A reward–punishment feedback control strategy based on energy information for wrist rehabilitation

Jiajin Wang1,2, Jiaji Zhang1, Guokun Zuo1, Changcheng Shi1, and Shuai Guo2

Abstract
Based on evidence from the previous research in rehabilitation robot control strategies, we found that the common feature of the effective control strategies to promote subjects’ engagement is creating a reward–punishment feedback mechanism. This article proposes a reward–punishment feedback control strategy based on energy information. Firstly, an engagement estimated approach based on energy information is developed to evaluate subjects’ performance. Secondly, the estimated result forms a reward–punishment term, which is introduced into a standard model-based adaptive controller. This modified adaptive controller is capable of giving the reward–punishment feedback to subjects according to their engagement. Finally, several experiments are implemented using a wrist rehabilitation robot to evaluate the proposed control strategy with 10 healthy subjects who have not cardiovascular and cerebrovascular diseases. The results of these experiments show that the mean coefficient of determination ($R^2$) of the data obtained by the proposed approach and the classical approach is 0.7988, which illustrate the reliability of the engagement estimated approach based on energy information. And the results also demonstrate that the proposed controller has great potential to promote patients’ engagement for wrist rehabilitation.

Keywords
Wrist rehabilitation, robot, adaptive control, engagement estimated, reward–punishment feedback

Introduction
Stroke has become one of the major diseases that threaten people’s physical and mental health in the world.1 Loss of control of upper limbs is a common impairment underlying disability after stroke for patients, which seriously affects their daily activities.2 Traditional physical therapy is labor intensive and requires great energy of therapists.3 With the development of robotics, the emergence of rehabilitation robots provides a new way for rehabilitation.4 Rehabilitation robots are able to assist patients to complete training tasks without therapists. In addition, rehabilitation robots are capable of estimating patients’ rehabilitation status accurately through a variety of sensors, which helps therapists to develop a follow-up treatment plan for patients.

Control of rehabilitation robots, however, remains an open-ended research area. Control strategies, which target subjects ranging from the mildly impaired and severely impaired to the severely impaired and healthy individuals, are urgently required. Various control strategies for rehabilitation robots have been proposed, ranging from model-based control, model-free control, to hybrid control strategies.5–9 Each strategy has its own advantages and disadvantages, and the selection of the appropriate strategy depends on the specific application and desired performance指标.

1 Cixi Institute of BioMedical Engineering, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Science, Ningbo, Zhejiang, China
2 Department of Mechanical Automation Engineering, School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China
Corresponding author: Jiaji Zhang, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Science, Ningbo, Zhejiang 315201, China. Email: zhangjiaji@nimte.ac.cn
impaired, are the most extensively investigated controller paradigm in the rehabilitation robotics community and have been proved to be the most promising techniques for promoting recovery after stroke.\textsuperscript{5,6} There is strong evidence that high engagement in rehabilitation training induces neural plasticity.\textsuperscript{7} Therefore, great attention is paid on investigating how to use robot control strategies to promote patients’ active engagement in robotic therapy.

Assist-as-needed (AAN) control strategy is one of the most popular research topics in the field of rehabilitation robots control strategies and is considered promising to promote patients’ engagement. As the name suggested, AAN control strategy emphasizes that robots only supply as much effort as a patient needs to accomplish training tasks by estimating his/her performance in real time.\textsuperscript{8} Impedance control first proposed by Hogon was applied in AAN control strategy primitively.\textsuperscript{9} Representatively, Krebs et al. proposed an AAN controller based on impedance control with MIT-Manus,\textsuperscript{10,11} which can update impedance parameters according to patients’ performance. In this case of robotic therapy, the robot provides assistance based on specific impedance parameters when the subject is not able to track the desired trajectory and does not provide assistance when the subject is able to track or exceed the desired trajectory so as to allow the subject to move voluntarily. This kind of mechanism encourages subjects to get rid of the limitations of the desired trajectory, which can be regarded as a reward and make them more active. But some subjects showed signs of slack behavior that they rely too much on the robot’s assistance to complete the task without any punishments.\textsuperscript{12} In other words, giving only rewards without punishments will cause subjects’ slackness in rehabilitation training. Therefore, it is necessary to develop control strategies exhibiting the reward–punishment feedback.

Wolbrecht et al. proposed an adaptive controller including a forgetting term to create the reward–punishment feedback mechanism.\textsuperscript{13} The adaptive law is made up of an error-based adaptive law and a forgetting law. The standard adaptive law dominates when there is a major tracking error so as to assist the subject to complete the task, while the forgetting law dominates when there is a minor tracking error so as to decay the assistance force to promote the subject’s active engagement, which forms a mechanism that gives a reward feedback to subjects by exhibiting a minor tracking error when they are highly engaged and gives a punishment feedback by exhibiting a major tracking error when they are slack. But the adaptive controller is model-based, it does not perform well when it is applied to wrist or finger rehabilitation because minor modeling deviations affect the wrist or finger much more than the upper limb. The tracking error will not change significantly regardless of the degree of the subject’s engagement.

Another improvement to the AAN control strategy was proposed by Pehilivan et al., who introduced a minimum AAN control strategy, which relied on Kalman filter to estimate subjects’ capability.\textsuperscript{14,15} According to the estimated results, the controller updates the derivative feedback gain to modify the bounds of allowable error on the desired trajectory, which also reflects the idea of reward–punishment feedback. Subsequently, Kalman filter was replaced by nonlinear disturbance observer, and the electromyography (EMG) sensors were used to estimate the subjects’ engagement.\textsuperscript{16}

To sum up, in order to promote the engagement of subjects, the common feature of the above control strategies is that they can create a reward–punishment feedback mechanism according to the subjects’ current engagement or performance. To the best of our knowledge, previous researchers have not specifically identified this mechanism. More control strategies for rehabilitation robots support this point of view.\textsuperscript{17–25}

In this article, we proposed a reward–punishment feedback control strategy to promote subjects’ engagement for wrist rehabilitation. Firstly, we utilize the energy contributed by the subject to estimate his/her engagement. The energy can be obtained by calculating the integral of the torque contributed by the subject against the position. Secondly, an adaptive controller including a reward–punishment term was proposed. Unlike the adaptive controller above,\textsuperscript{13} the included term is not constant. Instead, it updates based on the estimated results so that the controller can give reward or punishment feedback to subjects by reflecting different tracking error, which is suitable for wrist rehabilitation. Finally, the control strategy was demonstrated through experiments on healthy subjects without cardiovascular and cerebrovascular diseases operating a wrist rehabilitation robot. The contributions of this work include the development of an engagement estimated approach without any extra sensors, which greatly reduces development costs. This work also proposed an improved adaptive controller including a reward–punishment term for wrist rehabilitation, which has great potential to promote subjects’ engagement.

This article is organized as follows. The second section presents an engagement estimated approach based on energy information and a human robot coupled system modeling. The third section proposes an adaptive controller including a reward–punishment term and details the Lyapunov stability analysis. The fourth section introduces the specific implementation methods of three experiments. The fifth section presents and analyzes experimental results. Eventually, the discussion and conclusion are presented in the sixth section.

**Engagement estimated based on energy information**

We have developed a wrist rehabilitation robot, a three degree-of-freedom (DOF) device, as shown in Figure 1(a). The device is capable of independently actuating all three DOFs of subject’s forearm and wrist. Relatively, the device has three joints: flexion/extension joint,
radial/ulnar deviation joint, and pronation/supination joint can all be controlled. Each joint of the device employs both a brushless DC motor with a conveyor belt to drive. Therefore, the control methods of the three joints are similar, and this article only describes the control strategy of the flexion/extension joint.

We use the energy contributed by the subject to characterize their engagement in the process of rehabilitation training. The energy is the work done by the torque contributed by the subject, which can be calculated as in equation (1)

\[ E_s = \int \tau_s d\theta \]  

where \( E_s \) describes the energy contributed by the subject, \( \tau_s \) describes the torque contributed by the subject, \( \theta \) is the angle of the flexion/extension joint, which is shown in Figure 1(b). We can identify \( \tau_s \) from the inverse dynamics model of the human–robot coupled system as in equation (2)

\[ \tau_s + \tau_m = M\ddot{\theta} + \tau_g(\theta) + \tau_f(\dot{\theta}, \dot{\theta}) + \tau_c(\theta, \dot{\theta}) + \tau_w \]  

where \( \tau_m \) describes the torque induced by the robot, \( \tau_g(\theta) \) describes the torque induced by the robot’s gravity, \( \tau_f(\dot{\theta}, \dot{\theta}) \) describes the torque induced by the robot’s friction, \( \tau_c(\theta, \dot{\theta}) \) describes Coriolis torque, \( \tau_w \) represents the inverse dynamic model of the human wrist joint. We can obtain \( E_s \) by substituting equation (2) into equation (1)

\[ E_s = \int [M\ddot{\theta} + \tau_g(\theta) + \tau_f(\dot{\theta}, \dot{\theta}) + \tau_c(\theta, \dot{\theta}) + \tau_w - \tau_m] d\theta \]  

Friction effects acting on the system are described by \( \tau_f(\dot{\theta}, \dot{\theta}) \). Generally, we just identify the model friction in the motor and decelerator. But we observed a high influence of the conveyor belt during the identification process of friction. This influence on the system could also be modeled and identified, so the friction effects can be described by

\[ \tau_f(\dot{\theta}, \dot{\theta}) = \tau_{fd}(\dot{\theta}) + \tau_{fb}(\dot{\theta}, \dot{\theta}) \]  

where \( \tau_{fd}(\dot{\theta}) \) describes the torque induced by friction of the motor and decelerator and \( \tau_{fb}(\dot{\theta}, \dot{\theta}) \) describes the torque induced by friction of the conveyor belt.

Friction of the motor and the decelerator is only directly related to velocity. We controlled the velocity range of the joint from 0.1° s\(^{-1}\) to 30.0° s\(^{-1}\) and recorded the required motor torque to maintain a constant velocity during the movement. We used the data to fit a standard friction model that can be described by

\[ F_f(v) = F_{fc} + (F_{fs} - F_{fc}) \cdot e^{-\nu v} + k_f v \]  

where \( F_{fc} \) is Coulomb friction, \( F_{fs} \) is static friction, \( k_f \) is viscous friction, and \( \nu \) is an empirical constant. Because \( F_{fs} \) cannot be measured accurately with the position sensor. We reduced the model to the Coulomb and viscous part. In our case, friction induces a torque, so equation (5) can be reformulated to

\[ \tau_{fd}(\dot{\theta}) = \tau_{fc} \cdot (1 - e^{-\nu \dot{\theta}}) + k_f v \]  

where \( \tau_{fc} \) and \( \theta \) correspond to \( F_{fc} \) and \( v \) in equation (5).

Friction of the conveyor belt is derived from its elastic deformation. According to the force analysis, the closer hand pallet is to the horizontal position, the greater tension generated in the conveyor belt, and the greater elastic deformation will appear. So, the friction of the conveyor belt is position dependent. We controlled the joint with a specific trajectory and recorded the motor torque, which has subtracted the friction effects of the motor and the

---

**Figure 1.** The mechanical structure of the wrist rehabilitation robot. (a) The directional view of the human robot coupled system. (b) The side view of the wrist rehabilitation robot.
decelerator to fit an optimal model. In addition, there are also minor differences in different directions of motion because of the mechanism gap. Finally, we conclude the conveyor belt friction model as

\[
\tau_{fb}(\dot{q} > 0) = \begin{cases} 
    a_1 \sin(b_1 \dot{q} + c_1), & \dot{q} > 0 \\ 
    a_2 \sin(b_2 \dot{q} + c_2), & \dot{q} < 0 
\end{cases}
\]  

(7)

where \(a_1, b_1, c_1\) are the model parameters during flexion, and \(a_2, b_2, c_2\) are the model parameters during extension.

For our robot, we decided to neglect Coriolis torque due to the relatively slow movements and its minor influence on the dynamics of the robot. \(t_m\) can be calculated by the controller.

The inverse dynamic model of the human wrist is complex and not all effects are yet fully understood. For convenience, the model should be easily identified and it is not necessary to capture all effects, we decided to model the human wrist as a passive rigid body system. The anthropometric data from Winter were used to establish the wrist model as in equation (8) 

\[
\tau_w = \tau_{gw} + \tau_{fw}
\]  

(8)

where \(\tau_{gw}\) is the gravitational torque derived from the Winter data and \(\tau_{fw}\) describes the probable effects that are wrist stiffness and damping. \(\tau_w\) is solved by a model-based adaptive control law, as in equation (9)

\[
\tau_m = \gamma \ddot{\theta} + k \theta_e + b \dot{\theta}_e
\]  

(9)

where \(\gamma\) is the human–robot coupled inverse dynamics nominal model as in equation (10), \(\ddot{\theta}\) is an estimated value of real system parameter \(\dot{\theta}\), which describes the rate of the real model and the nominal model, \(k\) and \(b\) are positive constants, \(\theta_e = \theta_d - \theta\), as shown in Figure 2, \(\theta_d\) represents the desired trajectory, which is generated from a virtual environment and displayed as a red cursor on a screen. \(\theta\) represents the actual trajectory and displayed as a purple cursor on a screen. \(\theta_e\) represents the tracking error displayed as a displacement between the red cursor and the purple cursor.

Considering that the model may not be accurate and the movement capability of a subject who is in different states of recovery. We could use \(\dot{\theta}\) to estimate the real system parameter \(\dot{\theta}\) based on the standard adaptive control law. In addition, we added a reward–punishment term in adaptive law to create the reward–punishment feedback.
mechanism. The modified parameter estimates $\dot{a}$ are updated according to

$$\dot{a} = \gamma(\dot{\theta}_e + c\theta_e) - \frac{1}{\beta(E_s)} \dot{a}$$

(11)

where $c$ is a positive constant and $\beta(E_s)$ is the reward–punishment factor, which is the function of $E_s$. The first and second terms in equation (11) create an adaptive controller with the reward–punishment feedback mechanism. The first term is the standard adaptive control law and attempt to reduce tracking error. The second term attempt to reflect tracking error based on the energy contributed by the subject, which could form a training effect of reward–punishment feedback.

This form of controller was first proposed for use in upper limb rehabilitation training. The function of the second term is to make the subject’s engagement and the tracking error has a negative correlation, thus giving the reward–punishment feedback to subjects. However, the included factor in the second term was a constant, which was independent of the energy contributed by the subject. Because of the lightly quality of human wrist, tracking error will not change significantly no matter the subject’s engagement is high or low when the included factor remains constant. To make the subject’s engagement and the tracking error to have a negative correlation, we designed a linear relationship as in equation (12), which is shown in Figure 3

$$\bar{\theta}_e = -k_0 E_s + b_0$$

(12)

where the subject’s engagement is described by the energy contributed by him/her. Because $E_s$ obtained after a training cycle, the tracking error is also averaged within the training cycle and denoted by $\bar{\theta}_e$. $k_0$ and $b_0$ are positive constants. Their values determine the characteristics of the linear relationship we expect.

In equation (11), we named the included factor as the reward–punishment factor and found that adjusting the factor can change the tracking error through a series of experiments, their relationship can be represented by a discrete function as in equation (13), which is shown in Figure 4

$$\bar{\theta}_e = F(\beta(E_s))$$

(13)

Therefore, using $\bar{\theta}_e$ as an intermediate variable, we can get a series of discrete data like $\{\beta(E_s), E_s\}$ from equations (12) and (13). Since $E_s$ is a continuous variable, we fit the above discrete data and get the continuous function of $\beta(E_s)$ on $E_s$ as in equation (14), which can be displayed more intuitively in Figure 5

$$\beta(E_s) = \begin{cases} 
0.01 & E_s \leq 0.1 \\
0.214E_s^3 - 0.106E_s^2 + 0.095E_s + 0.01 & 0.1 < E_s \leq 0.6 \\
0.723E_s^2 - 1.761E_s^2 + 1.641E_s - 0.444 & 0.6 < E_s \leq 1.2 \\
0.041e^{1887E_s} - 0.041e^{1887E_s} + 4.07 \cdot 10^{-16}e^{20.917E_s}, & E_s > 1.2 
\end{cases}$$

(14)

In equation (9), $k$ and $b$ are feedback terms to increase the compliance of rehabilitation training. The values of $k$ and $b$ were chosen by trial-and-error experimentally, so that the robot could show a feel of compliant and soft. The controller diagram is shown in Figure 2.

Lyapunov stability analysis

Stability is a crucial consideration in human–robot interaction. This section describes the Lyapunov stability analysis for the proposed adaptive controller. The analysis is based on the adaptive algorithms developed by Slotine and Li and Spong and Vidyasagar. The convenience of analysis, we defined the new tracking error $s$ and the parameter estimate error $a_e$ in equations (15) and (16)
From equations (2), (9), and (10), the error equation of the system can be derived in equation (17)

\[ M_\delta s + bs = \gamma a_e \]  

The Lyapunov function can be designed as in equation (18)

\[ V(t) = \frac{1}{2}Ms^2 + \frac{1}{2}a_e^2 \]  

Taking the derivative of equation (18) along system trajectories yields

\[ \dot{V}(t) = M_\delta \dot{s} + a_e \dot{a}_e \]  

We assume that \( a \) is a slowly changing quantity whose rate of change is 0. Substituting for \( M_\delta \dot{s} \) from equation (17) and for \( \dot{a} \) from equation (11) into equation (19) gives

\[ \dot{V}(t) = -bs^2 + \frac{1}{\beta(E_s)} \dot{\gamma}a_e \]  

Because the second term in equation (20) may be positive, the system is not asymptotically stable. However, the system is stable in the sense of uniform boundedness. We see that \( \dot{V}(t) < 0 \) when

\[ \min_\delta (bs^2) > \max_\delta (\frac{1}{\beta(E_s)} \dot{\gamma}a_e) \]  

The maximum of \( \frac{1}{\beta(E_s)} \dot{\gamma}a_e \) occurs when

\[ \dot{\gamma} = \frac{1}{2}a \]  

Substituting equation (22) into equation (21) and solving for \( |s| \) gives

\[ |s| < |a| \left( \frac{1}{2b_{\min} \beta(E_s)} \right)^{\frac{1}{2}} \]  

Equation (23) describes the boundary of the human–robot system. Larger magnitudes of \( s \) will have \( \dot{V}(t) < 0 \) and the system will converge to this boundary. Also, we found a negative correlation between \( \beta(E_s) \) and \( s \), which further supports the phenomenon in Figure 4.

**Experimental protocol**

We tested the proposed controller with 10 volunteer subjects (seven males and three females, age 24–39 years (28.1 ± 2.0198)) without cardiovascular and cerebrovascular diseases. All experiments were approved by the Institutional Review Board of the Ningbo Institute of Industrial Technology, Chinese Academy of Sciences. All subjects provided informed consent.

Three separate experiments, named the contrast experiment on energy-EMG estimation, the contrast experiment on tracking error, and the contrast experiment on subject’s engagement, were implemented. The primary goal of the contrast experiment on energy-EMG estimation was to validate that the engagement estimated approach based on energy information was reliable. The goal of the contrast experiment on tracking error was to validate that the proposed controller could create the reward–punishment feedback mechanism. The goal of the contrast experiment on subject’s engagement was used to validate that the proposed controller had great potential to promote the engagement of subjects in rehabilitation training.

**Contrast experiment on energy-EMG estimation**

In the contrast experiment on energy-EMG estimation, the robot works in the conventional passive training mode, and the subject performs the wrist flexion–extension training with the aid of the robot. One flexion–extension training is a training cycle and we set the training range of the wrist to 80°. We told the subjects to slowly increase the force from a state of no force to a higher level and then gradually reduced the force to the initial state. The entire process could record almost all cases of subjects’ engagement. At the end of each training cycle, the engagement was estimated by two approaches based on energy information and EMG information.

Comparing two approaches is to verify the reliability of the engagement estimated approach based on energy information. To simplify the experimental procedure, the EMG sensor (DELSYS, Trigno Wireless EMG System) was only installed on the wrist extensor muscle of the subject’s forearm to obtain the EMG information. The engagement was estimated based on EMG information according to equation (24). At the same time, the engagement was estimated based on energy information according to equation (3)
$E_{\text{EMG}} = \frac{\text{EMG}_{\text{RMS}}}{\theta}$ (24)

where $E_{\text{EMG}}$ represents the estimated engagement based on EMG information, $\text{EMG}_{\text{RMS}}$ is the root-mean-square (RMS) of EMG signal, which can represent the energy of EMG signal.

**Contrast experiment on tracking error**

In the contrast experiment on tracking error, we tested the tracking error of two controllers facing different engagement of subjects. One has the variable reward–punishment factor updating as equation (14), the other has the constant reward–punishment factor set to 10. Similarly, we told the subjects to slowly increase the force from a state of no force to a higher level in wrist flexion–extension training. The mean tracking error for each training cycle will be recorded as the result of the contrast experiment.

**Experimental results**

We proposed an engagement estimated approach based on energy information and incorporated the estimated results into the controller. Based on the previous adaptive controller including the constant reward–punishment factor, we designed the adaptive controller with the variable reward–punishment factor for wrist rehabilitation robot, which could create the reward–punishment feedback mechanism according to different engagement of subjects. Three experiments were performed to evaluate the effectiveness of the proposed controller.

**Contrast experiment on energy-EMG estimation**

Example data from one of subjects were normalized and shown in Figure 7. On the whole, the estimated results of the two approaches are similar in trend. Further, we used the data to calculate the RMS error ($E_{\text{RMS}}$) and the coefficient of determination ($R^2$). Table 1 listed the results of 10 subjects. The mean $E_{\text{RMS}}$ of 10 subjects is 0.1165 and the mean $R^2$ is 0.7988, which can indicate that the estimated result by the two approaches has minor differences. It also verifies the reliability of the engagement estimated approach based on energy information.

**Contrast experiment on tracking error**

We calculated the mean tracking error for each training cycle and the results are shown in Figure 8. When the reward–punishment factor is constant, the mean tracking
error has barely changed whether the energy contributed by the subject is more or less. From equation (23), the value of mean tracking error is directly related to the selection of the reward–punishment factor. In contrast, when the reward–punishment factor is automatically updated by equation (14), the mean tracking error performed a significant negative correlation with the energy contributed by the subject, which can form the reward–punishment feedback mechanism. Whether this mechanism is capable of promoting subjects’ engagement has been demonstrated in the next experiment.

Contrast experiment on subject’s engagement

Two sets of experiments have proved that the engagement estimated approach based on energy information was reliable and the proposed controller could create the reward–punishment feedback mechanism. Now we need to verify whether the proposed controller can promote the engagement of subject. Figure 9 shows the energy contributed by one of the subjects per training cycle. Obviously, in the condition that the reward–punishment factor is constant, the subject initially tries to complete the training task. It is gradually found that the task can be completed well without contributing much effort. At the end of the training, no effort is contributed. However, in the condition that the reward–punishment factor varied based on the estimated results, human–robot interactive features have been improved. During the sixth training cycle, the subject developed slackness and contributed 0.1031 J, which is less than the energy contributed in previous training. The proposed controller will show a significant tracking error to remind the subject once he/she is slack. So, the energy contributed in the seventh cycle would be significantly promoted to 2.0958 J. The same situation happened in the 9th and 10th training cycle. When the subject becomes familiar with this characteristic, they will not be slack easily. Even in the later stages of training, a high level of engagement can be guaranteed.

Figure 10 shows the mean energy contributed by all subjects over 20 training cycles. The mean energy contributed by all subjects with the controller including the variable reward–punishment factor is 1.9106 J, and with the controller including the constant factor is 0.4314 J. Indeed, the proposed controller could motivate the subject to contribute more energy, significantly increasing the level of engagement.

Discussion and conclusion

In this article, the ultimate goal of our research on rehabilitation robot control strategies is to promote the
engagement of the subject in rehabilitation training. From the previous research, we summarized the following rules: the robot gives the reward and punishment feedback based on the engagement of the subject is conducive to achieve the goal. In short, if the robot rewards the subject when he/she is highly engaged, the subject will be encouraged to contribute more. If the robot punishes the subject when he/she is less engaged, the subject will be reminded to avoid slacking. Obviously, creating the reward–punishment feedback mechanism needs to solve two key technical problems: (1) The engagement of the subject needs to be quantitatively estimated in real time and (2) the controller needs to control the robot to achieve the above dynamic behavior.

To solve the first problem, we proposed an engagement estimated approach based on energy information, which is able to calculate the energy contributed by the subject to estimate the engagement after each training cycle for wrist rehabilitation. The reliability of the proposed approach was validated by contrasting with the classical approach based on EMG information.

To solve the second problem, we proposed an improved adaptive controller including a variable reward–punishment term. Because the subject’s task is to track the reference cursor as accurately as possible, the robot rewards the subject by exhibiting a minor tracking error when his/her engagement is in higher degree and punishes the subject by exhibiting a major tracking error when his/her engagement is in lower degree. Next, we selected 10 healthy subjects without cardiovascular and cerebrovascular diseases to test the two controllers above and found a significant difference. The contrast experiment on tracking error has validated that the proposed controller is able to show a significant negative correlation between engagement and tracking error for wrist rehabilitation. This dynamic behavior exhibits the reward–punishment feedback mechanism.

In summary, the significance of this article lies in (1) an engagement estimated approach based on energy information is developed. This approach relies on the motion information returned by the encoder without other expensive sensors, which greatly saves development costs. (2) An improved adaptive controller with a reward–punishment factor is proposed. Compared to the previously adaptive controller with a constant factor, the proposed controller can stimulate the subject to contribute more energy, which has great potential to promote the engagement of patients in rehabilitation training.

There are still some aspects to be improved in this study. One of them is that we used the average model to identify the torque contributed by the subject of different conditions, which may cause some inaccuracy. We try to use the neural network to learn the inverse dynamic model of the wrist in the future. On the other hand, we estimated the engagement of the subject at the end of per training cycle and the controller will give feedback based on the estimated results in the next training cycle, which may cause some delay in reward–punishment feedback and affect the interaction effect. Therefore, improving the real-time nature of the estimation will serve as the next research direction.

Acknowledgement
The author(s) showed sincere appreciation to the anonymous reviewers for their concern and suggestions to the presentation and correctness of the article.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China [No. 61973205]; National Key R&D Program of China [No. 2018YFC2001600]; Key Research and Development Program of Zhejiang Province [No. 2019C030990]; Ningbo Major Projects for Scientific and Technological Innovation [No. 2018B10073]; and National Natural Science Foundation of Ningbo [No. 2018A610071].

ORCID iD
Jiaji Zhang https://orcid.org/0000-0002-4939-8676

References
1. Feigin VL, Forouzanfar MH, Krishnamurthi R, et al. Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet 2014; 383: 245–255.
2. Bourbonnais D and Noven SV. Weakness in patients with hemiparesis. Am J Occupat Therapy 1989; 43: 313–319.
3. Dobkin BH. Strategies for stroke rehabilitation. Lancet Neurol 2004; 3: 528–536.

Figure 10. The mean energy contributed by 10 subjects over 20 training cycles.
4. Hou Z, Zhao X, Cheng L, et al. Recent advances in rehabilitation robots and intelligent assistance systems. *Acta Automatica Sinica* 2016; 42: 1765–1779.

5. Blank AA, French JA, Pehlivan AU, et al. Current trends in robot-assisted upper-limb stroke rehabilitation: promoting patient engagement in therapy. *Curr Phys Med Rehabil* 2014; 2: 184–195.

6. Basteris A, Nijenhuis SM, Stienen AHA, et al. Training modalities in robot-mediated upper limb rehabilitation in stroke: a framework for classification based on a systematic review. *J NeuroEng Rehabil* 2014; 11: 15.

7. Warraich Z and Kleim JA. Neural plasticity: the biological substrate for neurorehabilitation. *Pm&R* 2010; 2: S208–S219.

8. Cao J, Xie SQ, Das R, et al. Control strategies for effective robot-assisted gait rehabilitation: the state of art and future prospects. *Med Eng Phys* 2014; 36: 1555–1566.

9. Hogan N. Impedance control: an approach to manipulation. *J Dyn Syst Meas Control* 1985; 107: 1–23.

10. Krebs HI, Palazzolo JJ, Dipietro L, et al. Rehabilitation robotics: performance-based progressive robot-assisted therapy. *Auton Robot* 2003; 15: 7–20.

11. Krebs HI, Volpe BT, Williams D, et al. Robot-aided neuro-rehabilitation: a robot for wrist rehabilitation. *IEEE Trans Neural Syst Rehabil Eng* 2007; 15: 327–335.

12. Todorov E. Optimality principles in sensorimotor control. *Nat Neurosci* 2004; 7: 907–915.

13. Wolbrecht ET, Chan V, Reinkensmeyer DJ, et al. Optimizing compliant, model-based robotic assistance to promote neuro-rehabilitation. *IEEE Trans Neural Syst Rehabil Eng* 2008; 16: 286–297.

14. Pehlivan AU, Losey DP, and O’Malle MK. Minimal assist-as-needed controller for upper limb robotic rehabilitation. *IEEE Trans Robot* 2016; 32: 113–124.

15. Pehlivan AU, Sergi F, and O’Malle MK. A subject-adaptive controller for wrist robotic rehabilitation. *IEEE-ASME Trans Mechatron* 2015; 20: 1338–1350.

16. Pehlivan AU, Losey DP, Rose CG, et al. Maintaining subject engagement during robotic rehabilitation with a minimal assist-as-needed (mAAN) controller. In: 2017 international conference on rehabilitation robotics (eds Amirabdollahian F, Burdet E, and Masia L), New York, USA, 2017, pp. 62–67. New York: IEEE.

17. Hussain S, Xie SQ, and Jamwal PK. Adaptive impedance control of a robotic orthosis for gait rehabilitation. *IEEE Trans Cybern* 2013; 43: 1025–1034.

18. Perez-Ibarra JC, Siqueira AAG, and Krebs HI. Assist-as-needed ankle rehabilitation based on adaptive impedance control. In: *Proceedings of the IEEE international conference on rehabilitation robotics* (eds Yu HY, Braun D, and Campolo D), New York, USA, 2015, pp. 723–728. New York: IEEE.

19. Luo L, Peng L, Wang C, et al. A greedy assist-as-needed controller for upper limb rehabilitation. *IEEE Trans Neural Netw Learn Syst* 2019; 3433–3443.

20. dos Santos WM and Siqueira AAG. Optimal impedance control for robot-aided rehabilitation of walking based on estimation of patient behavior. In: 2016 6th IEEE international conference on biomedical robotics and biomechatronics, New York, USA, 2016, pp. 1023–1028. New York: IEEE.

21. Kim Y. Hybrid-mode impedance control for position/force tracking in motor-system rehabilitation. *Int J Adv Robot Syst* 2015; 12: 10.

22. Xing L, Wang XF, and Wang JH. A motion intention-based upper limb rehabilitation training system to stimulate motor nerve through virtual reality. *Int J Adv Robot Syst* 2017; 14: 8.

23. Taherifar A, Vossoughi G, and Ghafari A S. Variable admittance control of the exoskeleton for gait rehabilitation based on a novel strength metric. *Robotica* 2018; 36(3): 427–447.

24. Mounis SYA, Azlan NZ, and Sado F. Assist-as-needed control strategy for upper-limb rehabilitation based on subject’s functional ability. *Meas Control*. 2019; 52(9–10): 1354–1361.

25. Quy-Thinh D and Yamamoto S-I. Assist-as-needed control of a robotic orthosis actuated by pneumatic artificial muscle for gait rehabilitation. *Applied Sciences-Basel* 2018; 8(4): 499.

26. Guidali M, Keller U, Klamroth-Marganska V, et al. Estimating the patient’s contribution during robot-assisted therapy. *J Rehabil Res Dev* 2013; 50: 379–393.

27. Olsson H, Aström KJ, de Wit CC, et al. Friction models and friction compensation. *Eur J Control* 1998; 4: 176–195.

28. Piras G, Cleghorn WL, and Mills JK. Dynamic finite-element analysis of a planar high-speed, high-precision parallel manipulator with flexible links. *Mech Mach Theory* 2005; 40: 849–862.

29. Winter DA. *Biomechanics and motor control of human movement*. New York: Wiley, 1990.

30. Dolan JM, Friedman MB, and Nagurka ML. Dynamic and loaded impedance components in the maintenance of human arm posture. *IEEE Trans Syst Man Cybern* 1993; 23: 698–709.

31. Slotine J-JE and Li W. *Applied nonlinear control*. Upper Saddle River: Prentice-Hall, 1991.

32. Spong MW and Vidyasagar M. *Robot dynamics and control*. New York: Wiley, 1989.

33. Li C, Rusak Z, Horvath I, et al. Influence of complementing a robotic upper-limb rehabilitation system with video games on the engagement of the participants: a study focusing on muscle activities. *Int J Rehabil Res* 2014; 37: 334–342.

34. Vitiello N, Lenzi T, Roccella S, et al. NEUROExos: a powered elbow exoskeleton for physical rehabilitation. *IEEE Trans Robot* 2013; 29: 220–235.