking on ground, (2) walking on slope, and (3) walking on stairs. The lower-limb exoskeleton is activated to help the user complete the assigned motion.

Video image sequences of the volunteer's motion wearing powered exoskeleton are captured utilizing Coda Motion cx1 analysis system (Charnwood Dynamics Ltd.). The dual-camera system is used to reconstruct the motions in Cartesian coordinates. Each motion is required to last 4 seconds. And the motions are recorded by 200 frames (camera frame rate 50 fps).

For the motion capture system, we place 22 marker points on volunteer's lower body in advance. They are arranged as follows: 4 markers on each side of thigh, 4 markers on each side of crus, and each marker on every hip, knee, and ankle. The marker points and their locations are shown in Figure 2.

The positions of landmarks are digitized from stereo vision. And the sequences of human skeletons are calculated automatically by the Coda Motion analysis system based on linear regression between markers, as shown in Figures 3(a)–3(f). Notice that every group of 4 points on the thighs and cruses are replaced by a virtual point for visualization. We take one motion sequence of walking on ground as an example. Firstly, markers' velocities are calculated from frame to frame. Secondly, the velocities are input to the GPLVM model to reduce the dimension from 66-dimensional matrix (22 points*3 coordinates) to 2-dimensional latent vector. In the 2-dimensional latent space, each point is corresponded to the motion matrix in each frame in observation space, as shown in Figure 3(g).

Obvious cyclical trends can be found in Figure 3(g). As a result, prediction based on the representations of training data in the latent space is made. We assume that early 75% of the captured data (150 frames of poses) as train data and last

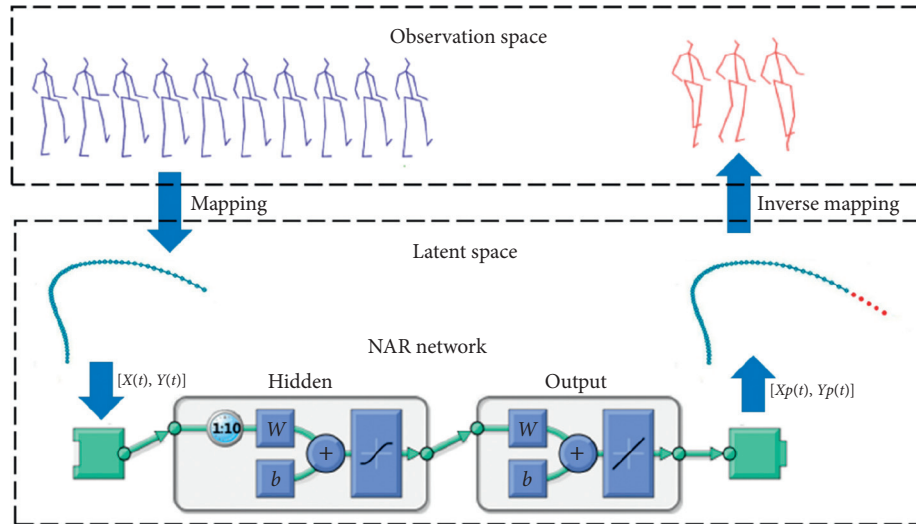


FIGURE 1: Strategy of the GPLVM-NAR method. Mapping motion data into low-dimensional latent space, predicting coordinates in the latent space using the NAR neural network, and obtaining future motions nondestructively by inverse mapping.



FIGURE 2: Diagrammatic sketch of the marker points and their locations: left/right thigh (marker no. 1–4/no. 5–8), left/right crus (marker no. 9–12/no. 13–16), left/right hip (marker no. 17/no. 20), left/right knee (marker no. 18/no. 21), and left/right ankle (marker no. 19/no. 22). Note that markers on the thigh are in front of them. So, they are invisible from this visual angle.

25% of the captured data (50 frames of poses) as true reference data. The NAR neural network is used to forecast future value of the representations in this paper. The NAR neural network is constructed with 10 input layer nodes, 10 hidden layer neurons, and 1 output layer node. It uses the

last 10 dataset as feedback delays to estimate the next future data and then iterate over to the end. Additionally, future motion data in observation space are obtained by inverse mapping of the data in latent space. Prediction results are shown in Figure 4. It can be observed that the NAR neural

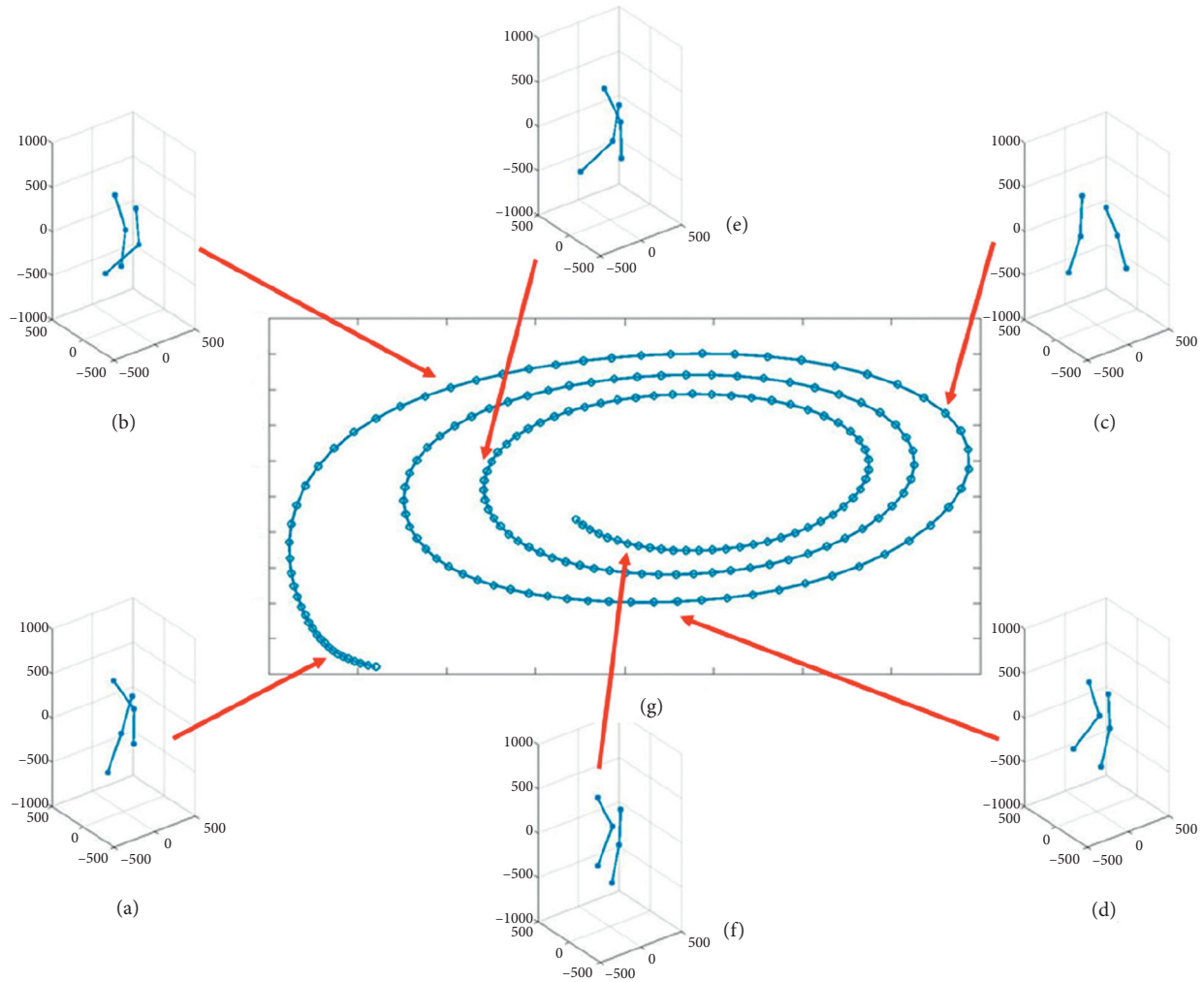


FIGURE 3: GPLVM training result of a sequence of walking on ground. (a)–(f) Frames of motion poses in observation space. Human leg skeletons are calculated based on captured marker data. (g) Representations of the training data in latent space. Cyclicity of the representations reflects the periodicity of human movement.

network method performs very high predicting accuracy in latent space, as shown in Figure 4(e). Consequently, predicted motion in observation space shows the same high similarity with human truth, as shown in Figure 4(a)–4(d).

Different motions of human user wearing powered exoskeleton walking on ground, on slope, and on stairs are experimented. 10 best performances of each motion are picked up for prediction practice. Comparisons are made with the existing method including Res-GRU and LSTM-3LR [12]. The prediction results using different methods are shown in Figure 5. We introduce mean relative error (MRE) and their standard deviation (SD) to evaluate the ability of the proposed method for human motion predicting. The averages of relative errors of all the marker points between

predicted motion and true value are first calculated. Then, their mean values and standard deviations are listed as a reference value. The comparison of prediction results are shown in Table 1. Furthermore, average calculation time of each method is calculated and shown in Figure 6. Computation time of the GPLVM-NAR method is less than other two methods, because of the data dimensionality reduction. Furthermore, the calculation time of the proposed method does not increase significantly with the increase of motion complexity.

From the comparison of prediction accuracy and average calculation time, conclusion can be made that our method is more competitive than other competitors. One reason is that GPLVM has its advantage in the training of one single set of motion data, especially when the amount

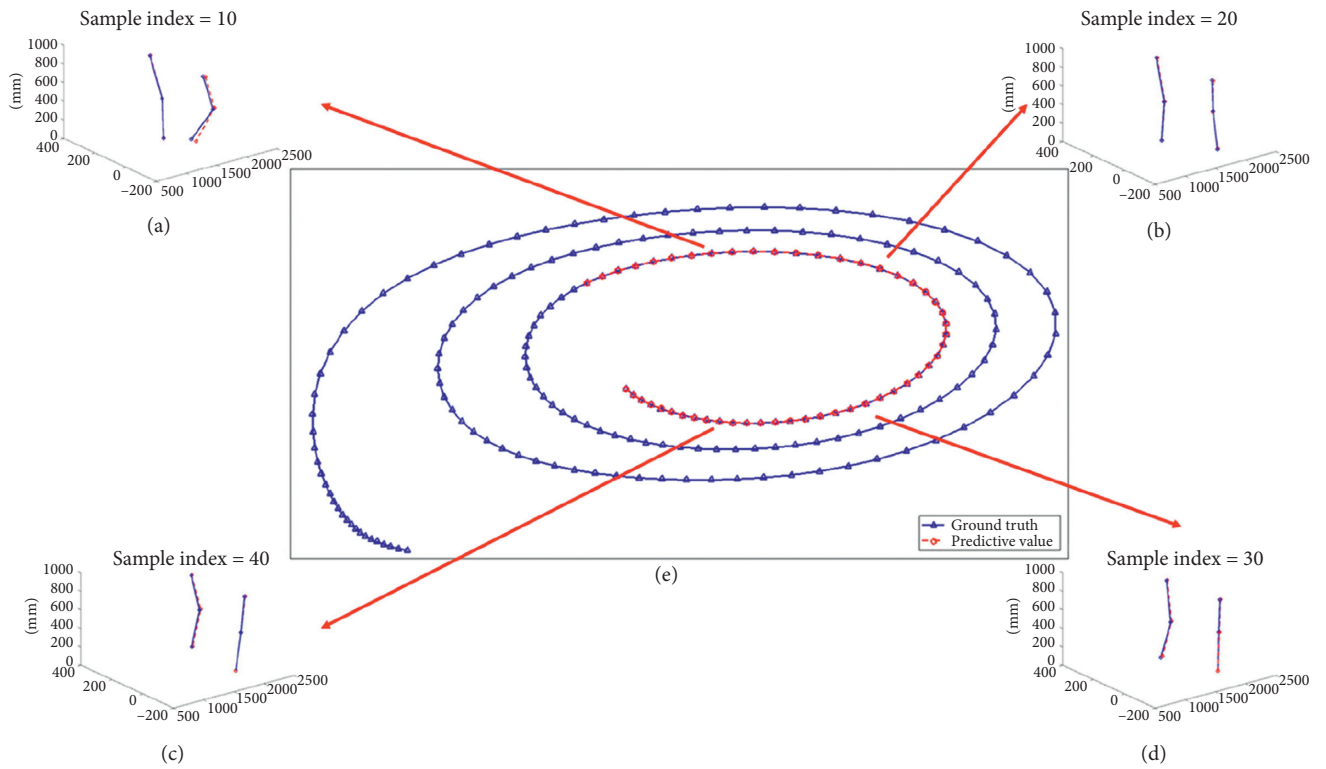


FIGURE 4: Comparison of the proposed method results and ground-truth poses. (a)–(d) Examples of particular sample indexes of predicted motions (dotted line) and ground-truths (solid line). (e) Exhibition of predicted data and test data in latent space. The predicted data (circle) are very close to test data (triangle) in latent space.

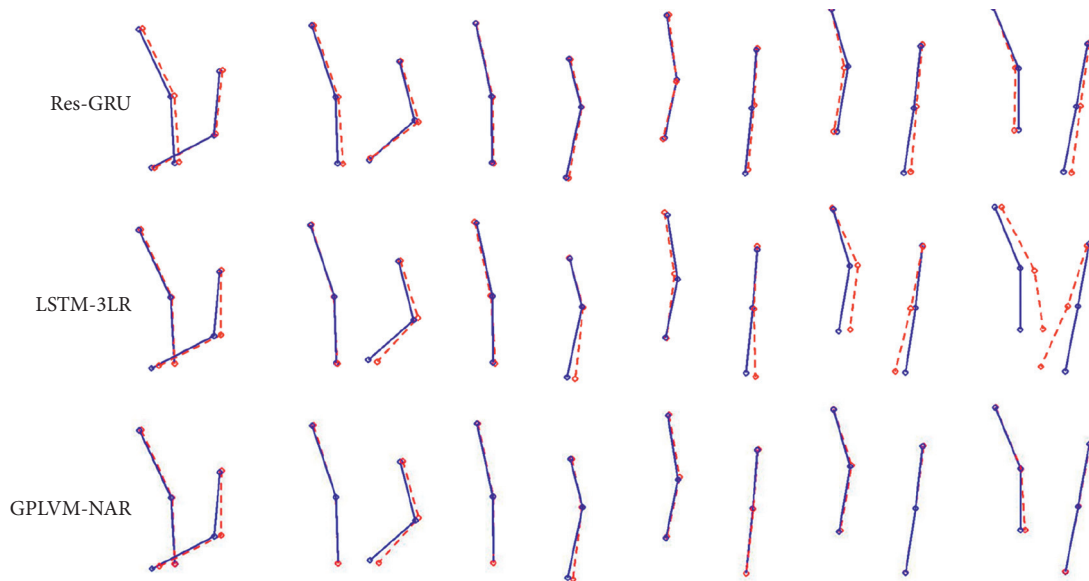


FIGURE 5: Prediction results using different methods, Res-GRU, LSTM-3LR and GPLVM-NAR. The predicted motion results are in dotted lines and ground-truths are in solid lines.

TABLE 1: Prediction comparison between different methods with various human motions via mean relative error and standard deviation. Each motion type has 10 datasets. The best performance is highlighted in boldface.

Methods	Walking on ground		Walking on slope		Walking on stairs	
	MRE	SD	MRE	SD	MRE	SD
Res-GRU	0.0875	0.0217	0.0995	0.0212	0.1322	0.0267
LSTM-3LR	0.1398	0.0128	0.1658	0.0143	0.2078	0.0466
GPLVM-NAR	0.0197	0.0057	0.0323	0.0043	0.054	0.0058

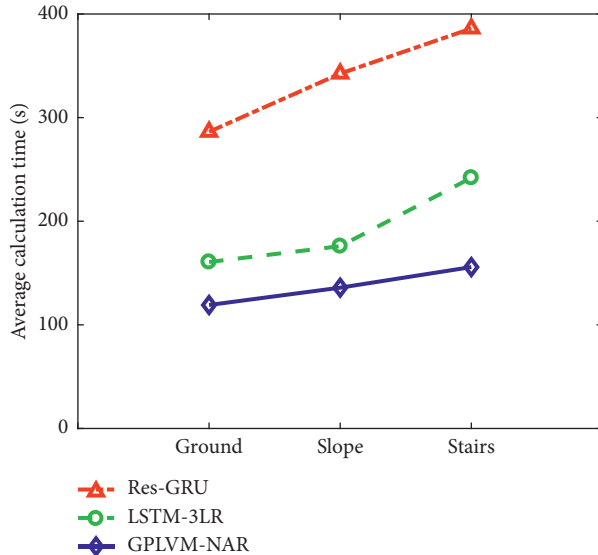


FIGURE 6: Average calculation time of different methods: Res-GRU (triangle dots with dot dash line), LSTM-3LR (circle dots with dotted line), and GPLVM-NAR (diamond dots with solid line).

of training data is also very small, while other methods may lose their convergences. Although the method is less accurate when predicting human motion of walking on stairs due to the motion complexity, it still outperforms others.

4. Conclusions

In this paper, a GPLVM-NAR method is proposed to predict future motions of human wearing powered exoskeleton. With the help of the GPLVM method, the dimension of observed human motion data is reduced. Then, prediction is made by employing the NAR network algorithm. Experiment results demonstrate that the proposed algorithm outperforms existing methods with advantages in relative error of 2%~5% and average calculation time of 120 s~155 s.

The GPLVM method generates a smooth map from the latent space to configuration space. It shows great potential in the prediction with a small amount of data. Because the predicting process is performed on the low-dimensional latent coordinates, the proposed method achieves higher model accuracy, computational efficiency, and generalization ability compared to the conventional neural network approaches. Therefore, it is more applicable for the real assistive strategies of powered exoskeleton.

Data Availability

The data used to support the findings of this study are all included within the article.

Ethical Approval

The study was approved by the Ethical Committee of Nanjing University of Science and Technology, and all methods were carried out in accordance with relevant guidelines and regulations.

Consent

Informed consent was taken from subjects before enrollment.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank Professor Xu, Professor Guan, and their team for providing the exoskeleton prototype. The powered exoskeleton is provided by the cooperation work of Xiaorong Guan (Nanjing University of Science and Technology) and Wei Dong (Harbin Institute of Technology). This study was funded by the National Defense Basic Scientific Research Program of China, grant no. B1020132012.

References

- [1] Q. Fang, T. G. Li, and H. J. CaiZhao, "A simplified inverse dynamics modelling method for a novel rehabilitation exoskeleton with parallel joints and its application to trajectory tracking," *Mathematical Problems in Engineering*, vol. 2019, no. 5, 10 pages, Article ID 4602035, 2019.
- [2] G. Zhu, X. Kong, G. Zhao et al., "Modeling and simulation of an unpowered lower extremity exoskeleton based on gait energy," *Mathematical Problems in Engineering*, vol. 2020, Article ID 4670936, 15 pages, 2020.
- [3] R. A. R. C. Gopura, D. S. V. Bandara, K. Kiguchi, and G. K. I. Mann, "Developments in hardware systems of active upper-limb exoskeleton robots: a review," *Robotics and Autonomous Systems*, vol. 75, pp. 203–220, 2016.
- [4] D.-X. Liu, X. Wu, W. Du, C. Wang, and T. Xu, "Gait phase recognition for lower-limb exoskeleton with only joint angular sensors," *Sensors*, vol. 16, no. 10, p. 1579, 2016.
- [5] Y. Fan, Z. Guo, and Y. Yin, "SEMG-based neuro-fuzzy controller for a parallel ankle exoskeleton with

- proprioception,” *International Journal of Robotics and Automation*, vol. 26, no. 4, p. 450, 2011.
- [6] K. Lee, D. Liu, L. Perroud, R. Chavarriaga, and J. d. R. Millán, “A brain-controlled exoskeleton with cascaded event-related desynchronization classifiers,” *Robotics and Autonomous Systems*, vol. 90, pp. 15–23, 2017.
- [7] Y. Du, Y. Wong, Y. Liu et al., “Marker-less 3D human motion capture with monocular image sequence and height-maps,” in *Proceedings of the European Conference on Computer Vision*, pp. 20–36, Springer, Cham, Switzerland, October 2016.
- [8] M. Nixon and A. Aguado, *Feature Extraction and Image Processing for Computer Vision*, Academic Press, Cambridge, MA, USA, 2019.
- [9] W. Ma, X. Zhang, and G. Yin, “Design on intelligent perception system for lower limb rehabilitation exoskeleton robot,” in *Proceedings of the 2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 587–592, IEEE, Xi’an, China, August 2016.
- [10] P. Ghosh, J. Song, E. Aksan, and O. Hilliges, “Learning human motion models for long-term predictions,” in *Proceedings of the 2017 International Conference on 3D Vision (3DV)*, pp. 458–466, IEEE Computer Society, Qingdao, China, October 2017.
- [11] Y. Cheng, W. Zhao, C. Liu et al., “Human motion prediction using semi-adaptable neural networks,” in *Proceedings of the 2019 American Control Conference (ACC)*, pp. 4884–4890, IEEE, Philadelphia, PA, USA, July 2019.
- [12] Y. Tang, L. Ma, W. Liu, and W. Zheng, “Long-term human motion prediction by modeling motion context and enhancing motion dynamic,” in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*, Stockholm, Sweden, July 2018.
- [13] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” *Advances in Neural Information Processing Systems*, <https://arxiv.org/abs/1409.3215>, 2014.
- [14] J. Martinez, M. J. Black, and J. Romero, “On human motion prediction using recurrent neural networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2891–2900, IEEE, Honolulu, HI, USA, July 2017.
- [15] J. Butepage, M. J. Black, D. Kragic, and H. Kjellstrom, “Deep representation learning for human motion prediction and classification,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6158–6166, Honolulu, HI, USA, July 2017.
- [16] G. Song, S. Wang, Q. Huang et al., “Harmonized multimodal learning with gaussian process latent variable models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 411 page, 2019.
- [17] W. E. Leithead, E. Solak, and D. J. Leith, “Direct identification of nonlinear structure using gaussian process prior models,” in *Proceedings of the European Control Conference*, IEEE, Linz, Austria, July 2015.
- [18] P. Li and S. Chen, “A review on Gaussian process latent variable models,” *CAAI Transactions on Intelligence Technology*, vol. 1, no. 4, pp. 366–376, 2016.
- [19] A. Mallasto, S. Hauberg, and A. Feragen, “Probabilistic Riemannian submanifold learning with wrapped Gaussian process latent variable models,” in *Proceedings of the The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 2368–2377, Naha, Okinawa, Japan, April 2019.
- [20] N. D. Lawrence, “Gaussian process latent variable models for visualisation of high dimensional data,” *Advances in Neural Information Processing Systems*, vol. 15, pp. 329–336, 2004.
- [21] J. Wang, A. Hertzmann, and D. J. Fleet, “Gaussian process dynamical models,” in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1441–1448, British Columbia, Canada, December 2006.
- [22] C. E. Rasmussen, *Gaussian Processes in Machine Learning*, In *Summer School on Machine Learning*, Springer, Berlin, Heidelberg, 2003.
- [23] N. Lawrence, “Probabilistic non-linear principal component analysis with Gaussian process latent variable models,” *Journal of Machine Learning Research*, vol. 6, pp. 1783–1816, 2005.
- [24] N. D. Lawrence and J. Quiñero-Candela, “Local distance preservation in the GP-LVM through back constraints,” in *Proceedings of the 23rd International Conference on Machine Learning*, pp. 513–520, ACM, Pittsburgh, PA, USA, June 2006.
- [25] K. W. Wang, C. Deng, J. P. Li, Y. Y. Zhang, X. Y. Li, and M. C. Wu, “Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network,” *Epidemiology and Infection*, vol. 145, no. 6, pp. 1118–1129, 2017.
- [26] R. Sarkar, S. Julai, S. Hossain et al., “A Comparative study of activation functions of NAR and NARX neural network for long-term wind speed forecasting in Malaysia,” *Mathematical Problems in Engineering*, 10.1155/2019/6403081, vol. 2019, Article ID 6403081, 14 pages, 2019.