Journal of Computational Science &Engineering 49 (2020) 1160-1165

Available at http://www.asocse.org ISSN 1710-4068

Modeling and Evaluation of Gait Phase Identification System based on Decision Tree and Transition Sequence

Rui Xue, Shixian Zeng, Jiangbo Chen

Qingdao Huanghai College, Qingdao in China

Abstract: **The ageing of the population and high-risk diseases associated with age have contributed to the focus on rehabilitation and exercise assist devices. In order to provide perfect control of the wearable knee orthosis at the appropriate time, proper Gait Phase Recognition (GPR) is necessary to reduce the incidence of knee collapse and fall events.In the study, the gait data of knee flexion angle, thigh segment angular velocity and acceleration were trained and tested by Logistic Model Decision Tree (LMDT) classifier.The Transition Sequence Verification and Correction (TSVC) algorithm is applied to correct the continuous class prediction model and improve the GPR performance.Finally, the sensitivity, specificity, accuracy and other correlation coefficients of the system are verified. Compared with other methods, the results show that the LMDT-GPR system can be used for real-time orthosis control with high feasibility.Combined with the TSVC algorithm, the complexity of the sensor system can be reduced. It is concluded that the gait phase identification system is designed to provide more convenience for users.**

Keywords: real-time aligner; logic model decision tree; gait recognition; transition sequence

Received 26 October 2020, Revised 26 November 2020, Accepted 30 December 2020

1.Introduction

Attitude Control Knee Orthosis is a wearable walking aid that prevents the knee from collapsing during weight bearing and it can provide unobstructed knee movement during work. The knee orthosis improves joint performance and provides a more natural gait compared to conventional fixed orthoses for patients with knee extensor disease[1-3]. Fewer gait compensation mechanisms can reduce associated joint loads, reduce energy consumption [4,5], increase foot clearance, increase walking speed, and improve overall user satisfaction [6,8].

The embedded system-controlled knee orthosis guides the accurate movement of the knee joint by adjusting the posture and swing using a plurality of electronic sensors connected to various orthopaedic limbs [5,7].The system combines electronic sensors with algorithms to enhance the knee function and reliability of the orthosis, increasing the versatility of the surface and walking speed, and the knee orthosis can also switch between different gait modes [9-11].Ideally, an effective posture control system can provide support during patient standing and provide unobstructed knee movement during swing to achieve natural gait. Knee stability is critical for safe gait, but knee extensor movement is not always required [12,13]. In addition, the standing control system must recognize the load response to safely lock and prevent knee collapse [14]. If the system can provide a release transition point to the knee, then a safe gait can be ensured without applying a load to the knee extensor movement.

The system proposes a method of identifying gait phases that provides local gait data from the thighs and knees accurately.Accurate gait phase identification controls the horizontal and vertical stride speeds across different horizontal surfaces, providing real-time basic information to the control system. The project study proposes that the new knee orthosis combines the signal characteristics of the extracted thighs and knees with the data characteristics of the decision tree estimates.The logistic model decision tree is improved by the transformation sequence verification and correction algorithm. The horizontal surface velocity, walking speed or individual gait changes are classified in stages, which can provide highly accurate GPR performance and more suitable system design.

2.Data Collection

Walking data from multiple healthy testers was collected in a rehabilitation virtual environment system. The system includes a six-degree-of-freedom force plate platform, a twin tread instrument treadmill with 1m×2m, a screen virtual projection with 180°, and a 3D motion capture system.All joint and body positions were defined by the tester's whole body marker data set [15,16] and the ground reaction force was recorded. Each tester was tested at a fixed tilt angle and speed, and the joint and body part trajectories were imported into the 3D motion capture system for bio mechanical analysis.The fourth-order low-pass Butter Worth Filter with a small cutoff frequency is used to mark the data before calculating the joint and segmentation motion data.The knee flexion and extension angle, the angular velocity and acceleration of the thigh segment of each axis were extracted from each walking state, and finally the data was imported into MATLAB modeling

3.Gait Phase Recognition

The segmentation of gait phase data can be achieved by a custom Matlab program, which is decomposed according to the gait movements such as initial contact,

2020@ The American Computational Science Society. All rights reserved.

Available at http://www.asocse.org ISSN 1710-4068

standing, foot and maximum knee flexion [10,12], knee flexion and extension angle (KA), per The angular velocity (AV) (x, y, z) and acceleration (AC) (x, y, z) signals of the thigh segments of the axes are divided into four gait phases: load response (LR), start (PO), swing (Sw) and terminal wobble (TSw). Among them, the thigh segment acceleration and angular velocity results are used to supplement the signal, and the features are extracted for estimation. It is a velocity composite signal in equation 1, where $\upsilon_x, \upsilon_y, \upsilon_z$ are acceleration and angular velocity signals from x, y, z directions respectively.

$$\sqrt{\upsilon_x^2 + \upsilon_y^2 + \upsilon_z^2} \qquad (1)$$

The correlation feature was selected for preliminary analysis [12], and 20 features were extracted: mean value of knee flexion and extension angle (x-axis), variance (x-axis), maximum difference (x-axis), and minimum value (x-axis), thigh Average angular velocity (x-axis, y-axis), maximum difference (x-axis), and minimum value (y-axis), average thigh acceleration (y-axis, z-axis), maximum (y-axis), and peak-to-peak value The sum (x-axis, y-axis, z-axis), and the main frequency (y-axis) of the angular velocity of the thigh segment, the correlation coefficient between the z-axis, and the y-axis acceleration. The sum of the scores of each data point in the sliding window is the sum of the signals, where the negative value is decremented by one and the positive value is incremented by one. All features are extracted from the 0.1-s sliding window with 90% overlap in increments of 0.01 s and the phased gait parameters are marked according to the last data point [10,12] in each window.

3.1Decision Tree Design

The feature vector is extracted for data estimation, and the logistic model decision tree [13] classifier is used to train the data hierarchically and tested by cross-validation. The performance set is provided by the training set, and the decision tree generates a logistic regression model at the terminal leaf node. The function takes the feature vector computed from the sliding window as input and returns the gait phase of the decision tree model as the predicted value with the greatest probability. The hyperbolic function equation is shown in equation 2 and is used for the calculation of the gait phase probability.

$$P(\lambda_i) = \frac{e^{\lambda_i}}{e^{\lambda_i} + 1} \qquad (2)$$

Among them, λ is linear combination coefficient of the multivariate feature for logistic regression, i is coefficient in the gait phase, λ is corresponding to i.

3.2 Conversion Sequence Verification and Correction

After the gait phase recognition GPR model performs class output, "Transform Sequence Verification and Correction" (TSVC) algorithm is used to calculate the discontinuous class function, and check whether several consecutive sequence signals in the current gait phase are the same. If they are different, the outliers are classified and re-marked. The combination of the GPR model and the TSVC algorithm can also indirectly identify the gait characteristics, and the data of the 20 data points of the knee joint flexion and extension angle (KA), the thigh segment angular velocity (AV) and the acceleration (AC) sensor are stored in the sliding window, and then Logistic regression model decision tree is used for feature extraction. Finally, the algorithm is used to verify and correct the continuous gait phase sequence and output the classification. At this time, the system extends the sliding window for a certain period of time and classifies the next set of data. The gait recognition flow chart is shown in Figure 1.

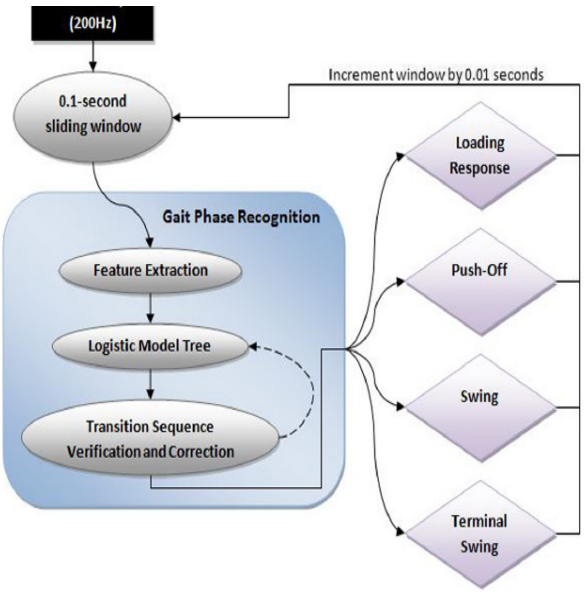

Fig. 1 Gait phase recognition flow diagram

3.3Evaluation of decision tree classifiers

The performance indicators of classifier include tree size, number of leaf nodes, overall accuracy of the classifier (Acc), sensitivity (Sens), specificity (Spec), accuracy (Prec), etc. The feature matrix that is provided by the system is used as a test set, and the classification index of the LMDT with TSVC is calculated according to the confusion matrix in Matlab2016b. The weighted average is as follows:

$$W_\psi = \frac{\sum_i \psi_i^* I_i}{\sum_i I_i} \qquad (3)$$

2020@ The American Computational Science Society. All rights reserved.

Available at http://www.asocse.org ISSN 1710-4068

Wherein, Ψ is the classification indicator, i is specific gait phase, I is the total number of instances in the total feature matrix related to each gait phase.

3.4 Analysis of results

The gait phase recognition LMDT has a size of 1643 with 822 leaf nodes and there is a logistic regression function at each leaf node. The average knee angle (degrees) and standard deviation (shaded area) for each horizontal surface condition (row) and walking speed (column) are shown in Figure 2. The average thigh angular velocity (rad/s) and standard deviation for each horizontal surface condition (row) and walking speed (column) are shown in Figure 3. The average thigh acceleration (m/s^2) and standard deviation for each horizontal surface condition (row) and walking speed (column) are shown in Figure 4. The gait phase of the classification is shown in Figure 2, 3, 4, include load response (blue), push (red), swing (cyan), and terminal swing (blue-green). The gait phase during the stride segmentation by the decision tree is in good agreement with the actual measured data. The shaded area in the Figure is the measured data, and the vertical line displayed between the shaded areas indicates the gait used for stride segmentation. signal. Classification by LMDT indicates that the GPR model studied can detect these signals indirectly under walking conditions. As shown in Figure 2, knee motion is reduced when standing slowly, but regardless of the horizontal surface condition and walking speed, the GPR model is able to distinguish all gait phases.

The statistics of the LMDT classification indicators in the training set matrix and the verification set matrix are shown in Table 1. It can be seen from the gait classification indicators, the overall accuracy rate of the training indicators is 98.38%, and all the classification indicators are greater than 0.98, and the effect is ideal.

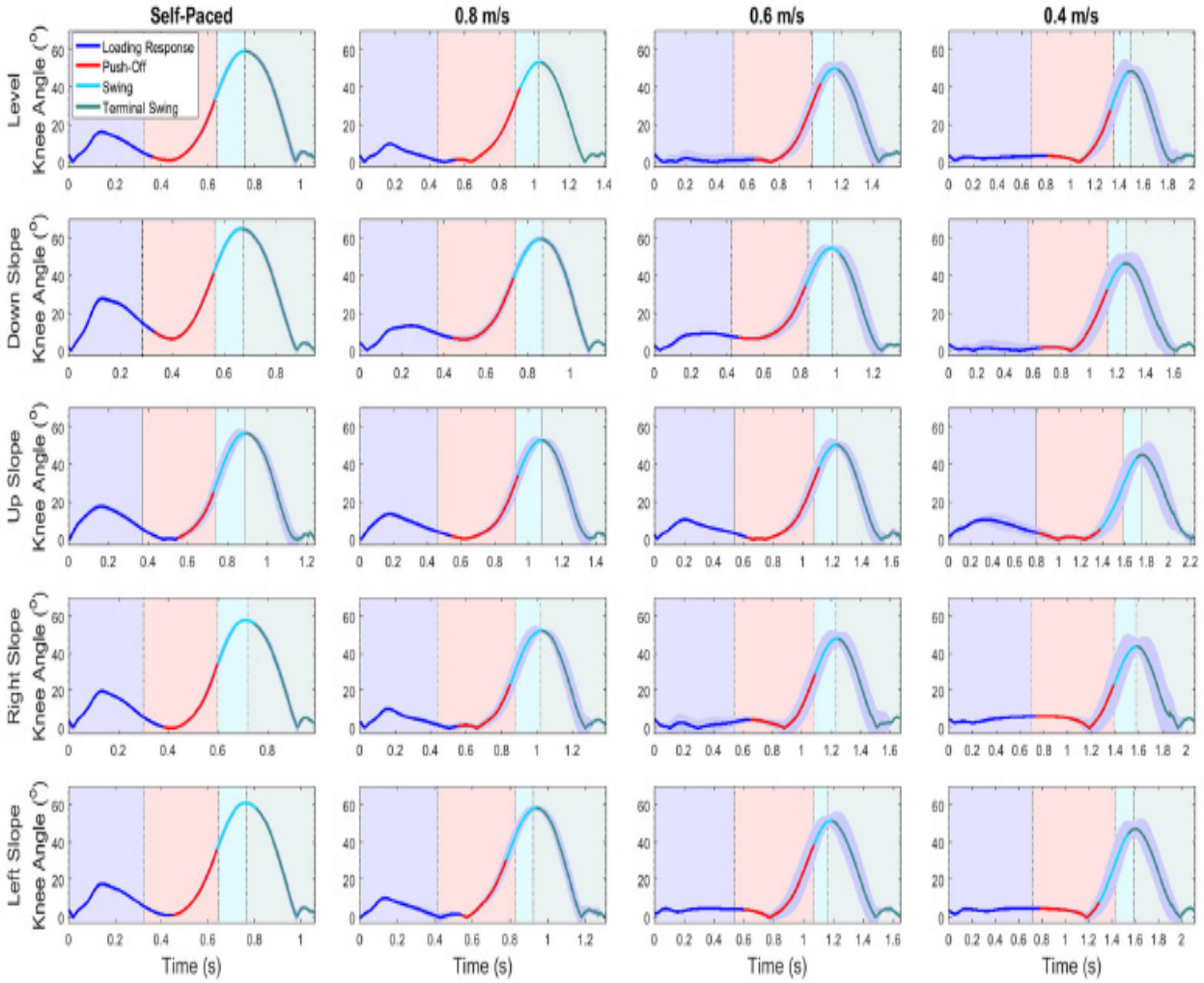

Fig. 2 Average knee angle (degrees) and standard deviation from a single participant for each surface condition (rows) and walking speed (columns),showing classified gait phases: Loading Response (blue), Push-Off (red), Swing (cyan), Terminal Swing (teal) for the training set.

2020@ The American Computational Science Society. All rights reserved.

Available at http://www.asocse.org ISSN 1710-4068

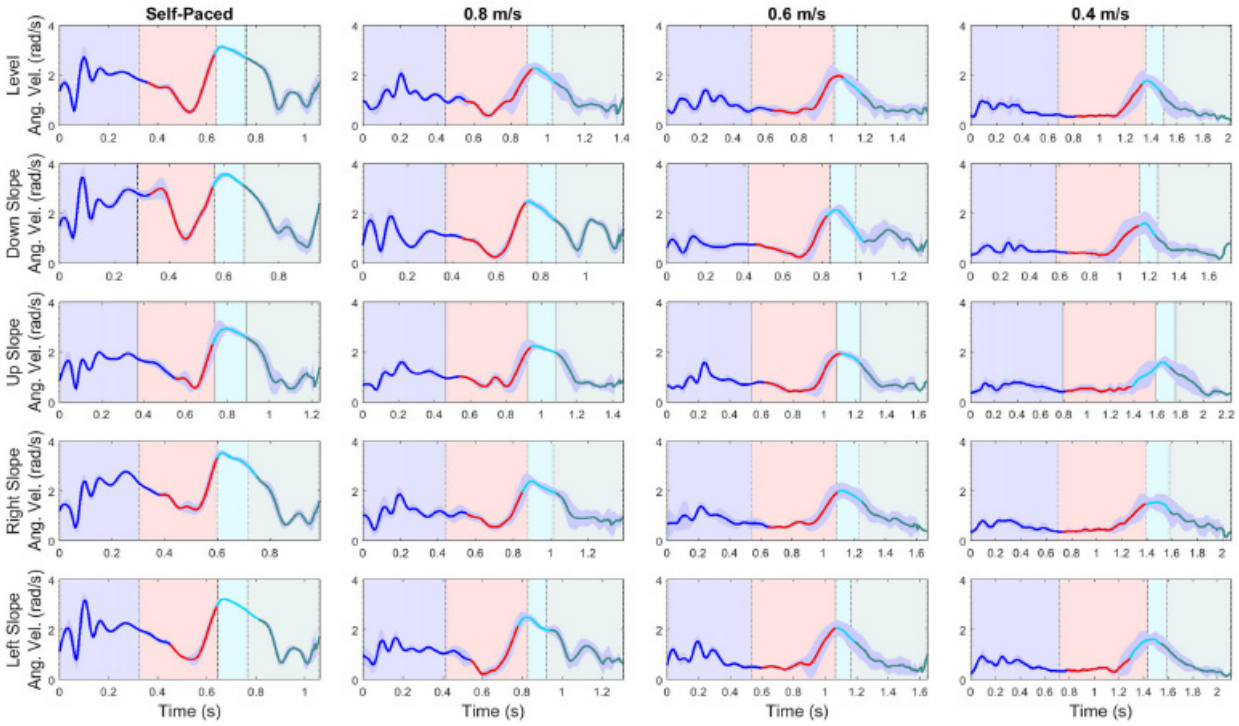

Fig. 3 Average thigh-segment angular velocity (rad/s) and standard deviation (shaded) from a single participant for each surface condition (rows) and walking speed (columns), showing classified gait phases: Loading Response (blue), Push-Off (red), Swing (cyan), Terminal Swing (teal) for the training set.

Accuracy was improved with TSVC at 0.38%, the sensitivity was reduced by 1% and the specificity remained unchanged. After applying TSVC, all validation set classification indicators were improved, sensitivity and accuracy were increased by 6%, specificity was increased by 2%, and all results were greater than 0.97.

The classification index statistics for each gait stage are shown in Table 2. The classification indicators of all gait parameters in the training set are greater than or equal to 0.97. TSw performed best, that all classification indicators at 0.99, and Swing had the lowest accuracy. Except for PO, the classification index of the verification set is greater than 0.85, the sensitivity of PO is the lowest, and the specificity of LR is the lowest. In terms of classification performance indicators, Swing and TSw performed similarly and they were superior to LR and PO.

The performance index parameter values of a specific gait phase are shown in Table 3. All results in the training set were greater than 0.91. Compared with the results in Table 2, the sensitivity of PO and Swing was lowered, and the LR and TSw data were unchanged. TSw performed best on all classification indicators, equal to 0.99. For the validation set, all metrics except Swing's sensitivity reduction are greater than 0.90. All classification indicators for LR and TSw increased compared to the training set. The LR classification performance in the verification set is better than other gait indicators.

Table 1 Logistic model tree classification performance

odel	Data Set	Accuracy	Sensitivity	Specificity	Precision
LMDT	Training	0.9838	0.98	0.99	0.98
	Validation	0.906	0.91	0.97	0.91
LMDT+TSVC	Training	0.9876	0.97	0.99	0.97
	Validation	0.9861	0.97	0.99	0.97

2020@ The American Computational Science Society. All rights reserved.

Available at http://www.asocse.org ISSN 1710-4068

Table 2 LMT-GPR performance by gait phase

Data Set	Gait Phase	Sensitivity	Specificity	Precision
Training	Loading Response	0.98	0.99	0.98
	Push-Off	0.98	0.99	0.98
	Swing	0.98	0.99	0.97
	Terminal Swing	0.99	0.99	0.99
Validation	Loading Response	0.91	0.95	0.87
	Push-Off	0.82	0.97	0.93
	Swing	0.96	0.99	0.92
	Terminal Swing	0.99	0.96	0.91

Table 3 LMT TSVC + GPR performance by gait phase

Data Set	Gait Phase	Sensitivity	Specificity	Precision
Training	Loading Response	0.98	0.99	0.97
	Push-Off	0.96	0.99	0.94
	Swing	0.91	0.99	0.97
	Terminal Swing	0.99	0.99	0.99
Validation	Loading Response	0.99	0.99	0.99
	Push-Off	0.98	0.98	0.90
	Swing	0.87	0.99	0.95
	Terminal Swing	0.97	0.99	0.99

This study validated the feasibility of the gait phase identification system through modeling and guided the control of the orthosis. The gait phase recognition classifier design combined with the Logistic control model decision tree can reliably identify the gait phase of different walking speeds [9]. Since our GPR model is trained from a data set of simple data features extracted from a small data sliding window, the data comes from the measured thigh and knee signals and can be performed under various walking conditions. The results can be directly converted to real-time knee joints. Orthosis control. In the study, the classifier generates a logistic regression model at each terminal node to process multiple types of target variables. Correct orthopaedic data produces a smaller tree than the normal classification tree [16], and smaller decision trees generate simpler models. For GPR in real-time orthotic control, as long as the classification decision can be made within an appropriate time period, high precision and universality are more important than tree size, and the classification performance of all indicators is also changed good. The GPR model has good classification performance, and the model can be used as the basis for the local intelligent sensor system control in the attitude control knee orthosis. Extracting the sensor signal characteristics of the thigh and knee from the LMT model can provide highly accurate GPR performance, and then improve the classification results by the TSVC algorithm. Finally, different data validation sets are used to evaluate the GPR model.

4. Conclusion

The accuracy of the gait classification signal is improved by the TSVC algorithm through the verification set, because the gait phase is continuous during walking, and during the gait phase transition, the gait signal is exited from one stage and proceeds to the next stage for continuation. Therefore, in a region where the gait phase is converted to another gait, if there is no significant gradient change in the signal, classification does not occur. For the knee orthosis, the knee joint is ideally released during PO, and during the pre-swing, the knee flexion data begins to occur when the weight is lowered, which is a safe knee joint release point. The system can help achieve this ideal PO swing transition point by adjusting the control of the orthosis. In summary, the use of TSVC improves the classification performance of GPR models with validation sets and is close to the training

2020@ The American Computational Science Society. All rights reserved.

Available at http://www.asocse.org ISSN 1710-4068

set results. The gait phase recognition based on the logistic regression model decision tree successfully identifies the loading response of multiple different surfaces, knee push, swing and terminal swing gait phases and a series of walking speeds. The input data is located in the thigh and knee. . To achieve higher GPR performance, gait phase sequence verification and correction algorithms are essential. The study provides a basis for real-time orthotic control in multiple real-world walking scenarios, enhancing the accuracy of gait phase recognition. Local sensors at the thighs and knees reduce the complexity of the sensor system and provide an integrated modular unit for the microprocessor-controlled knee orthosis.

5.Acknowledgments

This work was financially supported by the Science and Technology Plan Projects of Higher Education in Shandong Province (J17KB151).

References

[1]W.J. Cao, X.D. Wei, W.L. Zhao, Q.L. Meng, H.L, Yu. Journal of Biomedical Engineering, (2018) 35(05):733-739.

[2]Z.J. Liu, X.Z.Gao, X.D. Zhao, L.L. Chen. Chinese Journal of Scientific Instrument, (2018) 39(07):74-82.

[3]B.K. Xuan, Z.J. Liu, L.L. Chen, P. Yang. Journal of Southeast University(Natural Science Edition), (2017) 47(06): 1107-1116.

[4]H. Yue, P. Yang, L.L. Chen, Y.L. Geng. Journal of Clinical Rehabilitative Tissue Engineering Research, (2012) 16(26):4872-4876.

[5]Z.J. Liu. Intelligent lower limb prosthesis system based on flexible iterative learning control [A]. Intelligent Automation Professional Committee of China Automation Society. Proceedings of China Intelligent Automation Conference 2011 (Part 1) [C]. Intelligent Automation of China Automation Society Professional Committee: Intelligent Automation Professional Committee of China Automation Society, 2011: 6.

[6]D.C. Shi, L.Q. Jia, C. Liang, X.X. Wang. Computer Engineering, (2017) 43(10): 198-202+208.

[7]Y.Y. Gao, M. Meng, Z.Z. Luo, Q.S. Qi. Journal of Transduction Technology, (2011) 24(11):1574-1578.

[8]F. Tian, M.S. Hefzy, M. Elahinia. Ann Biomed Eng. (2015) 43:427–41.

[9]E.D. Lemaire, R. Samadi, L. Goudreau, J. Kofman. J Rehabil Res Dev, (2013) 50:43–52.

[10]J.D. Farah, E.D. Lemaire, N. Baddour. IEEE Symp Med Meas Appl, (2017) 263–268

[11]J.D. Farah, E.D. Lemaire, N. Baddour. Trans neural Syst Rehabil Eng Publ, (2017).

[12]J. Thiele, B. Westebbe, M. Bellmann, M. Kraft. Biomedizinische Technik. Biomed Eng, (2014) 59(1):65–77.

[13]M.J. Lawn, M. Takashima, M. Ninomiya, J. Yu, K. Soma, T. Ishimatsu. IEEE Sens, (2015) 1–4.

[14]T. Schmalz, E. Pröbsting, R. Auberger, G. Siewert. Prosthetics Orthot Int, (2016) 40:277–286.

[15]J.M. Wilken, K.M. Rodriguez, M. Brawner, B.J. Darter. Gait Posture, (2012) 35:301–307.

[16]N. Landwehr, M. Hall, E. Frank. Mach Learn, (2005) 59:161–205.

2020@ The American Computational Science Society. All rights reserved.