Journal of Southeast University (English Edition) Vol. 39, No. 2, pp. 187 – 193 June 2023 ISSN 1003—7985

Motion estimation for cable-driven distal end-effectors using attention-based bi-directional gated recurrent neural networks

Xu Xinzhou[1] Chen Yongfa[1] Liu Guangming[2] Li Ziqian[2] Zhao Li[3] Wang Zhengyu[2]

([1]School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)
([2]School of Mechanical Engineering, Hefei University of Technology, Hefei 230009, China)
([3]School of Information Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: A data-driven motion estimation approach based on attention-based bi-directional gated recurrent neural networks was proposed to adaptively estimate the motion of a cable-driven distal end-effector. First, the data construction was performed to obtain short-term temporal sequences as training samples. The data were then processed using the bi-directional gated recurrent neural networks with self-attention modules for building sequential models on the samples. Finally, based on the motion dataset of the cable-driven distal end-effectors, the estimation-performance comparison experiments were performed using the motor's position, speed, and input time sequence for the system-control signal as the sample features. The results show that compared with conventional sequence modeling regression approaches, the proposed approach can achieve better performance for estimating the motion of the end-effector. Therefore, it can effectively estimate the motion of cable-driven distal end-effectors under complex conditions.

Key words: cable-driven distal end-effectors; motion estimation; bi-directional gated recurrent neural networks; attention mechanism

DOI: 10.3969/j.issn.1003 – 7985.2023.02.010

Compared to rigid links, flexible cables are conventionally used in robots for minimally-invasive surgery[1–3], rehabilitation training[4–5], and trajectory planning[6–7], thus benefiting from flexibility, low inertia, and distant mechanical driving[1–2,5]. Nevertheless, cable-driven robots may be influenced by nonlinear time-delay factors such as cables' flexibility, fatigue, mechanical transmission, and external interferences[2,8], which typically appear in complex surgical backgrounds[1,9]. Therefore, estimating the motion of a distal end-effector driven by cables allow machines to perceive the movement of a mechanical manipulator. Through the estimation procedure, it is further useful in motion compensation for distal end-effectors[10–11].

Estimating the motion of a distal end-effector is crucial for the precise control of robotic arms or instruments[10,12–13]. Due to the high stability in rigid cases, early studies on rigid-link robots mainly use prior knowledge in modeling an end-effector's motion[14–15]. Additionally, to perform precise control and manipulation, cable-driven end-effectors require force estimation from internal grasping[16] and external sensing[1]. Furthermore, works on motion estimation resort to the inclusion of various modalities, such as computer vision[13] and movement-sensing measurements[12]. In addition to these above-mentioned works, time-delay estimation is also considered in cable-driven end-effectors' control[10].

However, despite theoretical accuracy, most existing studies on motion estimation for cable-driven robots focus on using prior knowledge mainly from cables' deformation and mechanical friction[10,12–13], which may result in inaccurate estimation due to the adaptation to various conditions. Furthermore, it is difficult to model a cable-driven end-effector's motion due to the nonlinear influence within the mechanical system[2,8].

Therefore, a motion estimation approach for cable-driven distal end-effectors is proposed, which uses attention-based bi-directional gated recurrent neural networks (Bi-GRNNs) to jointly learn bi-directional and long-distance dependency in sequences. Gated recurrent neural networks improve sequence modeling performance by using specifically designed memory units for learning sufficient sequential knowledge[17]. Moreover, within GRNN, the units can be chosen as long short-term memory (LSTM) or gated recurrent units (GRUs), typically used in speech analysis and natural language processing (NLP)[18–19], and can contain attention mechanisms to further improve GRNN learning performance[20–22].

The proposed approach first segments three-dimensional motion series into samples with labels corresponding to the end-effector's real rotation angles, with the input expected rotation degree and the motor's rotation degree and speed. Subsequently, the samples are carefully fed to a specifically designed deep Bi-GRNN to build the mapping

Received 2022-10-10, **Revised** 2023-05-14.

Biographies: Xu Xinzhou (1987—), male, doctor, associate professor; Wang Zhengyu (corresponding author), male, doctor, associate professor, wangzhengyu_hfut@hfut.edu.cn.

Foundation items: The China Postdoctoral Science Foundation (No. 2022M711693), the Natural Science Foundation of China (No. 52175221, 61801241, 62071242), the Natural Science Foundation of Jiangsu Province (No. BK20191381).

Citation: Xu Xinzhou, Chen Yongfa, Liu Guangming, et al. Motion estimation for cable-driven distal end-effectors using attention-based bi-directional gated recurrent neural networks[J]. Journal of Southeast University (English Edition), 2023, 39(2): 187 – 193. DOI: 10.3969/j.issn.1003 – 7985.2023.02.010.

from these series to the end-effector's real motion. Following that, the trained Bi-GRNN model is used for the motion estimation task.

1 Proposed Strategy

1.1 Cable-driven series elastic actuator

Considering the end-effector for cable-driven surgical robots, Wang et al. [23] built a cable-driven series elastic actuator (CDSEA) as a simplified cable-driven instrument prototype, as shown in Fig. 1. The actuator includes a motor & cable-driven module, a back-spring & cable-driven module, a cable-pulley system, and a long-distance driven revolute joint. Specifically, the tension sensor at the end of the back-spring is used to adjust the initial cable tension. Note that the joint uses a micromagnetic encoder to measure the rotation angle in a 2D space and the gratings to collect the initial displacement information of the drive and reset cables, respectively. Furthermore, Fig. 1 illustrates the overall experiment system, including CDSEA, host computer, and dSPACE real time control system. dSPACE controls CDSEA to execute the running track and collect the data information of each sensor in real time. Therefore, relevant experiments are performed based on CDSEA to obtain real-world data for end-effector control.

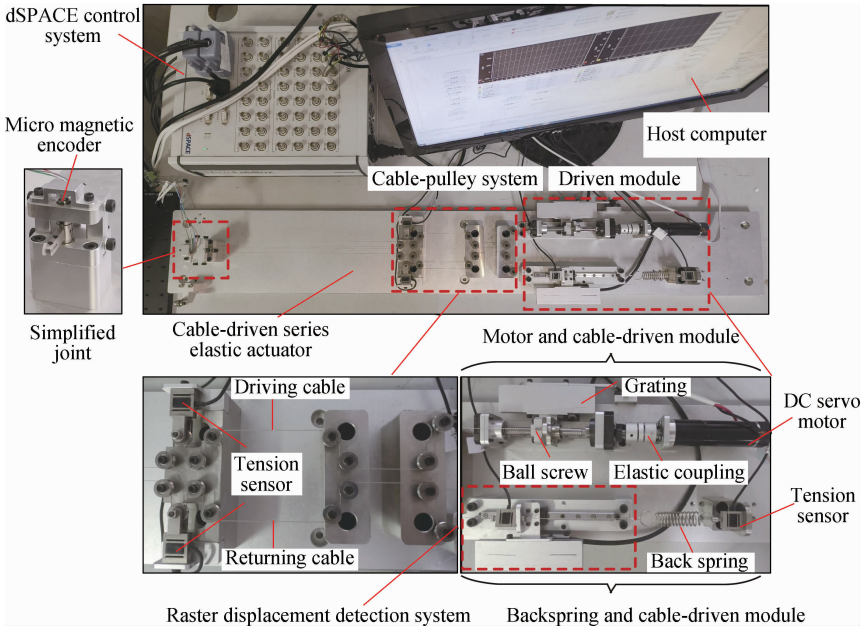


Fig. 1 Overview of the cable-driven series elastic actuator

Additionally, the actuator uses a DC servo motor (Maxon 118752 RE 25 with the incremental encoder Maxon 225780 and planetary reducer Maxon 406764 GP 26A) driven by a Copley ACJ-055-09 driver. The driver communicates through the RS232 interface, with the control system as dSPACE/MicroLabBox connected by MATLAB/Simulink for signal input[2]. The connection of the servo motor and ball screw (SFKR, pitch 2 mm and diameter 6 mm) takes into account GCPSW20-6-6 double diaphragm elastic coupling. The cable type is set as Carl Stahl TECHNOCABLES CG719045 (nominal diameter 0.45 mm), the sliding module as MISUMI SSE-BV8G-130, and the back-spring as MISUMI DS8723.

The simplified dynamic model based on CDSEA is depicted in Fig. 2, in which the cable-pulley driving module and the cable-pulley returning module are simplified into mass-spring-damping systems[2]. The system driving mechanism is as follows. Input the theoretical angle to the system. The ball screw-motor module drives the cable to realize the single-direction rotation of the joint, which is the system's actual joint angle output, and the motor encoder feeds back the motor position and speed informa-

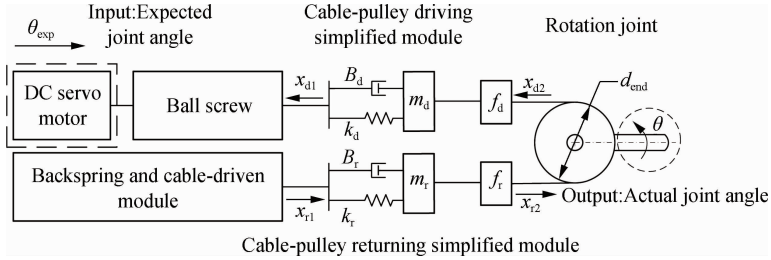


Fig. 2 Simplified dynamic model based on CDSEA

tion in real time. The motor drives the cable in reverse, and the back-spring provides the reverse driving force for the system to realize the reset function.

The control system receives a periodic sinusoidal signal of the expected rotation angles (with fixed maximum angles), and the sequential data of the input signal, motor's position, and speed are recorded accordingly, with a sampling rate of 100 Hz. The resultant three time series are inputted into the estimation system as feature sequences. The system also records the real rotation angles corresponding to each time point of the input as the output of the estimation system; these angles are collected by the rotation-measure sensor.

1.2 Attention-based bi-directional gated recurrent neural networks

When obtaining the feature sequences, the proposed approach first segments these sequences into the samples, each with $T = 50$ time points (0.5 s), using the last time point within these T points in the corresponding output sequence as the sample's annotation. Note that it is optional to add a time-delay operator to perform temporal compensation[10]. This results in the motor-state features and corresponding output motion depicted in Fig. 3.

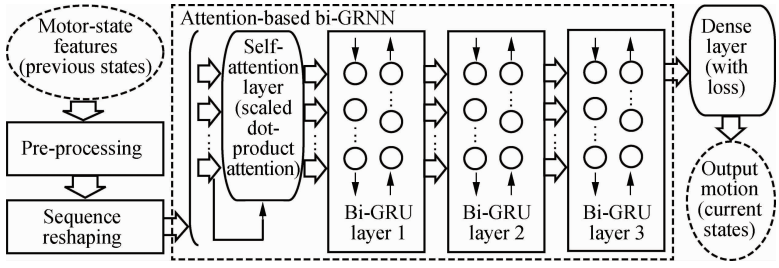


Fig. 3 Overview of the proposed attention-based Bi-GRNN for motion estimation

The training samples' features are denoted as $\boldsymbol{\chi} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\} \subset \mathbf{R}^{T\times 3}$, with the annotations $\boldsymbol{\Upsilon} = \{y_1, y_2, \ldots, y_N\} \subset \mathbf{R}$, where N refers to the number of the training samples. Moreover, the approach includes pre-processing using min-max normalization on each sample's features for x. Afterwards, the sequence-reshaping procedure aims to set the input time series to the network by reshaping a sequence to $\frac{T}{d}\times 3d$, where $d \geqslant 1$ represents the reshaping scale. Hence, the attention-based Bi-GRNN's input training set can be obtained as $\tilde{\boldsymbol{\chi}} = \{\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, \ldots, \tilde{\boldsymbol{x}}_N\} \subset \mathbf{R}^{\frac{T}{d}\times 3d}$.

Next, the approach employs a self-attention layer using scaled dot-product attention[24] as the first layer of the attention-based Bi-GRNN, resulting in its output for the i-th sample $\tilde{x}_i$ in a batch including $N^{(0)}$ samples as follows:

$$(\boldsymbol{O}^{(0)})_i = \text{softmax}\left(\frac{\boldsymbol{Q}_i\boldsymbol{K}_i^{\mathrm{T}}}{\sqrt{m}}\right)\boldsymbol{V}_i \in \mathbf{R}^{\frac{T}{d}\times m} \tag{1}$$

setting the self-attention layer's query, key, and value for $\tilde{x}_i$ to

$$\left.\begin{aligned}\boldsymbol{Q}_i &= \tilde{\boldsymbol{x}}_i\boldsymbol{W}^{(Q)} \in \mathbf{R}^{\frac{T}{d}\times m}\\ \boldsymbol{K}_i &= \tilde{\boldsymbol{x}}_i\boldsymbol{W}^{(K)} \in \mathbf{R}^{\frac{T}{d}\times m}\\ \boldsymbol{V}_i &= \tilde{\boldsymbol{x}}_i\boldsymbol{W}^{(V)} \in \mathbf{R}^{\frac{T}{d}\times m}\end{aligned}\right\} \tag{2}$$

with the corresponding learnable linear-mapping matrices of $\boldsymbol{W}^{(Q)}$, $\boldsymbol{W}^{(K)}$, $\boldsymbol{W}^{(V)} \in \mathbf{R}^{3d\times m}$, considering an m-dimensional space for $(\boldsymbol{O}^{(0)})_i$.

Afterwards, the component at time t of the batch's initial input for the first Bi-GRU layer $\boldsymbol{O}^{(0)}$ is noted as $\boldsymbol{O}_t^{(0)} = [(\boldsymbol{O}_t^{(0)})_1 \quad (\boldsymbol{O}_t^{(0)})_2 \quad \ldots \quad (\boldsymbol{O}_t^{(0)})_{N^{(0)}}] \in \mathbf{R}^{m\times N^{(0)}}$[25], and the output of the three Bi-GRU layers can be written as follows:

$$\boldsymbol{O}_t^{(3)} = f_{\text{Bi-GRU}}^{(3)}(f_{\text{Bi-GRU}}^{(2)}(f_{\text{Bi-GRU}}^{(1)}(\boldsymbol{O}_t^{(0)}))) \subset \mathbf{R}^{n^{(3)}\times N^{(0)}} \tag{3}$$

where $f_{\text{Bi-GRU}}^{(1)}(\cdot)$, $f_{\text{Bi-GRU}}^{(2)}(\cdot)$, and $f_{\text{Bi-GRU}}^{(3)}(\cdot)$ indicate the output mapping of the three Bi-GRU layers with $n^{(1)}$, $n^{(2)}$, and $n^{(3)}$ dimensions, respectively. Note that for the k-th of the three Bi-GRU layers in Fig. 2 ($k = 1, 2, 3$), the time-t output for $f_{\text{Bi-GRU}}^{(k)}(\cdot)$ can be represented as follows:

$$\boldsymbol{O}_t^{(k)} = [\boldsymbol{H}_t^{(+)\mathrm{T}} \quad \boldsymbol{H}_t^{(-)\mathrm{T}}]^{\mathrm{T}} \in \mathbf{R}^{n^{(k)}\times N^{(0)}} \tag{4}$$

where $\boldsymbol{H}_t^{(+)}$ and $\boldsymbol{H}_t^{(-)}$ are the backward and forward directional hidden outputs of a Bi-GRU layer, respectively. For either of the two directions, its hidden outputs ($\boldsymbol{H}_t^{(+)}$ or $\boldsymbol{H}_t^{(-)}$) can be represented as follows:

$$\boldsymbol{H}_t = \boldsymbol{U}_t \odot \boldsymbol{H}_{t\pm 1} + (\boldsymbol{E} - \boldsymbol{U}_t) \odot \widetilde{\boldsymbol{H}}_t \in \mathbf{R}^{\frac{n^{(k)}}{2}\times N^{(0)}} \tag{5}$$

considering $t \pm 1$ for the two directions, where $\odot$ represents element-wise product operator and $\boldsymbol{E} \in \boldsymbol{R}^{\frac{n^{(k)}}{2}\times N^{(0)}}$ is an all-one matrix. Note that the two directions' outputs are calculated separately, without shared learnable variables. Within the equation, the candidate hidden state is defined with the input $\tilde{\boldsymbol{x}}_t^{(0)}$ as follows:

$$\widetilde{\boldsymbol{H}}_t = \tanh(\boldsymbol{W}^{(h_1)}\tilde{\boldsymbol{x}}_t^{(0)} + \boldsymbol{W}^{(h_2)}(\boldsymbol{R}_t \odot \boldsymbol{H}_{t\pm 1}) + \boldsymbol{b}^{(h)}) \tag{6}$$

in which the linear weights obey $\boldsymbol{W}^{(h_1)} \in \mathbf{R}^{\frac{n^{(k)}}{2}\times n^{(k-1)}}$ ($n^{(0)} = m$ for $k = 1$) and $\boldsymbol{W}^{(h_2)} \in \mathbf{R}^{\frac{n^{(k)}}{2}\times\frac{n^{(k)}}{2}}$, and the bias $\boldsymbol{b}^{(h)} \in \mathbf{R}^{\frac{n^{(k)}}{2}\times N^{(0)}}$. The update gate $\boldsymbol{U}_t$ and reset gate $\boldsymbol{R}_t$ can be obtained through

$$\left.\begin{aligned}\boldsymbol{U}_t &= g(\boldsymbol{W}^{(u_1)}\tilde{\boldsymbol{x}}_t^{(0)} + \boldsymbol{W}^{(u_2)}\boldsymbol{H}_{t\pm 1} + \boldsymbol{b}^{(u)})\\ \boldsymbol{R}_t &= g(\boldsymbol{W}^{(r_1)}\tilde{\boldsymbol{x}}_t^{(0)} + \boldsymbol{W}^{(r_2)}\boldsymbol{H}_{t\pm 1} + \boldsymbol{b}^{(r)})\end{aligned}\right\} \tag{7}$$

with the weights ($\boldsymbol{W}^{(u_1)}$, $\boldsymbol{W}^{(u_2)}$, $\boldsymbol{W}^{(r_1)}$, and $\boldsymbol{W}^{(r_2)}$) and the biases ($\boldsymbol{b}^{(u)}$ and $\boldsymbol{b}^{(r)}$), setting the activation $g(\cdot)$ as a sigmoid function.

When using the three Bi-GRU layers' output $\boldsymbol{O}^{(3)}$, the dense layer's output motion corresponding to the current batch can be presented as follows:

$$\boldsymbol{O} = \boldsymbol{w}\boldsymbol{O}^{(3)}_{\frac{T}{d}} + \boldsymbol{b} \in \mathbf{R}^{1\times N^{(0)}} \tag{8}$$

with the dense layer's weights $\boldsymbol{w} \in \mathbf{R}^{1\times n^{(3)}}$ and biases $\boldsymbol{b} \in \mathbf{R}^{1\times N^{(0)}}$. Then, it is possible to calculate the loss of mean square error (MSE) and perform optimization using the adaptive momentum (Adam) operator.

2 Experiments

2.1 Experimental setups

This study uses the data collected from Hefei University of Technology's Cable-driven End-effector's Motion (HFUT-CDEM) dataset using the designed cable-driven system[23] in the experiments, consisting of 23 input periodic sinusoidal time sequences corresponding to the maximum rotation angles from 2.5° to 57.5° with fixed intervals of 2.5°, respectively. Each sequence contains 200 s using the sampling rate of 100 Hz (20 000 sampling points), with additional 11 measured time sequences from the two gratings' position and speed (four sequences), four tension sensors (four sequences), the motor's position and speed (two sequences), and the effector's position (one sequence), in which the effector's position is set to the target estimated motion. Note that for each rotation angle, this work uses the motor's position and speed and the input periodic sinusoidal time sequence (three sequences in total) to form the features.

These 20 000 sampling points are divided into 12 000 and 8 000 points as training and test sets for each rotation angle. This can be seen as a procedure for imitating the estimation of the end-effector's future motion in real-world cases, considering the nonlinear time-varying factors for the cable-driven system. Furthermore, every 50 time points are defined as a sample, leading to 150-dimensional features (50×3). Considering the head and tail interference, a total of 11 250 training and 7 250 testing samples are chosen for each angle, resulting in 23 training-test trials. Prior to further processing, a min-max normalization on each feature is first employed. For the parametric setups of the proposed approach, $n^{(1)}$, $n^{(2)}$, and $n^{(3)}$ are set to 32, 32, and 16, respectively, while the dimensionality m for the self-attention layer is set to the range of {5, 10, 15}. For the sequence-reshaping procedure sets $d=2$, obtaining a 25×6 shape. The batch size and maximum epoch number are further set to 256 and 30, respectively, with the initial learning rate of {0.000 1, 0.001, 0.01, 0.1} to record the best results.

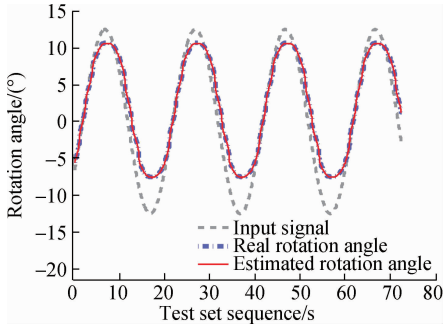
2.2 Experimental results

First, the proposed approach can be compared to conventional applicable regression approaches, which include support vector regression (SVR), LSTM/GRU, attention-based LSTM/GRU, Bi-LSTM/Bi-GRU, attention-based Bi-LSTM, and the proposed attention-based Bi-GRNN (using GRU). Note that the SVR considers the best performance when using a linear kernel with the regularization parameters ranging from {0.000 1, 0.001, 0.01, 0.1}. The LSTM and GRU-based networks used in the comparison all use the three recurrent layers with the same setups as the proposed approach. The MSE and Pearson's correlation coefficient (PCC) is among the indicators compared[26-27]. Therefore, the average MSE and PCC results across the 23 angles' training-test trials are presented in Tab. 1. Note that the table presents both the average and best-average results across the range of the attention dimensionality. The table shows that the proposed approach achieves better performance in most cases by jointly considering the bi-directional structure carefully. Further one-way analysis of variance for the 23 expected maximum rotation angles and the corresponding post hoc results using Scheffé multiple comparisons[28] show a significant difference between the proposed and SVR approaches ($p<0.005$). Considering the close performance between the proposed approach and Bi-LSTM with a focus on the PCC indicator, the results for the 23 expected maximum rotation angles are analyzed separately. The results show that the proposed approach performs better on 15 angles out of 23 ones, specifically for large-angle cases (over 40°).

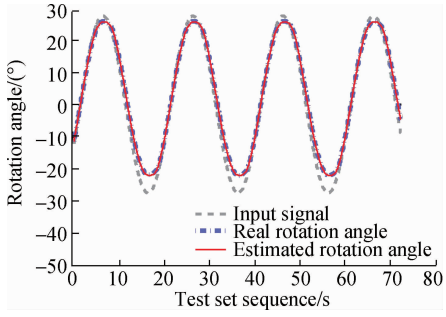
Tab. 1 Comparison between conventional regression and the proposed approach

Approaches/Indicators	MSE	PCC
W/O estimation	9.168 4	0.990 500 7
Linear SVR	0.318 8	0.999 049 1
LSTM	0.216 1	0.999 781 0
GRU	0.071 9	0.999 834 5
Bi-LSTM	0.067 1	0.999 841 7
Bi-GRU	0.040 0	0.999 844 2
Average LSTM + Attention	0.111 5	0.999 810 5
Best-average LSTM + Attention	0.048 7	0.999 839 7
Average GRU + Attention	0.055 0	0.999 808 0
Best-average GRU + Attention	0.043 7	0.999 839 7
Average bi-LSTM + Attention	0.034 1	0.999 841 6
Best-average bi-LSTM + Attention	0.028 6	0.999 855 5
Average proposed approach	0.032 2	0.999 837 1
Best-average proposed approach	0.021 7	0.999 855 8

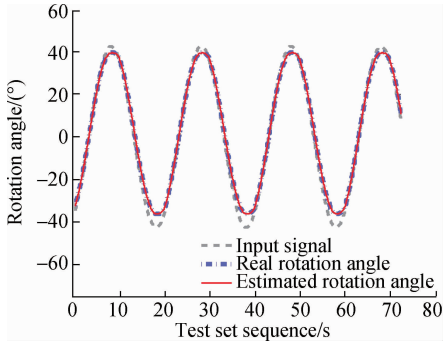
This study presents the rotation angle fitting curves of the proposed approach for different expected rotation angles in Fig. 4 to investigate the detailed performance for each maximum rotation angle. The results show a well-fitting performance for the proposed approach. Further-



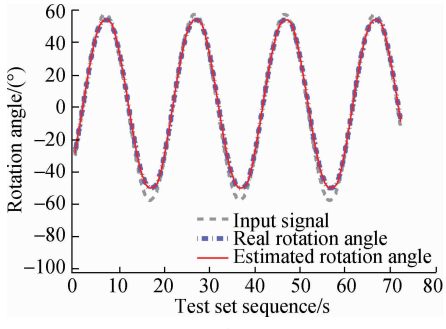
(a)



(b)



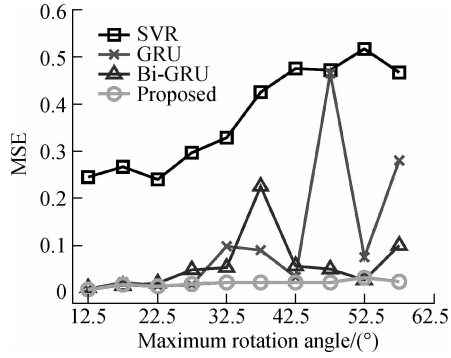
(c)



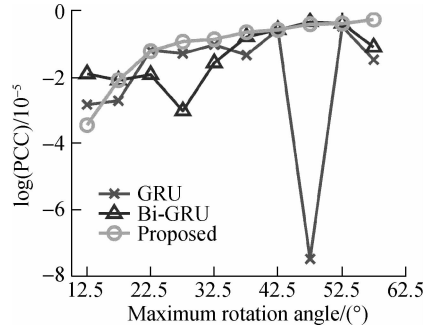
(d)

Fig. 4 Rotation angle fitting curves of the proposed approach for expected maximum rotation angles. (a) 12.5°; (b) 27.5°; (c) 42.5°; (d) 57.5°

more, Fig. 5 also shows the line charts for the proposed approach. The comparison results show that the proposed approach outperforms the other approaches for most of the expected maximum rotation angles. Furthermore, both the MSE and PCC indicators tend to increase with the increase in the expected maximum rotation angle, which is possibly due to the rotation angle scales. Due to this tendency, the proposed approach is more likely to achieve better performance when confronting larger expected maximum rotation angles.



(a)



(b)

Fig. 5 Line charts for the proposed approach considering the maximum rotation angle ranging from 12.5° to 57.5°. (a) MSE compared with SVR, GRU, and Bi-GRU; (b) PCC compared with GRU and Bi-GRU

Finally, it is critical to focus on examining the effect of different setups on the proposed approach. The setups considered include the reshaping scales ($d \in \{2, 5, 10\}$) and the locations of the self-attention layer (at the first, second, and third Bi-GRU layers), with the corresponding MSE and PCC presented in Tab. 2, resulting in a comparison between nine setups in total. These experimental results show that the current setups contribute to the performance of the proposed approach (attention at the first Bi-GRU layer with $d = 2$) when used for motion estimation for a cable-driven distal end-effector. The results also show that an appropriate reshaping scale can improve the performance of the motion estimation task.

Tab. 2 Parametric comparison between the proposed approach and possible setups

d	Attention position	MSE	PCC
2	1st bi-GRU layer	0.021 7	0.999 855 8
	2nd bi-GRU layer	0.022 6	0.999 852 3
	3rd bi-GRU layer	0.027 2	0.999 847 1
5	1st bi-GRU layer	0.026 4	0.999 850 3
	2nd bi-GRU layer	0.025 1	0.999 852 4
	3rd bi-GRU layer	0.024 2	0.999 852 5
10	1st bi-GRU layer	0.0296	0.999 848 2
	2nd bi-GRU layer	0.028 3	0.999 842 5
	3rd bi-GRU layer	0.023 4	0.999 849 0

3 Conclusions

1) This study proposed an approach focusing on estimating the motion of a cable-driven distal end-effector using attention-based Bi-GRNNs to support adaptive motion estimation.

2) The proposed approach used short-term data from input signals and motor motion, which were fed into a network containing three Bi-GRU layers with self-attention. The experimental results indicated that the proposed approach outperformed conventional regression approaches.

3) Despite the works completed in this study, our possible future direction may lie in exploring cross-domain motion estimation considering more complex practical cases.

References

[1] Wang Z Y, Zi B, Wang D M, et al. External force self-sensing based on cable-tension disturbance observer for surgical robot end-effector[J]. *IEEE Sensors Journal*, 2019, **19**(13): 5274 − 5284. DOI: 10.1109/JSEN.2019.2903776.

[2] Yu X, Liu G M, Wang Z Y, et al. Full-closed loop tracking control based on multi-factor coupling compensations using artificial neural network for a cable-pulley-driven surgical robotic manipulator[C]//*Proc USCToMM Symposium on Mechanical Systems and Robotics*. Rapid City, SD, USA, 2022: 43 − 53. DOI: 10.1007/978-3-030-99826-4_5.

[3] Wang W, Li J M, Wang S X, et al. System design and animal experiment study of a novel minimally invasive surgical robot[J]. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 2016, **12**(1): 73 − 84. DOI: 10.1002/rcs.1658.

[4] Chen B, Zi B, Wang Z Y, et al. Knee exoskeletons for gait rehabilitation and human performance augmentation: A state-of-the-art[J]. *Mechanism and Machine Theory*, 2019, **134**: 499 − 511. DOI: 10.1016/j.mechmachtheory.2019.01.016.

[5] Chen Q, Zi B, Sun Z, et al. Design and development of a new cable-driven parallel robot for waist rehabilitation[J]. *IEEE/ASME Transactions on Mechatronics*, 2019, **24**(4): 1497 − 1507. DOI: 10.1109/TMECH.2019.2917294.

[6] Xie F, Shang W W, Zhang B, et al. High-precision trajectory tracking control of cable-driven parallel robots using robust synchronization[J]. *IEEE Transactions on Industrial Informatics*, 2020, **17**(4): 2488 − 2499. DOI: 10.1109/TII.2020.3004167.

[7] Rasheed T, Long P, Roos A S, et al. Optimization based trajectory planning of mobile cable-driven parallel robots[C]//2019 *Proc IEEE/RSJ International Conference on Intelligent Robots and Systems* (*IROS*). Macau, China, 2019: 6788 − 6793. DOI: 10.1109/IROS40897.2019.8968133.

[8] Kuan J Y, Pasch K A, Herr H M. A high-performance cable-drive module for the development of wearable devices[J]. *IEEE/ASME Transactions on Mechatronics*, 2018, **23**(3): 1238 − 1248. DOI: 10.1109/TMECH.2018.2822764.

[9] Min S, Yi S. Development of cable-driven anthropomorphic robot hand[J]. *IEEE Robotics and Automation Letters*, 2021, **6**(2): 1176 − 1183. DOI: 10.1109/LRA.2021.3056375.

[10] Wang Y Y, Yan F, Zhu K W, et al. A new practical robust control of cable-driven manipulators using time-delay estimation[J]. *International Journal of Robust and Nonlinear Control*, 2019, **29**(11): 3405 − 3425. DOI: 10.1002/rnc.4566.

[11] Zhang J, Kan Z Y, Li Y, et al. Novel design of a cable-driven continuum robot with multiple motion patterns[J]. *IEEE Robotics and Automation Letters*, 2022, **7**(3): 6163 − 6170. DOI: 10.1109/LRA.2022.3166547.

[12] Korayem M H, Yousefzadeh M, Kian S. Precise end-effector pose estimation in spatial cable-driven parallel robots with elastic cables using a data fusion method[J]. *Measurement*, 2018, **130**: 177 − 190. DOI: 10.1016/j.measurement.2018.08.009.

[13] Peng H N, Yang X J, Su Y H, et al. Real-time data driven precision estimator for RAVEN-II surgical robot end effector position[C]// 2020 *Proc IEEE International Conference on Robotics and Automation* (*ICRA*). Paris, France, 2020: 350 − 356. DOI: 10.1109/ICRA40945.2020.9196915.

[14] Lertpiriyasuwat V, Berg M C. Adaptive real-time estimation of end-effector position and orientation using precise measurements of end-effector position[J]. *IEEE/ASME Transactions on Mechatronics*, 2006, **11**(3): 304 − 319. DOI: 10.1109/TMECH.2006.876515.

[15] Liu X, Zhao F, Ge S S, et al. End-effector force estimation for flexible-joint robots with global friction approximation using neural networks[J]. IEEE Transactions on Industrial Informatics, 2019, **15**(3): 1730 − 1741. DOI: 10.1109/TII.2018.2876724.

[16] Xue R F, Du Z J, Yan Z Y, et al. An estimation method of grasping force for laparoscope surgical robot based on the model of a cable-pulley system[J]. *Mechanism and Machine Theory*, 2019, **134**: 440 − 454. DOI: 10.1016/j.mechmachtheory.2018.12.032.

[17] Chung J, Gulcehre C, Cho K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[EB/OL]. (2014-12-11) [2022-09-01]. https://arxiv.org/1412.3555.pdf.

[18] Yin W, Kann K, Yu M, et al. Comparative study of CNN and RNN for natural language processing[EB/OL]. (2017-02-07) [2022-09-01]. https://arxiv.org/1702.01923.pdf.

[19] Shewalkar A, Nyavanandi D, Ludwig S A. Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU[J]. *Journal of Artificial Intelligence and Soft Computing Research*, 2019, **9**(4): 235 − 245. DOI: 10.2478/jaiscr-2019-0006.

[20] Basiri M E, Nemati S, Abdar M, et al. ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis[J]. *Future Generation Computer Systems*, 2021, **115**: 279 − 294. DOI: 10.1016/j.future.2020.08.005.

[21] Fernando T, Ghaemmaghami H, Denman S, et al. Heart sound segmentation using bidirectional LSTMs with attention[J]. *IEEE Journal of Biomedical and Health Informatics,* 2020, **24**(6): 1601 - 1609. DOI: 10.1109/JBHI.2019.2949516.

[22] Zhou L Y, Fan X J, Tjahjadi T, et al. Discriminative attention-augmented feature learning for facial expression recognition in the wild[J]. *Neural Computing and Applications*, 2022, **34**(2): 925 - 936. DOI: 10.1007/s00521-021-06045-z.

[23] Wang Z Y, Zi B, Wang D M, et al. Design, modeling and analysis of a novel backdrivable cable-driven series elastic actuator[C]//2019 *Proc IEEE International Conference on Nanotechnology* (*IEEE-NANO*). Macau, China, 2019: 179 - 183. DOI: 10.1109/NANO46743.2019.8993930.

[24] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//*Proc International Conference on Neural Information Processing Systems*. Long Beach, CA, USA, 2017:30. DOI: 10.5555/3295222.3295349.

[25] Liu X, You J L, Wu Y L, et al. Attention-based bidirectional GRU networks for efficient HTTPS traffic classification[J]. *Information Sciences*, 2020, **541**: 297 - 315. DOI: 10.1016/j.ins.2020.05.035.

[26] Xu X, Deng J, Cummins N, et al. Autonomous emotion learning in speech: A view of zero-shot speech emotion recognition[C]//*Proc Annual Conference of the International Speech Communication Association* (*INTERSPEECH*). Graz, Austria, 2019: 949 - 953. DOI: 10.21437/Interspeech.2019-2406.

[27] Dong X, Williamson D S. An attention enhanced multitask model for objective speech assessment in real-world environments[C]//*Proc IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*). Virtual Barcelona, Spain, 2020: 911 - 915. DOI: 10.1109/ICASSP40776.2020.9053366.

[28] Xu X, Deng J, Zhang Z, et al. Rethinking auditory affective descriptors through zero-shot emotion recognition in speech[J]. *IEEE Transactions on Computational Social Systems*, 2022, **9**(5): 1530 - 1541. DOI: 10.1109/TCSS.2021.3130401.

基于注意力双向门控循环神经网络的索驱末端效应器运动估计

徐新洲[1] 陈永发[1] 刘光明[2] 李紫千[2] 赵 力[3] 王正雨[2]

([1]南京邮电大学物联网学院, 南京 210003)
([2]合肥工业大学机械工程学院, 合肥 230009)
([3]东南大学信息科学与工程学院, 南京 210096)

摘要: 为实现对索驱末端效应器运动状态的自适应估计,基于注意力双向门控循环神经网络,提出了一种数据驱动的末端效应器运动估计方法. 首先进行数据构建,以获取短时序列作为训练样本. 然后,将数据输入包含带自注意力模块的双向门控循环神经网络,构造样本的序列模型. 最后,在索驱末端效应器运动数据集上,将电机位置、速度以及作为系统控制信号的输入时间序列作为样本特征,进行运动估计性能对比实验. 结果表明,相比常用的序列建模回归算法,所提方法能够取得更好的末端效应器运动估计性能,因而能有效实现复杂条件下对索驱末端效应器的运动估计.

关键词: 索驱末端效应器;运动估计;双向门控循环神经网络;注意力机制

中图分类号: TP241.3