

Anxiety detection and training task adaptation in robot-assisted active stroke rehabilitation

Guozheng Xu¹ , Xiang Gao¹, Lizheng Pan² , Sheng Chen¹,
Qiang Wang¹, Bo Zhu¹ and Jinfei Li³

Abstract

In the therapist-centered rehabilitation program, the experienced therapists can observe emotional changes of stroke patients and make corresponding decisions on their intervention strategies. Likewise, robotic-assisted stroke rehabilitation systems will be more appreciated if they can also perceive emotional states of the stroke patients and enhance their engagements by exploring emotion-based dynamic difficulty adjustments. Nevertheless, few research have addressed this issue. A two-phase pilot study with anxiety as the target emotion state was conducted in this article. In phase I, the motor performances and the physiological responses to the stroke subject's anxiety with high, medium, and low intensities were statistically analyzed, and anxiety models with three intensities were offline developed using support vector machine-based classifiers. In phase II, anxiety-based closed-loop robot-aided training task adaptation and its impacts on patient-robot interaction engagements were explored. As a comparison, a performance-based robotic behavior adaptation was also implemented. Experimental results with 12 recruited stroke patients conducted on the Barrett WAMTM manipulator verified that the rehabilitation robot can implicitly recognize the anxiety intensities of the stroke survivors and the anxiety-based real-time robotic behavior adaptation shows more engagements in the human-robot interactions.

Keywords

Rehabilitation robot, stroke patients, anxiety detection, human-robot interaction, training task adaptation, engagements investigation

Date received: 23 November 2017; accepted: 19 September 2018

Topic: Medical Robotics

Topic Editor: Arianna Menciassi

Associate Editor: Yukio Takeda

Introduction

Task-oriented repetitive rehabilitation training is becoming the state-of-the-art therapy approach for poststroke patients. These therapy tasks are traditionally implemented by physical therapists. In recent years, there is an increasing interest in using robotic devices to help providing motor rehabilitation therapy.¹ Compared with the therapist-centered therapy, robot-assisted stroke rehabilitation can not only provide a variety of highly repetitive movements and training protocols for stroke patients, but also offer objective measurements of stroke patients' functional improvements.

¹Robotics Information Sensing and Control Research Institute, Nanjing University of Posts and Telecommunications, Nanjing, China

²School of Mechanical Engineering, Changzhou University, Changzhou, China

³Department of Rehabilitation Medicine, Nanjing Tongren Hospital, Nanjing, China

Corresponding author:

Guozheng Xu, College of Automation, Robotics Information Sensing and Control Research Institute, Nanjing University of Posts and Telecommunications, No. 9, Wenyuan Road, Yadong new District, Nanjing 210023, China.

Email: xu_gz@hotmail.com



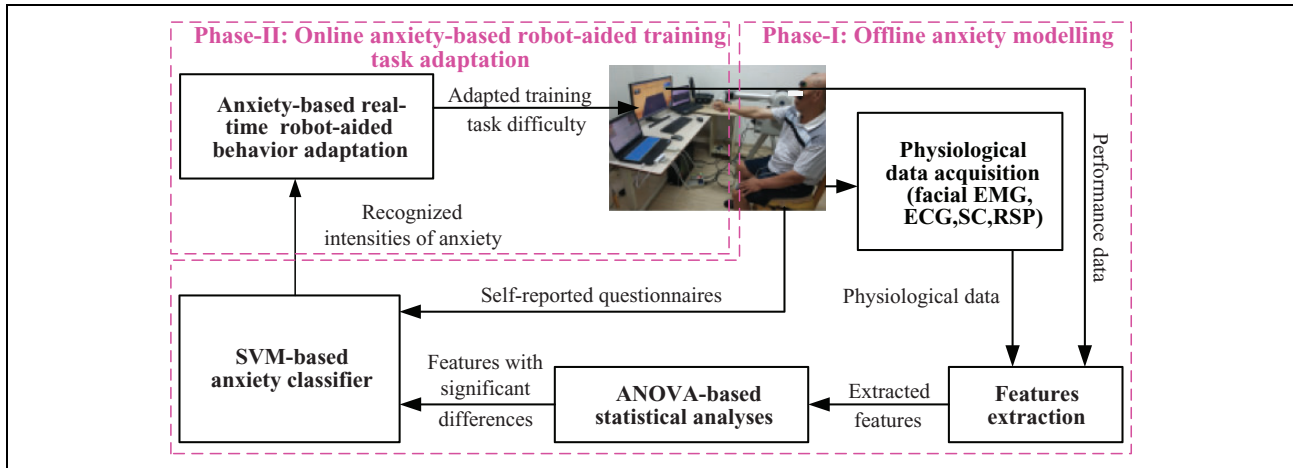


Figure 1. Overview of the anxiety-based robot-aided training task adaptation system. SVM: support vector machine; EMG: electromyogram; ECG: electrocardiogram; SC: skin conductance; ANOVA: analysis of variance.

Stroke patients' active engagements in rehabilitation training have been shown to be a very positive factor to the success of rehabilitation.² Early rehabilitation robots are able to provide active assistance to stroke patients, but do not take into account individual properties, spontaneous intentions, or voluntary efforts of that particular person. These problems were addressed by integrating the patients into the sensorimotor control loop. By recognizing the patients' active motor abilities or movement intentions, the human-in-the-loop rehabilitation robotic systems are able to optimize participation and support the patients only as little as needed.³ However, stroke patients' active involvements in the existing rehabilitation robotic systems are mostly considered from biomechanical and bioelectrical viewpoints, where the patients' active force/position signals^{4,5} or electrical activities of the brain and the muscles^{6,7} were recorded.

In the therapist-centered program, the therapists who work with the stroke patients can not only perceive the patients' active motor involvements, but also continuously monitor the patients' emotion changes in order to make appropriate decisions on their intervention strategies. The stroke patients are particularly vulnerable to anxiety and frustration, which requires to plan tasks at an appropriate level of difficulty. Likewise, robotic-assisted stroke rehabilitation training systems will be more appreciated if they can perceive the stroke patients' emotion changes and make emotion-based dynamic difficulty adjustment. Offering insights into the patients' emotional changes and adapting emotion-based behavioral interventions are known as another critical factor to successful stroke rehabilitation.⁸ Nonetheless, very few research on robot-assisted stroke rehabilitation are specifically addressed how to automatically recognize and respond to the emotion changes of the stroke survivors. One possible reason is that there are some difficulties in perceiving the stroke patient's emotion states.

There are several modalities such as facial expression,⁹ vocal intonation,¹⁰ body gesture,¹¹ and physiology¹² that can be utilized to recognize the emotion states of individuals in human-robot/computer interaction. Nevertheless, the patients with chronic stroke are often characterized by dull facial expression, severe aphasia, and limb motor dysfunction. These vulnerabilities place limits on observational, conversational, and limb methodologies to recognize the stroke survivors' emotional states. Physiology-based measurements are far more robust against these difficulties because they are noninvasive and further the psycho-physiological signals can be continuously available without the stroke patient's active cooperation. Besides, evidences show that the transition from one emotion state to another is accompanied by dynamic shifts in indicators of autonomous and central nervous system activity.^{13,14}

In this article, anxiety, which can be easily evoked by training tasks with different difficulties in clinical rehabilitation therapy, was chosen as the target emotion state. The primary focuses of the current research were firstly to offline evaluate anxiety with high, medium, and low intensities and then to carry out real-time anxiety-based robot-aided rehabilitation training task adaptation.

Methods

The block diagram of the anxiety detection and subsequent anxiety-based robot-aided training task adaptation system are shown in Figure 1. It consists of two consecutive phases: offline anxiety modeling (phase I) and online anxiety-based robot-aided training task adaptation (phase II). In phase I, the features, from the physiological and the motor performances recordings of the stroke subjects under anxiety with high, medium, and low intensities, were firstly extracted and then subjected to analysis of variance (ANOVA)-based statistical analyses to obtain the features

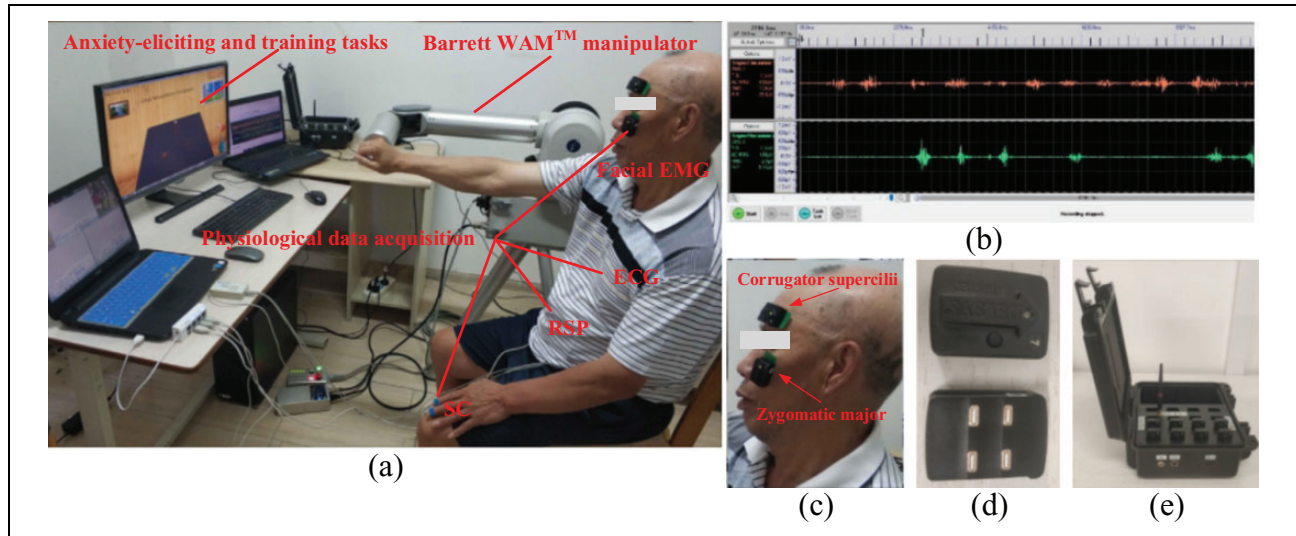


Figure 2. Anxiety detection and anxiety-based robot-aided rehabilitation task adaptation system. (a) Experimental system setup, (b) facial EMG data collections, (c) position, (d) sensors, and (e) wireless base. EMG: electromyogram.

with significant differences among three anxiety intensities. Anxiety with three intensities was further offline evaluated using support vector machine (SVM)-based anxiety classifier, in which the features with significant differences were adopted as inputs while the self-reported questionnaires as outputs. In phase II, robot-aided rehabilitation training tasks were online adapted to the recognized intensities of anxiety from the stroke subjects. Further, to demonstrate the effect of introducing anxiety of stroke patients into robot-assisted stroke rehabilitation, the impacts of anxiety-based robot-aided behavior adaptation on the stroke patient's engagements were explored using the performance-based robot-aided training task adaptation as a comparison. The details on the enrolled subjects are given in the "Subjects" section while the experimental system setup is depicted in the "Experimental setup" section. "Phase I: Offline anxiety modeling" section demonstrates the offline modeling of the anxiety with high, medium, and low intensities (phase I). The online anxiety-based robot-aided rehabilitation training task adaptation and its impacts on the stroke patient's engagements are shown in the "Phase II: Online anxiety-based robot-aided rehabilitation training task adaptation" section (phase II).

Subjects

The stroke patients, with upper extremity motor impairments and similar Brunnstrom Recovery Scale (BRS) evaluation scores, were recruited as representative of hemiparesis participants. Participants were excluded from the study if they had severe neurological disorder, senile dementia, or cognitive intact.¹⁵ Twelve stroke participants (mean age: 53.6 years, 7 males, 5 females, mean stroke time: 12.6 months, 6 right-sided hemiplegia, 6 left-sided hemiplegia, 10 stage-4 and 2 stage-3 BRS scores of upper

extremity, and 9 stage-4 and 3 stage-3 BRS scores of hand) were recruited, and all were received motor rehabilitation therapy at the Rehabilitation Medicine Center of the Nanjing Tongren Hospital of China. Before the tasks began, ethical approval was obtained from the Medical Ethics Committee of the Nanjing Tongren Hospital of China, and all subjects were informed about the procedure and that they would be video-recorded and photo-taken during the experiment. All subjects gave written informed consent concerning the use of their video footage and questionnaire data for further analysis. Of the 12 stroke participants, one stroke subject (upper extremity and hand BRS scores were both stage-3) was not able to complete phase I experiments, and the rest also took part in phase II closed-loop experiments.

Experimental setup

The experimental setup for phases I and II experiments is shown in Figure 2(a), and its configuration mainly consisted of three major parts: the rehabilitation robotic hardware, the physiological data acquisition, and the anxiety-eliciting and training tasks.

Rehabilitation robotic hardware. The current rehabilitation robotic system was constructed using the Barrett WAM™ (Newton, MA, USA) manipulator. The Barrett WAM™ manipulator is a four degree-of-freedom highly dexterous, naturally back-drivable manipulator. Its human-like kinematics and high back drivability enable inherent force-control, haptic interaction, and rehabilitation application. The Barrett WAM™-aided rehabilitation training system was extensively developed in our previous work, in which passive,⁵ active-assisted,¹⁶ and resistance¹⁷ training modes were explored. More detailed descriptions of the usages

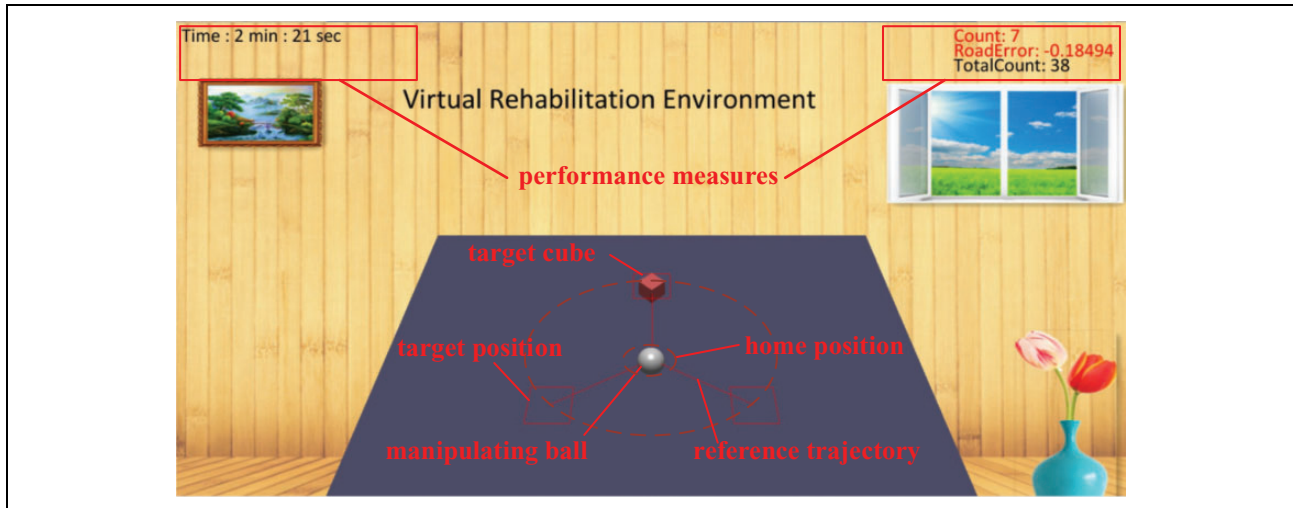


Figure 3. Target-hitting virtual rehabilitation training task.

and control of the WAM™ robot could be referred in our previous work. In the current anxiety-based robot-aided training task adaptation, an active training mode aiming to improve upper limb motor coordination was investigated for the WAM™ rehabilitation manipulator. The qualified stroke participants were required to actively manipulate the WAM™ manipulator (controlled with full gravity compensation mode) by themselves, and no assistive or resistive forces from the robot were provided. An internal PC offered by the Barrett Technology Inc. (Newton, MA, USA), running with the Ubuntu Linux-Xenomai real-time system, was responsible for executing threads including physiological and performance features extraction, anxiety intensities recognition, and anxiety-based robot-aided training task adaptation. All real-time communication between the external PC and the Barrett motor Pucks™ (motor controllers) was conducted by wireless Ethernet.

Physiological data acquisition. Participants were equipped with biosensors to acquire their anxiety-related nervous system activities including facial electromyogram (EMG from corrugator and zygomatic), electrocardiogram (ECG), skin conductance (SC), and respiration (RSP). EMG activities were collected from the zygomatic major and the corrugator supercilii muscle regions using a Delsys Trigno Wireless EMG system (Delsys Inc., Boston, Massachusetts, USA), as shown in Figure 2(b) to (e). They are often used to find the correlation between emotion and physiological reactions. In the work by Kuriakose and Lahiri,¹⁸ the EMG features were used to recognize three different emotional states for children with autism spectrum disorder. ECG activities were measured from the chest using the standard electrode configuration sensor. SC signals were acquired by positioning two electrodes on the distal phalanges of the index and middle fingers. RSP rate and intensity were captured by placing a respiration belt around the belly. Many studies have

verified that the measured peripheral parameters from the autonomous nervous system, ECG, SC, and RSP, are highly related to stress, arousal, and emotions.^{13,19} During the physiological recording, the sample rates for the facial EMG, ECG, SC, and RSP were set as 1296, 200, 50, and 50 Hz, respectively.

Anxiety-eliciting and training tasks. A virtual target-hitting training task, which aimed to coordinate multi-joint arm movements of stroke patients, as shown in Figure 3, was designed and implemented to evoke anxiety with different intensities. A $62.5 \times 37.6 \text{ cm}^2$ screen with back projection was used to display the anxiety-eliciting and rehabilitation training tasks. The stroke subjects were instructed to manipulate the ball by moving the Barrett WAM robot end-effector and were required to hold the manipulating ball at the home position until the target cube object appeared. The target position was spread evenly along a circle at a constant radius from the home position. The target cube always appeared and stayed in one of the three target positions for a predefined time interval. When the target cube appeared, the stroke subjects should move the manipulating ball to hit the target as quickly as possible while trying to track the predefined trajectory. Three task parameters, updating intervals, sizes, and sequence appearances of the target cube, could be manipulated to determine the task difficulty level and further to elicit anxiety with different intensities. In the current research, task difficulty levels were predefined from 1 to 15 based on the updating intervals. Before the experiments in phase I began, each stroke subject needed to play the target-hitting tasks several times to determine his/her skill level. This was achieved by starting from a low difficulty level and progressively increasing it until the skill level was reported by the subject or starting from a high level and decreasing it. The average of the skill levels was then considered as the subject's challenge level. Depending on this challenge level, over-

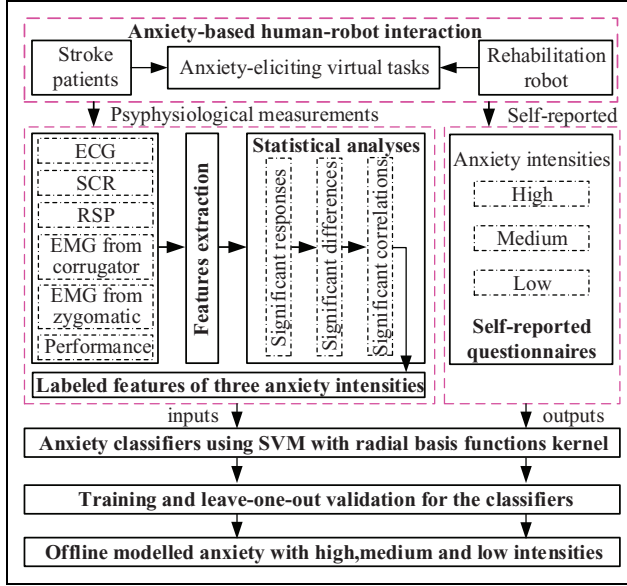


Figure 4. Overview of offline anxiety modeling with high, medium, and low intensities.

challenge and under-challenge levels were determined by adding or subtracting five levels from the subject's challenge level. Three anxiety-evoking tasks, hard (task with over-challenge level), medium (task with challenge level), and easy (task with under-challenge level), were then determined. To eliminate the habitual effects in phase I, the subjects should reexplore their three challenge tasks in phase II closed-loop studies.

Phase I: Offline anxiety modeling

Figure 4 describes the flowchart of anxiety modeling with high, medium, and low intensities. The physiological and performance data were firstly recorded using psychophysiological measurements from anxiety-based human-robot interaction and then were subjected to features extraction. The features that could differentiate three anxiety intensities were obtained through three-step statistical analyses (significant responses, differences, and correlations). Using the labeled features as inputs and the self-reported questionnaires on high, medium, and low intensities of anxiety as outputs, anxiety classifiers using SVM with radial basis function (RBF) kernel were trained. After leave-one-out cross-validations, the offline anxiety models with three intensities were evaluated.

Experimental procedure. Task-evoking experiments and offline modeling for the anxiety with different intensities were conducted in phase I studies. Experiments were carried out in a specialized room at the Rehabilitation Medical Center of the Nanjing Tongren Hospital of China. The stroke subjects, the experimenter, and the physical therapist were present. At the first visit, the stroke subjects were informed of the purpose of the experiment and signed a consent form.

When the familiarization process began, the stroke subjects explored their hard, medium, and easy tasks. Then, the subjects were asked to wear biosensors and take part in nine consecutive sessions shown in Figure 5. Each session consisted of three sequential periods (baseline, task-performing, and self-report questionnaires) and was randomly assigned three times. The physiological signals and the performance data with the optimum scores in each task were recorded. In the baseline period, the subjects took a 3-min rest and physiological measurements were record as baseline data. For the task-performing period, the subjects performed target-hitting task with one of the three tasks and the related physiological and performance data were measured. At the end of each session, the questionnaires were made, which were composed of a battery of self-reported questions related to the intensity of anxiety they felt, and the answer to each question was rated on a nine-point Likert-type scale.^{20,21} For the scales, 1 indicated the lowest intensity of anxiety while 9 indicated the highest intensity of anxiety. Meanwhile, the experimenter and the physical therapist also gave their opinions about how they thought the subject was feeling based on their visual observations. During the anxiety models development, the reported intensity levels were normalized to [0, 1] and further evenly discretized into three labeled intensities. Specifically, normalized intensity scales in [0, 0.33] were defined as low intensity of anxiety, [0.34, 0.66] were medium intensity of anxiety, and [0.67, 1] were labeled as high intensity of anxiety. According to the emotion and flow theories,²² anxiety with high, medium, and low intensities could be evoked by hard, medium, and easy tasks, respectively.

Feature extraction. A total of 45 features were extracted from the physiological signals and task performance data for each 3-min task period. They were divided into the following three categories:

Facial EMG activity from the zygomatic major and the corrugator supercilii muscle regions (18 features). Nine features were extracted from time and frequency domains by measuring the corrugator supercilii and zygomatic major muscle activities, respectively. These features were the integration (EMG_{cor_INTG} , EMG_{zyg_INTG}), root mean square (EMG_{cor_RMS} , EMG_{zyg_RMS}), standard deviation (EMG_{cor_STD} , EMG_{zyg_STD}), and mean absolute value slope (EMG_{cor_MAVP} , EMG_{zyg_MAVP}) of the corrugator supercilii and zygomatic major muscle; the mean ($EMG_{cor1stDiff_MEAN}$, $EMG_{zyg1stDiff_MEAN}$), median ($EMG_{cor1stDiff_MED}$, $EMG_{zyg1stDiff_MED}$), and standard deviation ($EMG_{cor1stDiff_STD}$, $EMG_{zyg1stDiff_STD}$) of the first difference of the corrugator supercilii and zygomatic major muscle; the mean (EMG_{cor_MPF} , EMG_{zyg_MPF}) and median (EMG_{cor_MEDPF} , EMG_{zyg_MEDPF}) values of power frequency of the corrugator supercilii and zygomatic major muscle.

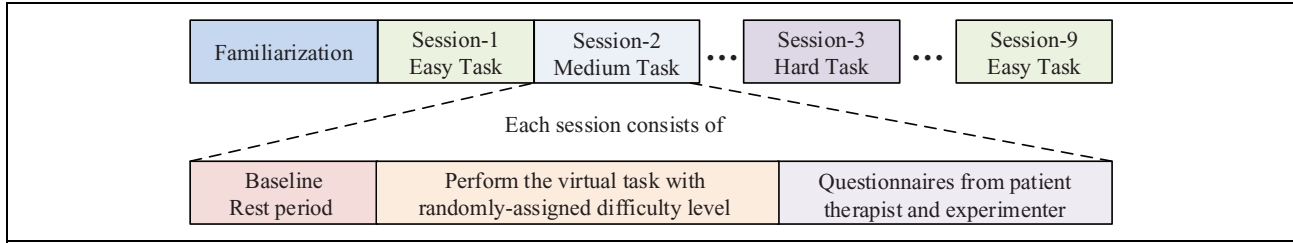


Figure 5. Experimental procedures of the anxiety intensities evoking studies in phase I.

Peripheral signals from ECG, SC, and RSP activities (22 features). From the ECG signals, mean heart rate from the intervals of the normal-to-normal (NN) heartbeats (ECG_{hr_MEAN}) was calculated. Besides, the standard deviation of heart rate (ECG_{hr_STD}) as well as several measures of heart rate variability in time-domain and spectral-domain were extracted: the standard deviation of NN intervals (ECG_{hrv_SDNN}), the root mean square of the difference of successive NN-intervals (ECG_{hrv_RMSSD}), the percentage of interval differences of successive NN intervals greater than 50 ms (ECG_{hrv_pNN50}), the low-frequency band (ECG_{hrv_LF}), the high-frequency band (ECG_{hrv_HF}), and the ratios of the LF/HF bands ($ECG_{hrv_LF/HF}$).

As to the SC signals, it has two separate components: skin conductance level (SCL) and skin conductance response (SCR). Features extracted from the SC signals included the mean (SCR_{MEAN}), standard deviation (SCR_{STD}), and error between maximum and minimum amplitude (SCR_{amp_MMERR}) of SCR; the mean ($SCR_{1stDiff_MEAN}$), standard deviation ($SCR_{1stDiff_STD}$), and error between maximum and minimum amplitude ($SCR_{amp1stDiff_MMERR}$) of the first difference of SCR; the number of SCR response peaks (SCR_{num_PEAK}); and the mean (SCL_{MEAN}) of SCL.

Concerning the RSP signals, the mean respiration rate (RSP_{rate_MEAN}), the respiration variability (RSP_{var_MEAN}), the mean amplitude (RSP_{amp_MEAN}), and the error between maximum and minimum amplitude (RSP_{amp_MMERR}) of the respiration, the mean ($RSP_{1stDiff_MEAN}$) and standard deviation ($RSP_{1stDiff_STD}$) of the first difference of RSP were calculated from the RSP signal.

Task performance (5 features). Performance measurements, as ultimate concerns, are the direct reflection of the efficacy of rehabilitation therapy program on the stroke subjects. Cai and Lin²³ have revealed that the operator's task performance varies as a result of different emotional states in a driving context. Features calculated from the task performance were the mean of a successful target hitting time ($PeRF_{time_MEAN}$), the mean ($PeRF_{trjerr_MEAN}$), and standard deviation ($PeRF_{trjerr_STD}$) of the trajectory tracking error and the number of successful target hitting ($PeRF_{num_SUC}$) and the trajectory smoothness²⁴ of a successful target hitting ($PeRF_{trj_SMOOTH}$).

Statistical analysis. All physiological and performance features were subjected to ANOVA tests to search for the differences among high, medium, and low intensities of anxiety. The following three statistical analyses were performed.

Significant responses analyses for physiological features. These analyses were conducted between baseline and each task period using each baseline corrected relative values of physiological features. One-way repeated measures ANOVA was conducted with each training task period (baseline vs. each task period) as within-subject factors and the physiological features as measures. The goal of these analyses was to determine whether each task could evoke significant psycho-physiological response.

Significant differences analyses for physiological and performance features. These analyses were performed among different task periods using the relative values of the physiological and performance features. One-way repeated measures ANOVA was implemented with the different training task periods (hard, medium, and easy) as within-subject factors and the physiological and performance features as measures. The goal of these analyses was to search for the significant differences among three different tasks that evoking anxiety with high, medium, and low intensities, respectively.

Significant correlation analyses among physiological, performance features, and self-reported questionnaires. These analyses were carried out between the self-reported questionnaires and the physiological features without significant differences but with significant responses at least in one of three task periods, as well as between the self-reported questionnaires and the performance features with insignificant differences. The goal of these analyses was to further analyze the relationships between physiological and performance features with insignificant differences and the self-reported questionnaires. If there existed significant correlations, these features with insignificant differences could still be used to classify intensities of anxiety.

The Greenhouse-Geisser correction was used in the case of violation of the assumption of Sphericity. For the post hoc multiple comparisons between different tasks, Bonferroni correction was adopted. To test the correlations between the self-reported questionnaires and the physiological/performance features with insignificant differences,

Spearman rank-order correlation was used. All statistical results with $p > 0.05$ were not reported.

Offline anxiety modeling with high, medium, and low intensities. Determining the intensities of anxiety (high, medium, and low) from the physiological/performance features and the subject's self-reported intensities of anxiety resembles a typical multi-category classification problem. Concerning this issue, the attributes were the physiological and performance features with significant differences or correlations, and the target outputs were the subject's normalized self-reported intensities of anxiety. SVM with RBFs kernel was selected to differentiate the intensities of anxiety in this research. SVM approach is able to deal with noisy data and over-fitting by allowing for some misclassifications on the training set.²⁵ This characteristic makes it particularly suitable for affect recognition because the physiological data are noisy and the training set size is often small.¹⁹ According to the theory and learning method of SVM in the study by Vapnik and Vapnik,²⁶ the RBF kernel-based SVM was summarized as in Appendix 1.

A total of three-category binary SVM classifiers were constructed using one-versus-one method for three anxiety intensity pairs (high/medium, high/low, and medium/low). Subject-independent classifiers were designed, and the accuracy of the SVM classifiers was computed using leave-one-out cross-validation method. Since the SVM classifiers were tested on the data of subjects that are not present in the training sets, these subject-independent classifiers will certainly yield a lower accuracy than the subject-dependent classifiers. Nonetheless, less time-consuming will be achieved than the subject-dependent classifiers which is necessary to train SVM classifiers for each subject. Additionally, the accuracies of anxiety-based SVM classifiers with three different types of input (physiology, performance, and combination of physiology and performance) were investigated.

Phase II: Online anxiety-based robot-aided rehabilitation training task adaptation

Experimental procedure. The anxiety-based closed-loop robotic behavior adaptation and its impacts on the engagements of the stroke patients were investigated in phase II studies. As a comparison, the performance-based robot-aided training task adaptation was also implemented. The target hitting training task conducted in phase I was still adopted. To prevent habituation, the stroke subjects should reexplore their hard, medium, and easy tasks before phase II experiments initiated. Besides, an interval of at least 30 days between the two phases as well as anxiety/performance-based training sessions was enforced.

Before the robot-aided training task adaptation began, real-time predictive accuracy on classifiers with three different types of input data was further verified. The robot-aided task adaptation training session was consisted of 20

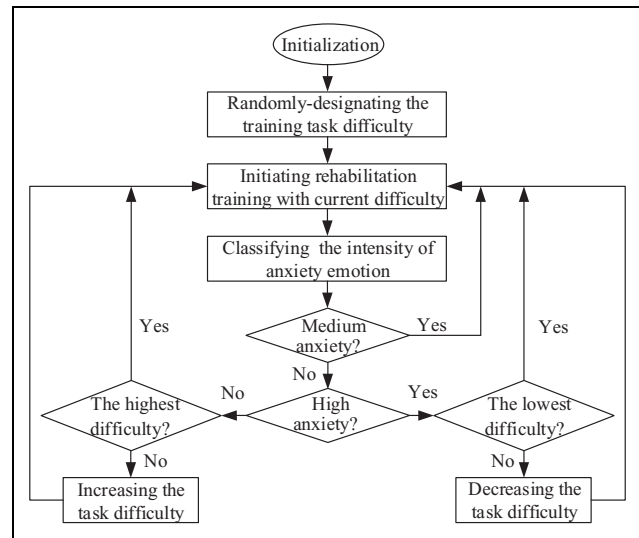


Figure 6. The flowchart of real-time anxiety-based training task adaptation.

epochs. During the experiments, the anxiety intensity model that yielded the highest real-time predictive accuracy was selected. At the end of each epoch, the task difficulty for the next epoch was determined based on the intensity prediction result in the current epoch and the anxiety-based training task adaptation mechanism. The stroke subjects were also asked about their preferences, but the task difficulty was modified according to the anxiety-based adaptor output. As to the performance-based human-robot interaction studies, similar experimental protocols were implemented. Additionally, the stroke subjects should report their task liking levels at the end of each training epoch and their overall engagement experiences at the end of the whole training session (see Section Engagements investigations for the stroke patients).

Robot-aided rehabilitation training task adaptation

Anxiety-based robot-aided training task adaptation. Figure 6 shows the flowchart of real-time anxiety-based training task adaptation mechanism. The training task difficulty level was randomly designated by the therapist at the initialization stage, and the training task adaptor was not activated until one of three different intensities of anxiety was detected. In the proposed anxiety-based training task adaptation mechanism, anxiety with high intensity resulted in a decrease in task difficulty level if the subjects were not at their lowest difficulty level, and anxiety with low intensity generated an increase in task difficulty level if the subjects were not at their highest difficulty level, while no change was made on task difficulty level if anxiety with medium intensity was recognized.

Performance-based robot-aided training task adaptation. As a comparison, performance-based robot-aided training task adaptation was investigated. According to the task requirements, the stroke subjects had to move the manipulating ball

to hit the target cube as quickly as possible while trying to track the predefined trajectory. Therefore, the two performance indices, the percentages of successful target hitting (positive index) and the mean of trajectory tracking errors (reverse index), were chosen, and their significant differences among hard, medium, and easy tasks were statistically verified (see “Performance features differences between different tasks” section).

When evaluating the subject’s task performances for a given difficulty, *excellent*, *good*, and *poor* rankings were assigned. They were determined by calculating a comprehensive performance index which was aggregated from the abovementioned two performance indices. The comprehensive performance index was generated as follows: (i) the scores of the two performance indices for all subjects were obtained by asking the subjects to play the tasks with the given difficulty; (ii) the upper and lower limits for the two performance indices were set according to the scores across all subjects; (iii) the reverse index was subject to positive and nondimensionalization treatments; and (iv) the value of the comprehensive performance index lying between 0 and 1 was calculated using weighted-sum method where each index has the same weight coefficients.

Using the abovementioned method, the thresholds of the comprehensive performance index for ranking the performances were set as 0.44 and 0.85, respectively. At the end of each training epoch, the stroke subject’s performance was rated as excellent (comprehensive performance $\in (0.85, 1)$), good (comprehensive performance $\in (0.44, 0.85)$), or poor (comprehensive performance $\in (0, 0.44)$). A similar adaptation mechanism as the anxiety-based robotic behavior adaptation was adopted in the performance-based robot-aided training task adaptation. Excellent performance resulted in an increase in task difficulty level, and poor performance caused a decrease in task difficulty level, while no changes were made if the stroke subject’s performance was rated as good.

Engagements investigations for the stroke patients. To investigate the impacts of the anxiety or performance-based training task adaptation on the human–robot interaction experience of stroke patients, several standard questionnaires referred in the work by Shirzad and Van der Loos²⁷ were made at the end of each training epoch or the whole training session. Specifically, a User Engagement Scale was used to measure the user engagement at the end of the whole training session. Five measures of the User Engagement Scale were adopted and presented with 5-point Likert-type scale questions (1 = *disagree* to 5 = *agree*; see the appendix in the work by Shirzad and Van der Loos²⁷): the task engagingness, the user’s willingness to continue the training, and the user’s focused attention, motivation, and feedback. Apart from the abovementioned questionnaires on the whole training experiences, liking level questionnaires with five-point Likert-type scale questions were made at the end of each training

epoch to gauge the level of the stroke subject’s liking on the task during each training epoch.

The pairwise *t-test* comparisons of the self-reported measures were statistically analyzed to verify whether the anxiety-based robot-aided training task adaptation has more engagement experiences than the one in the performance-based human–robot interaction. Before the *t-test* was conducted, the internal reliabilities for all the metrics of each measure were evaluated by calculating the Cronbach’s α values. The metrics of each measure with Cronbach’s $\alpha > 0.7$ were considered be internally reliable.

Results

Phase I: Offline anxiety modeling

Statistical analyses for physiological and performance features

Physiological features responses between baseline and each task. Comparisons of baseline corrected relative values of the physiological features between baseline and each task, using one-way repeated measure ANOVA, were done separately. Table 1 presents the significant physiological responses results between baseline and each task that evoking anxiety with high, medium, and low intensities, respectively. As to the extracted physiological features, it can be seen that each task can induce some significant physiological features responses.

To better illustrate how the physiological signals vary with the transition from one anxiety intensity to another, Figure 7 shows the physiological signals changes with a typical stroke subject during baseline and different task periods. As to the EMG signals from corrugator supercilii, the tasks evoking anxiety with high and low intensities can both cause sharp intensity increments compared to the baseline level, while there is no significant change for the task evoking medium anxiety. Concerning the ECG profiles, there are no significant heart rate differences for the tasks evoking high and medium intensities, but lower heart rate is found in the task with low intensity. Good differentiations between the baseline and each task can be found for the number of SCR peaks and the amplitude of the SCR signal. Compared to the baseline period, faster respiration rate is found in the tasks evoking high and medium intensities, but lower rate exists in the low intensity evoking task. Besides, the task evoking medium intensity causes more significant change than the other two intensities evoking tasks.

Physiological features differences between different tasks.

The univariate statistical tests for physiological features, using one-way repeated measure ANOVA, show that significant differences among three different tasks were found for EMG_{cor_INTG} ($F(2, 20) = 6.32, p = 0.007$), EMG_{cor_MPF} ($F(2, 20) = 3.54, p = 0.048$), EMG_{zyg_INTG} ($F(2, 20) = 3.46, p = 0.051$), ECG_{hr_rate} ($F(2, 20) = 4.66, p = 0.022$), SCR_{MEAN} ($F(2, 20) = 5.85, p = 0.01$), $SCR_{amp1stDiff_MMERR}$ ($F(2, 20) = 4.39, p = 0.026$),

Table 1. Results of significant physiological response between baseline and each task that evoking anxiety with high, medium, and low intensities, respectively.^a

	Hard evoking task				Medium evoking task				Easy evoking task			
	Mean	Standard deviation	F-value	p-Value	Mean	Standard deviation	F-value	p-Value	Mean	Standard deviation	F-value	p-Value
EMG _{cor_INTG}	5.32	1.59	11.2	0.007	2.27	1.89	6.4	0.028	5.12	1.96	9.2	0.015
EMG _{cor_MAVP}	-21.8	20.5	5.9	0.035	-15.1	19.6	17.5	0.002	-21.2	18.1	15.8	0.003
EMG _{cor_MPF}	5.22	1.68	7.4	0.023	0.25	1.18	8.9	0.013	2.64	1.68	5.4	0.043
EMG _{cor_RMS}	—	—	—	—	3.06	7.74	9.5	0.011	-8.21	9.29	12.3	0.006
EMG _{cor1stDiff_STD}	-22.5	6.07	9.2	0.015	-11.8	9.21	17.5	0.002	-22.1	6.05	13.3	0.005
EMG _{zyg_INTG}	26.9	13.1	5.2	0.046	59.1	81.8	4.8	0.053	—	—	—	—
EMG _{zyg1stDiff_MEAN}	-17.5	30.9	4.9	0.048	-22.7	40.9	6.9	0.023	-11.1	24.8	27.8	0.000
ECG _{hr_MEAN}	0.76	3.32	5.9	0.034	-3.98	4.25	15.8	0.002	0.96	4.69	7.3	0.022
ECG _{hrv_LF/HF}	26.6	36.2	5.3	0.044	—	—	—	—	60.2	79.2	5.5	0.041
ECG _{hrv_pNN50}	-28.9	19.8	4.9	0.049	—	—	—	—	-40.0	24.5	6.0	0.032
SCR _{MEAN}	35.3	11.0	5.2	0.046	29.2	8.97	20.7	0.001	24.6	9.05	5.7	0.037
SCR _{num_PEAK}	44.9	22.4	5.4	0.042	33.8	16.7	4.9	0.049	27.8	13.6	7.9	0.018
SCR _{amp1stDiff_MMERR}	81.7	46.7	5.0	0.05	55.2	63.4	4.9	0.048	—	—	—	—
RSP _{rate_MEAN}	-3.94	1.36	13.3	0.005	-2.76	1.41	11.2	0.006	0.96	1.81	18.0	0.002
RSP _{amp_MMERR}	—	—	—	—	13.7	12.8	6.7	0.025	8.98	14.5	5.1	0.051

EMG: electromyogram; ECG: electrocardiogram; SCR: skin conductance response.

^aConcerning the high and low intensities evoking tasks, each of them has an invalid sample data because one of the recruited stroke patients cannot complete the training session.

SCR_{num_PEAK} ($F(2, 20) = 5.69, p = 0.013$), RSP_{rate_MEAN} ($F(2, 20) = 4.21, p = 0.036$), and RSP_{amp_MMERR} ($F(2, 20) = 3.51, p = 0.049$). Table 2 gives the post hoc multiple comparisons for physiological features with significant differences. It can be seen that significant differences between tasks that evoke anxiety with high and medium intensities are found for EMG_{cor_INTG}, EMG_{cor_MPF}, EMG_{zyg_INTG}, ECG_{hr_rate}, SCR_{MEAN}, and RSP_{amp_MMERR}. There exist good differentiations between tasks evoking high and low intensities for EMG_{cor_MPF}, ECG_{hr_rate}, SCR_{MEAN}, SCR_{num_PEAK}, SCR_{amp1stDiff_MMERR}, and RSP_{rate_MEAN}. Concerning the tasks inducing medium and low intensities, EMG_{cor_INTG}, EMG_{zyg_INTG}, ECG_{hr_rate}, SCR_{num_PEAK}, and RSP_{rate_MEAN} cause significant differences.

Performance features differences between different tasks.

Figure 8 describes the target-hitting performance with a representative stroke subject during different task periods. It is clear to see that the best trajectory tracking performance is achieved in the low anxiety-evoking task, while the worst one is done in the high anxiety-evoking task. Compared to the performance in the task-evoking anxiety with medium intensity, the low intensity evoking task shows less trajectory tracking error, but intentionally deviating from the reference trajectory and frequently exceeding the predefined target position are also demonstrated which indicates that the skill of the stroke subject is increased and repetitive movements with the same difficulty level create boredom in the subjects. Further statistical analyses of performance features found significant differences among three different tasks for PeRF_{trjerr_MEAN} ($F(2, 20) = 4.51, p = 0.023$), PeRF_{trjerr_STD}

($F(2, 20) = 3.46, p = 0.051$), PeRF_{time_MEAN} ($F(2, 20) = 6.99, p = 0.005$), and PeRF_{num_SUC} ($F(2, 20) = 5.78, p = 0.011$). The post hoc multi-comparisons for performance features differences are described in Table 3. It is obvious that both PeRF_{num_SUC} and PeRF_{trjerr_MEAN} have good differentiations among three different tasks, while PeRF_{time_MEAN} and PeRF_{trjerr_STD} have no significant differences between the medium anxiety-evoking task and the low anxiety-evoking task.

Correlation among physiological, performance features and self-reported questionnaires. Spearman correlation coefficients (R -value) were computed for those physiological and performance features which have no significant differences among three different tasks but have significant responses at least in one of the three tasks compared to baseline period. These features with significant correlations were remained and merged into the feature sets with significant differences. Table 4 details the results of the Spearman correlation analysis. EMG_{zyg_MAVP} was found to be significantly correlated with the self-reported three anxiety intensities, while ECG_{hrv_LF/HF} has significant correlations with the subject's self-reported medium and low intensities, respectively. Both of them could be used to classify anxiety with different intensities.

Offline anxiety modeling with high, medium, and low intensities.

Concerning the SVM-based anxiety intensity classification, a total of 198 samples (66 samples for each task evoking different anxiety intensities) were collected from all the recruited stroke subjects except one who were not able to complete the task. As can be seen from Tables 2 and 3, a total of 10 features (6 physiological features and 4

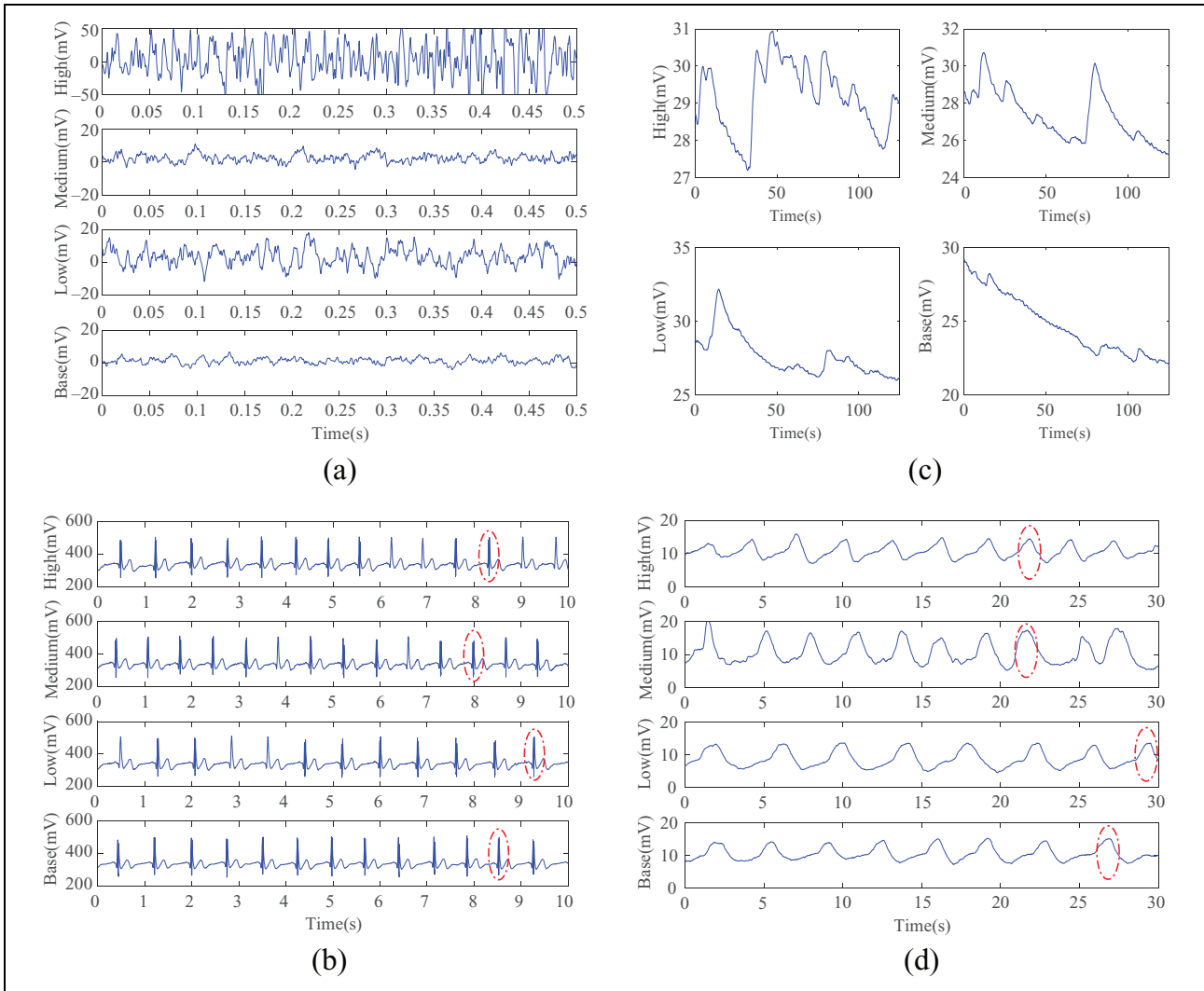


Figure 7. Physiological signals changes with a representative stroke subject during baseline and tasks evoking anxiety with high, medium, and low intensities. (a) EMG signals from corrugator supercilii, (b) ECG signals, (c) SCR signals, and (d) RSP signals. EMG: electromyogram; ECG: electrocardiogram; SCR: skin conductance response.

performance features) were found to have significant differentiations in both high/medium and high/low intensity classifiers, and 9 features (7 physiological features and 2 performance features) were statistically chosen for classifying medium and low intensities. Subject-independent anxiety intensity classifiers were developed where the features from all the subjects were merged and normalized. Discriminant performances were validated using leave-one-out cross-validation method, where a single observation taken from the collected samples was used as the test data while the others were used for training classifiers. This was done as many times as there were samples in each task. The accuracy was judged according to how often the classified intensity matched the subject's self-reported intensity ratings. The Gaussian kernel parameter γ and the soft margin parameter C were set 1.41 and 11, respectively. Tables 5 to 7 show the confusion matrices for classifiers with the

physiological, performance, and the combined physiological and performance features input.

As can be seen from these confusion matrices, the average correct rates of the SVM-based anxiety intensity classifiers with the physiological, performance, and the combined physiological and performance features were 73.2, 79.3, and 83.3%, respectively. Combining the physiological and performance features yields the most accurate classification. Although the average correct rate of the physiology-based classifier is inferior to the one of the performance-based classifier, the low intensity classification accuracy of the former (78.8%) is superior to the one in the latter (75.7%). Besides, better accuracy of medium intensity classification (78.8%) is achieved in the performance-based classifiers, which cannot well differentiated (60.6%) in the physiology-based recognizer. Fortunately, the combination of physiological and performance features yields a relative high accuracy for all of the high (87.9%), medium (78.8%), and low (83.3%) intensities.

Table 2. Post hoc multi-comparisons for physiological features difference.

	EMG _{cor_INTG}			EMG _{cor_MPF}			EMG _{zyg_INTG}		
	Mean error	Std	Significant differences	Mean error	Std	Significant differences	Mean error	Std	Significant differences
Hard/medium	3.05	1.09	**	4.97	2.40	**	−32.2	18.6	*
Medium/easy	−3.24	1.27	*	−2.39	1.23		31.3	13.1	*
Hard/easy	0.20	0.61		2.58	1.65	*	−0.9	5.85	
	ECG _{hr_rate}			SCR _{MEAN}			SCR _{amp1stDiff_MMERR}		
	Mean error	Std	Significant differences	Mean error	Std	Significant differences	Mean error	Std	Significant differences
Hard/medium	4.75	5.85	*	6.10	16.2	*	26.5	87.5	
Medium/easy	−4.94	2.13	*	4.61	14.1		5.4	11.2	
Hard/easy	−0.21	7.21		10.7	14.4	*	31.9	29.3	*
	SCR _{num_PEAK}			RSP _{rate_MEAN}			RSP _{amp_MMER}		
	Mean error	Std	Significant differences	Mean error	Std	Significant differences	Mean error	Std	Significant differences
Hard/medium	11.1	9.3		−1.18	1.07		2.71	1.68	**
Medium/easy	6.1	5.2	*	−3.72	1.23	*	4.72	6.37	
Hard/easy	17.1	12.6	*	−4.91	1.89	*	7.42	9.54	

EMG: electromyogram; ECG: electrocardiogram; SCR: skin conductance response; Std: standard deviation.

* $p < 0.05$.

** $p < 0.01$.

Phase II: Online anxiety-based robot-aided rehabilitation training task adaptation

Real-time predictive accuracy of anxiety model with three intensities. The offline anxiety intensities classifiers with three types of input data were further demonstrated to see whether these classifiers could be applied to online recognize the anxiety intensities of the stroke subjects. The accuracy of how closely the real-time anxiety intensity classifications from the SVM-based recognizers matched with the subjective anxiety intensity ratings made by the stroke subject was verified. Figure 9 demonstrates the real-time predictive accuracy for each stroke subject across 12 verification epochs.

Using classifiers with three types of input as a within-subject factor, the real-time predictive differences among physiology-based, performance-based, and physiology/performance-based classifiers were statistically analyzed through one-way repeated measures ANOVA. Results indicate that significant predictive differences are existed among classifiers with three types of input ($F(2, 20) = 22.79, p < 0.001$). The Bonferroni post hoc multiple comparisons further show that the classifier with the combined physiological and performance features (84.8 ± 5.05) yields more accurate classification than the physiology-based ($68.2 \pm 6.26, p = 0.000$) and the performance-based classifiers ($75 \pm 6.43, p = 0.004$), respectively, while there are no significant predict differences ($p = 0.092$) between the physiology-based and the performance-based intensity

classifiers. According to the experimental protocols in Section Experimental procedure of Phase II, the anxiety model with the combined physiology and performance features as inputs would be adopted in the subsequent anxiety-based robot-aided training task adaptation.

Robot-aided training task adaptation (anxiety versus performance-based). Anxiety-based robot-aided training task adaptation was implemented using the anxiety model with the combined physiological and performance features as inputs. As a comparable method, the performance-based robot-aided training task adaptation was also investigated. Figures 10 and 11 show the details of anxiety and performance-based robot-aided training task adaptation. Two important aspects (task difficulty selection and its consistency with the active willingness of the stroke patients) were demonstrated and summarized as follows.

The stroke subjects perceive more challenging in the anxiety-based robot-aided training task adaptation. Figures 10(a) and 11(a) describe the percentages of hard, medium, and easy tasks that were selected across all subjects in anxiety and performance-based robotic behavior adaptation. Averaged across all stroke subjects in the anxiety-based robotic behavior adaptation, the hard, medium, and easy tasks were chosen to be 11.8, 75.5, and 12.7% of the training epochs, respectively, while 14.1, 68.6, and 17.7% of the training epochs were made in the performance-based robot-aided training task adaptation. Further, pairwise

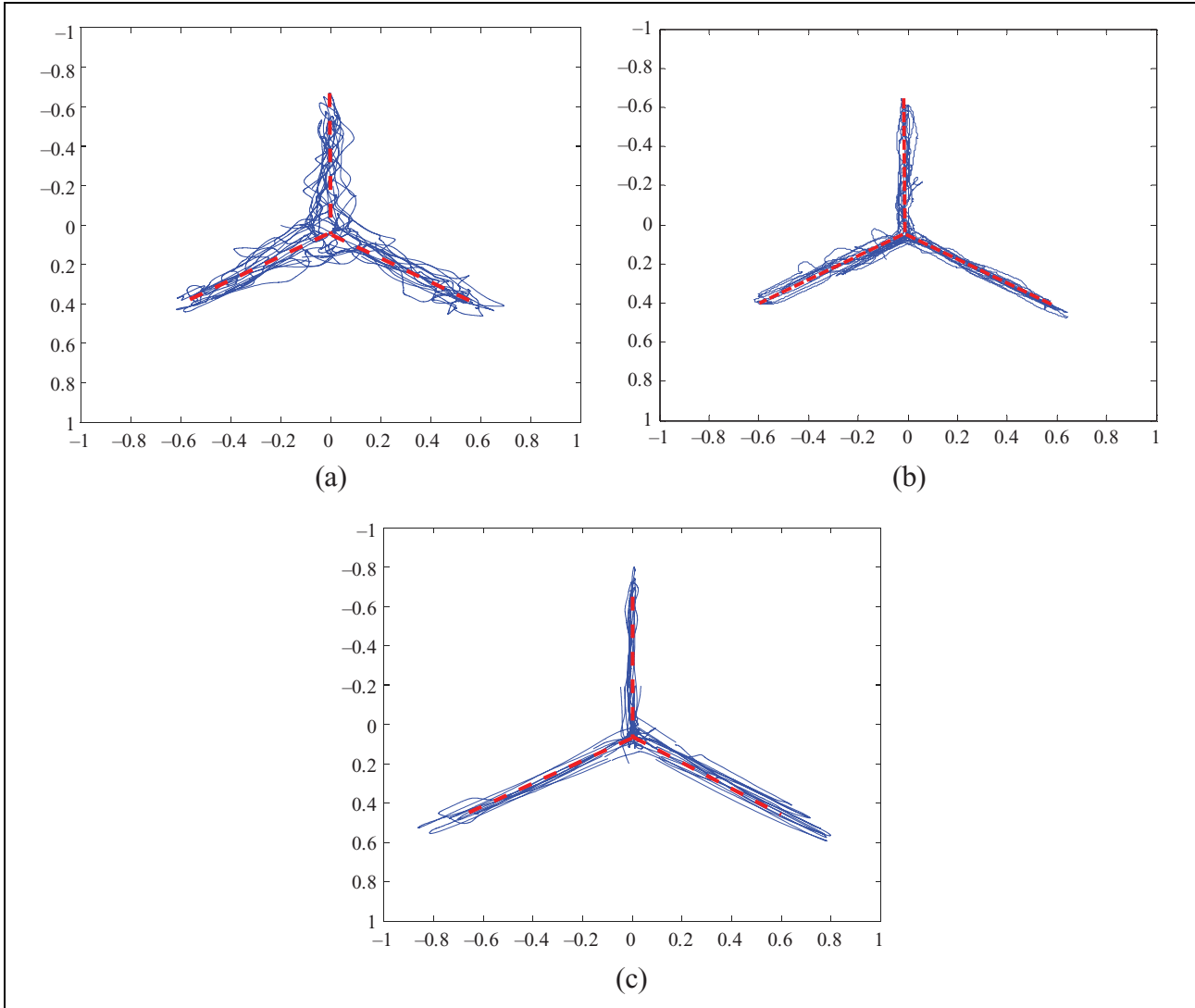


Figure 8. Target-hitting performance changes with a representative stroke subject among tasks evoking anxiety with high, medium, and low intensities (the red and blue lines denote the reference and actual trajectories, respectively). (a) High intensity evoking task, (b) medium intensity evoking task, and (c) low intensity evoking task.

t-test comparisons of the medium task with challenge difficulty between the anxiety and performance-based robotic behavior adaptation were made. Statistical results show that the anxiety-based robotic behavior adaptation ($75.5\% \pm 7.57\%$, $p = 0.016$) has more selections of medium task with challenge difficulty than the performance-based robotic behavior adaptation ($68.6\% \pm 5.05\%$). This indicates that the stroke subjects perceive more challenging in the anxiety-based robot-aided training task adaptation.

The robotic selections have more consistency with the active willingness of the stroke patients in the anxiety-based robot-aided training task adaptation. In order to demonstrate how consistently the robotic task selections matched with the active willingness of the stroke patients, the matching proportions across all stroke subjects between the robotic

selections and the active preferences of the stroke patients were calculated (shown in Figures 10(b) and 11(b)). In the anxiety-based robotic task selections, the averaged matching percentages across all the stroke patients for the three tasks, that is, hard task, medium task, and easy task, were 76.4, 9.9, 83.6, and 6.5%, respectively, while 66.8, 10.2, 80.7, and 9.1% were matched in the performance-based robotic task selections. Pairwise *t-test* comparisons of the whole matching for three tasks show that the anxiety-based robotic selections ($76.4\% \pm 6.36\%$, $p = 0.005$) have higher matching percentages than the performance-based robotic selections ($66.8\% \pm 5.13\%$). Further, κ statistical results also verify that the consistency coefficients in the anxiety-based robotic task selections ($\kappa = 0.513$, $p < 0.001$) are higher than the one in the performance-based robotic task selections ($\kappa = 0.276$, $p < 0.001$). The abovementioned statistical results indicate that the robotic selections in the

Table 3. Post hoc multi-comparisons for performance features differences.

	PeRF _{time_MEAN} (s)					PeRF _{num_SUC} (%)				
	Mean	Standard deviation	Mean error	Standard error	Significant differences	Mean	Standard deviation	Mean error	Standard error	Significant differences
Hard (H)/medium	13.38 (H)	5.66 (H)	5.67	1.93	**	42.6 (H)	21.8 (H)	−30.8	−13.5	*
Medium (M)/easy	7.71 (M)	1.62 (M)	1.42	0.41		73.4 (M)	35.2 (M)	−12.9	−10.2	*
Hard/easy(E)	6.29 (E)	1.42 (E)	7.09	1.89	**	86.3 (E)	19.3 (E)	−43.7	−18.6	**
	PeRF _{trjerr_MEAN} (mm)					PeRF _{trjerr_STD} (mm)				
	Mean	Standard deviation	Mean error	Standard error	Significant differences	Mean	Standard deviation	Mean error	Standard error	Significant differences
Hard (H)/medium	1.69 (H)	1.15 (H)	1.22	0.58	**	0.88 (H)	0.45 (H)	0.41	0.13	*
Medium (M)/easy	0.47 (M)	0.25 (M)	0.25	0.14	*	0.47 (M)	0.24 (M)	−0.08	0.06	
Hard/easy (E)	0.22 (E)	0.41 (E)	1.47	1.05	*	0.55 (E)	0.38 (E)	0.33	0.18	*

PeRF: performance.

* $p < 0.05$.** $p < 0.01$.**Table 4.** Spearman correlations between self-reported rating intensities and physiological features with insignificant response as well as performance features with insignificant difference.

	Self-reported high intensity		Self-reported medium intensity		Self-reported low intensity	
	R-value	p-Value	R-value	p-Value	R-value	p-Value
EMG _{cor_STD}	−0.089	0.795	−0.405	0.192	0.019	0.958
EMG _{cor_RMS}	0.180	0.596	−0.238	0.456	−0.332	0.349
EMG _{cor1stDiff_STD}	0.071	0.083	0.519	0.084	−0.806	0.089
EMG _{zyg_MAVP}	−0.201	0.045	−0.591	0.043	0.481	0.026
EMG _{zyg1stDiff_MEAN}	0.176	0.604	0.156	0.629	0.448	0.119
ECG _{hrv_LF/HF}	0.466	0.068	0.66	0.02	0.493	0.007
ECG _{hrv_pNN50}	0.333	0.172	0.387	0.613	−0.601	0.087
SCR _{STD}	0.152	0.066	−0.11	0.748	0.791	0.112
SCR _{1stDiff_STD}	0.183	0.591	0.801	0.086	0.857	0.076
RSP _{1stDiff_MEAN}	−0.437	0.179	0.823	0.077	0.611	0.108
RSP _{var_MEAN}	0.448	0.079	0.034	0.921	0.342	0.304
PeRF _{trj_SMOOTH}	0.579	0.062	0.845	0.069	0.393	0.232

EMG: electromyogram; ECG: electrocardiogram; SCR: skin conductance response.

Table 5. Confusion matrix using physiological features as the classifiers input and the average correct rate was 73.2%.

	High intensity	Medium intensity	Low intensity	Total samples	Accuracy rate
High intensity	53	5	8	66	80.3
Medium intensity	18	40	8	66	60.6
Low intensity	10	4	52	66	78.8

Table 6. Confusion matrix using performance features as the classifiers input and the average correct rate was 79.3%.

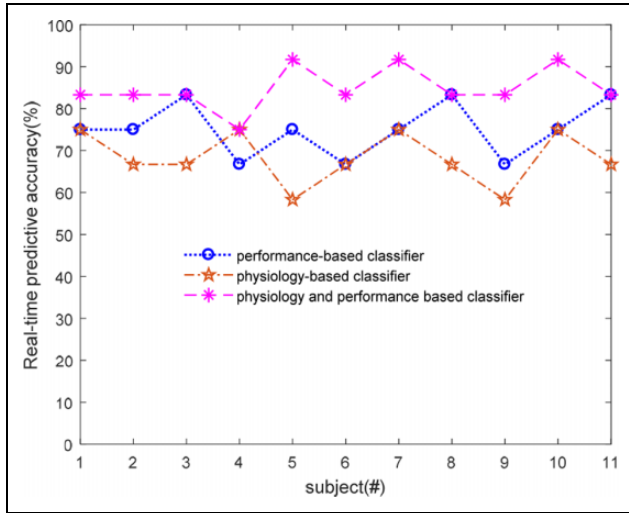
	High intensity	Medium intensity	Low intensity	Total samples	Accuracy rate
High intensity	55	5	6	66	83.3
Medium intensity	4	52	10	66	78.8
Low intensity	6	10	50	66	75.7

anxiety-based robotic behavior adaptation have more consistency with the active willingness of the stroke patients. Another finding is that the matching proportions of the

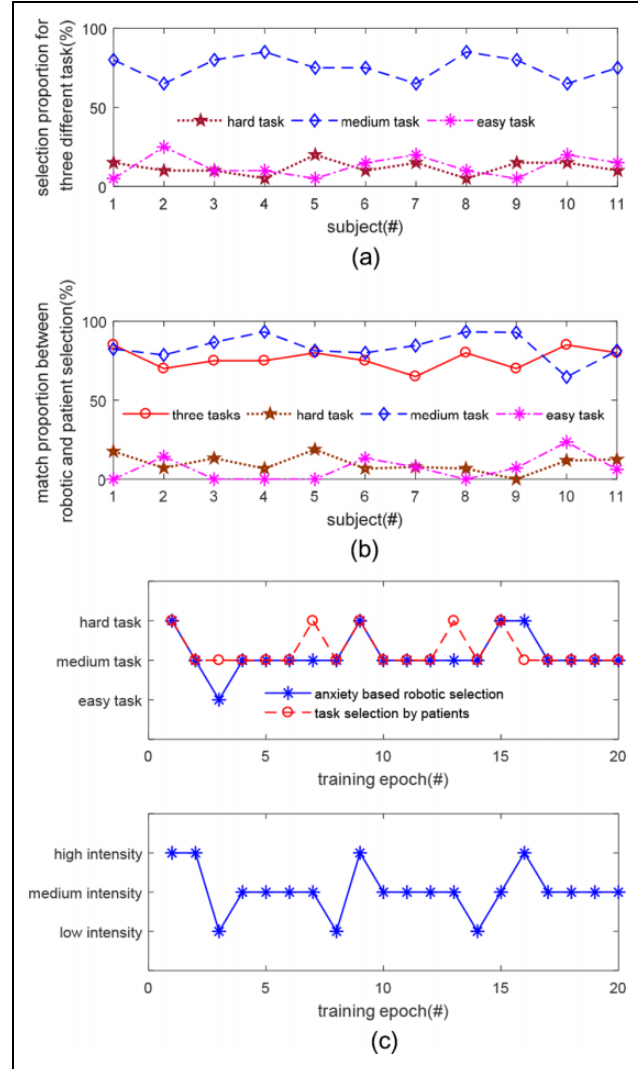
medium task with challenge difficulty are the highest among three tasks in anxiety and performance-based robotic behavior adaptations.

Table 7. Confusion matrix using the combined physiological and performance features as the classifiers input and the average correct rate was 83.3%.

	High intensity	Medium intensity	Low intensity	Total samples	Accuracy rate
High intensity	58	5	3	66	87.9
Medium intensity	5	52	9	66	78.8
Low intensity	4	7	55	66	83.3

**Figure 9.** Real-time predictive accuracy for classifiers with three types of input data across all subjects.

To under these results more clearly, robotic task selections for the same representative stroke subject in the anxiety and performance-based robot-aided training task adaptations are described in Figures 10(c) and 11(c). In the anxiety-based robotic behavior adaptation, a hard task was randomly initialized, and the high intensity of anxiety was observed from the subject. The first mismatching occurred in the third training epoch, where the rehabilitation robot still recognized high intensity of anxiety from the stroke subject at the end of the second training epoch and thus decreased the difficulty level in the third training epoch from medium to easy task. However, the stroke subject's active preference was still kept medium task. During the whole 20-epoch training epochs, the number of the robotic and the patient's selections for the hard, medium, and easy tasks were 4/15/1 and 5/15/0, respectively. The matching proportions of the hard, medium, and easy tasks were 18.7, 81.3, and 0%, respectively, and the whole matching proportion for the three tasks was 80%. In the performance-based robotic behavior adaptation, an easy task was randomly initialized and the excellent motor performance was obtained from the subject at the end of the first training epoch. According to the performance-based task adaptation mechanism, an increase of the task difficulty from the easy to medium task was determined for the subsequent second training epoch. Similar robot-aided training task adaptations were achieved in the rest of

**Figure 10.** Anxiety-based robot-aided training task adaptation. (a) Selection proportions for three different tasks (b) matching proportions between robotic and patient selections and (c) robotic behavior selection for a representative patient.

the training session. Across the whole training epochs, the number of the robotic and the patient's selections for the hard, medium, and easy tasks were 3/13/4 and 2/17/1, respectively. The averaged matching percentages for all three tasks was 70%, and the specific matching proportions for the hard, medium, and easy tasks were 7.1, 85.8, and 7.1%, respectively.

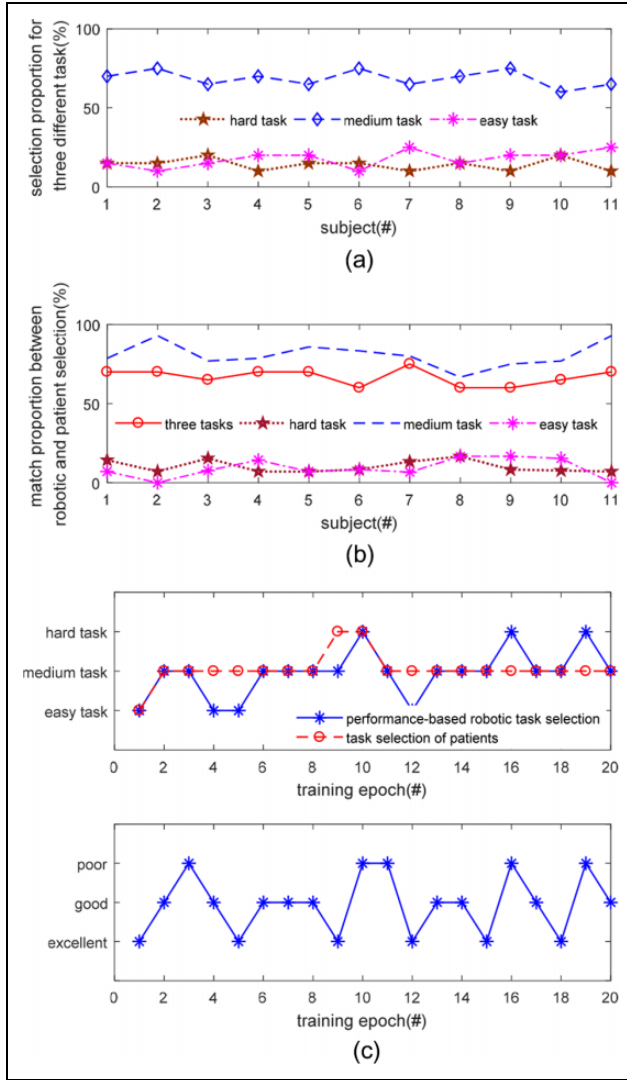


Figure 11. Performance-based robot-aided training task adaptation. (a) Selection proportions for three different tasks (b) matching proportions between robotic and patient selections and (c) robotic behavior selection for a representative patient.

Engagement investigations for the stroke patients (anxiety versus performance-based). As mentioned in Section Engagements investigations for the stroke patients, the User Engagement Scale including five measures was used to quantitatively compare the impacts of the anxiety or performance-based training task adaptation on the engagements of stroke patients. Additionally, the level of the stroke subject's liking on the task during each epoch was also statistically analyzed. Table 8 summarizes the statistical results for the perceived engagements and task liking levels in anxiety and performance-based robotic behavior adaptation. Only three of the engagement measures (engagingness, willingness to continue the exercise and focused attention) are internally reliable, while the other two measures (motivation and feedback) have Cronbach's α values

lower than 0.7 and were excluded from the subsequent pairwise *t*-test comparisons. In the anxiety and performance-based robotic behavior adaptations, the motivation measure was quantified as 3.75 ± 0.96 and 3.06 ± 0.47 , respectively, while the feedback measure was evaluated as 3.53 ± 0.49 and 3.25 ± 0.86 , respectively.

Pairwise *t*-test comparisons of the internally reliable engagement measures and task liking indicate that the anxiety-based robot-aided training task adaptation has more engagingness ($p = 0.033$), focused attention ($p = 0.032$), and task liking ($p = 0.006$) than the one in the performance-based robot-aided training task adaptation, while the willingness to continue the exercise ($p = 0.167$) has no significant difference. The duration (about 1 h) of the whole training session consisting of 20 epochs may be one of the possible reasons. This is especially true for the stroke patients. The engagement measures and task liking levels averaged across all the subjects and for each of the subjects are graphically shown in Figure 12(a) and (b), respectively.

Combined with the results in "Robot-aided training task adaptation (anxiety versus performance-based)" section, it can be seen that the training session with more challenging and more consistency with the active willingness of the stroke patients has higher engagements and task liking level. Of 11 subjects, the engagement measures and task liking levels of the second and the seventh subjects in anxiety-based robotic behavior adaptation were inferior to the one in the performance-based robotic behavior adaptation. These are also demonstrated in Figures 8 and 9. The averaged matching proportions of the seventh subject are higher in the performance (75%) than the anxiety-based (65%) robotic behavior adaptation. Although the same matching proportion (70%) of the second subject exists in anxiety and performance-based robotic behavior adaptation, the former matching proportions for the medium task with challenge difficulty (78.6%) are far lower than the latter one (92.9%).

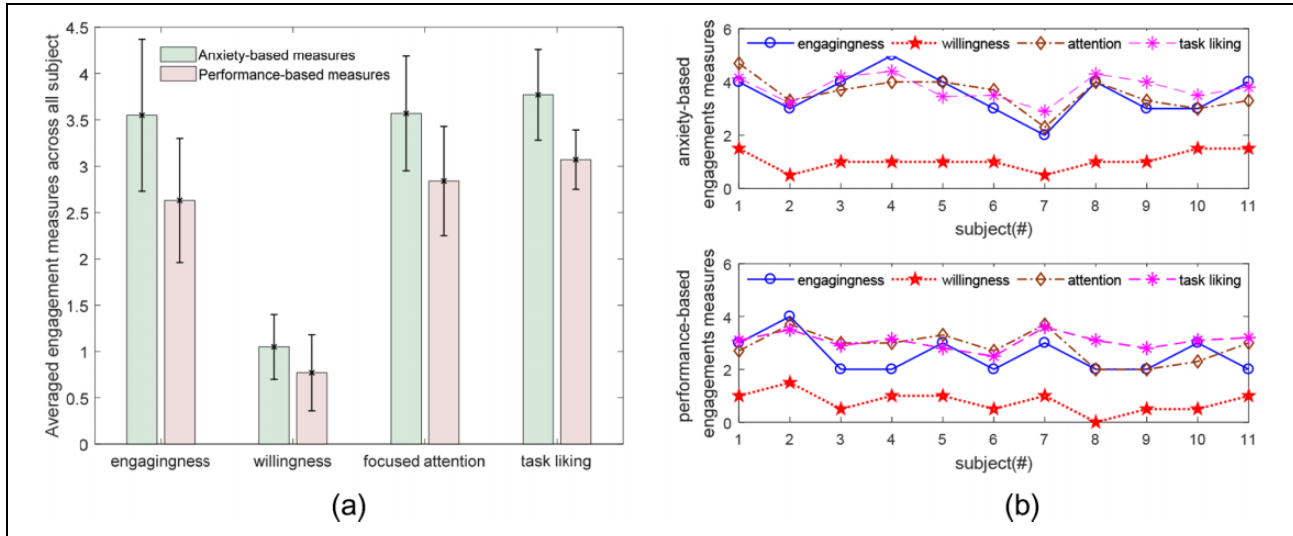
Discussions

Recognizing the stroke survivor's emotion, achieving emotion-based rehabilitation robotic behavior adaptation, and exploring the stroke patient's interaction engagements, as a new type of active involvement in rehabilitation training for stroke patients, receive little attention in the past few years. Most of the research are mainly focused on recognizing the active movement intention of the stroke subjects from the viewpoint of biomechanical (force, position, velocity, etc.) and bioelectrical (electroencephalogram and electromyography) signals. One of the key challenges is to collect and analyze emotion-related information data from the stroke patients.

Compared with the emotion recognition for healthy subjects, there exist more challenges. On the one hand, less emotion-based information can be adopted. The patients

Table 8. Statistical results for the engagement measures and task liking levels in anxiety-based (anxiety) and performance-based (performance) robotic behavior adaptation.

Measures	Anxiety	Performance	df	t-Value	p-Value
Engagingness	3.55 ± 0.82	2.63 ± 0.67	10	2.469	0.033
Willingness to continue	1.05 ± 0.35	0.77 ± 0.41	10	1.491	0.167
Focused attention	3.57 ± 0.62	2.84 ± 0.59	10	2.486	0.032
Task liking	3.77 ± 0.49	3.07 ± 0.32	10	3.492	0.006

**Figure 12.** The engagement measures and task liking averaged across all the subjects and for each of the subjects. (a) The engagement measures and task liking levels averaged across all subjects and (b) the engagement measures and task liking for each of the subjects.

with chronic stroke are often characterized by dull facial expression, severe aphasia, and limb motor dysfunction. These vulnerabilities place limits on observational (facial expressions), conversational (vocal intonations), and limb (body gestures) methodologies to recognize the stroke survivors' emotional states. On the other hand, although the physiology-based emotion modeling is more suitable for stroke survivors' characteristics, good classification accuracy cannot be achieved only using the physiology features. This is also indicated in the current research. Classifiers development is one possible reason, but another key issue is that it is difficult to obtain physiological features with significant responses and differences, which is subject to many factors from the stroke survivors themselves such as the damaged neural systems, the different recovery stages, and the leading causes of stroke. Fortunately, according to the current research findings, the classification accuracy achieves a substantial improvement by combining the physiological features with the motor performance features.

Emotion-based robotic behavior adaptation is an important and interesting issue in robot-aided neurorehabilitation. Apart from making comparisons with healthy subjects, the next step we would like to carry out is to make deep comparisons between the healthy and hemiplegia sides of the stroke patients as well as stroke patients with different recovery stages. Additionally, compared to the

performance-based robotic behavior adaptation, the preliminary results in the current research show that anxiety-based robot-aided training task adaptation has more challenging task selections, more consistency with the willingness of the stroke patients, higher engagements, and task liking levels. We are now conducting clinical experimental study, and the rehabilitation outcomes of anxiety and performance-based robotic behavior adaptation will be investigated.

Conclusions

A pilot study on anxiety-based human–robot interaction in robot-aided stroke rehabilitation was carried out in the current research. The main focuses were divided into two consecutive phases. In phase I, the motor performances and the physiological responses to the stroke subject's anxiety with high, medium, and low intensities were first statistically analyzed, and then anxiety models with three intensities were offline developed using SVM-based machine learning method. In phase II, anxiety-based closed-loop robot-aided training behavior adaptation and its impacts on patient–robot interaction engagements were explored. As a comparison, a performance-based robot-aided training task adaptation was implemented. The User Engagement Scale including five measures and the task liking level were used

to quantify the engagements of the whole training session and the task satisfaction in each training epoch, respectively. Experimental results with 12 recruited stroke patients conducted on the Barrett WAMTM manipulator verified that the rehabilitation robot can implicitly recognize the anxiety intensities of the stroke survivors, and the anxiety-based real-time robotic behavior adaptation shows more engagements and task satisfactions in the human–robot interactions.


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 jointly supported by the National Natural Science Foundation of China (61305095, 61673114, 61603195, and 61773078), the Key Research and Development Program of Jiangsu Province (BE2015701), the Natural Science Foundation of Jiangsu Province of China (BK20141426, BK20140878, and BK20170898), the Qing Lan Project of Jiangsu Province of China (QL00516014), the Informatization Project of Jiangsu Provincial Commission of Health and Family Planning (X201603) and the Jiangsu Overseas Research and Training Program for University Prominent Young and Middle-aged Teachers and Presidents.

ORCID iD

Guozheng Xu  <http://orcid.org/0000-0001-9973-8408>

Lizheng Pan  <http://orcid.org/0000-0002-3698-0122>

References

- Mazzoleni S, Turchetti G, Palla I, et al. Acceptability of robotic technology in neuro-rehabilitation: preliminary results on chronic stroke patients. *Comput Method Program Biomed* 2014; 116: 116–122.
- Duschau-Wicke A, von Zitzewitz J, Riener R, et al. Path control: a method for patient-cooperative robot-aided gait rehabilitation. *IEEE Trans Neural Syst Rehabil Eng* 2010; 118: 38–48.
- Hussain S, Jamwal PK, Ghayesh MH, et al. Assist-as-needed control of an intrinsically compliant robotic gait training orthosis. *IEEE Trans Ind Elect* 2017; 64: 1675–1685.
- Wakita K, Huang J, Di P, et al. Human-walking-intention-based motion control of an omnidirectional-type cane robot. *IEEE/ASME Trans Mech* 2013; 18: 285–296.
- Xu G, Song A, and Li H. Control system design for an upper-limb rehabilitation robot. *Adv Robot* 2011; 25: 229–251.
- Lenzi T, De Rossi SMM, Vitiello N, et al. Intention-based EMG control for powered exoskeletons. *IEEE Trans Biom Eng* 2012; 59: 2180–2190.
- Cui C, Bian GB, Hou ZG, et al. A multimodal framework based on integration of cortical and muscular activities for decoding human intentions about lower limb motions. *IEEE Trans Biom Cir Sys* 2017; 11: 889–899.
- Lohmann S, Strobl R, Mueller M, et al. Psychosocial factors associated with the effects of physiotherapy in the acute hospital. *Disabil Rehabil* 2011; 33: 2311–2321.
- Soleymani M, Asghari-Esfeden S, Fu Y, et al. Analysis of EEG signals and facial expressions for continuous emotion detection. *IEEE Trans Affect Comput* 2016; 7: 17–28.
- El Ayadi M, Kamel MS, and Karray F. Survey on speech emotion recognition: features, classification schemes, and databases. *Pattern Recogn* 2011; 44: 572–587.
- Kleinsmith A and Bianchi-Berthouze N. Affective body expression perception and recognition: a survey. *IEEE Trans Affect Comput* 2013; 4: 15–33.
- Koelstra S, Muhl C, Soleymani M, et al. Deap: a database for emotion analysis; using physiological signals. *IEEE Trans Affect Comput* 2012; 3: 18–31.
- Kreibitz SD. Autonomic nervous system activity in emotion: a review. *Biol Psychol* 2010; 84: 394–421.
- Khosrowabadi R, Quek C, Ang KK, et al. ERNN: a biologically inspired feedforward neural network to discriminate emotion from EEG signal. *IEEE Trans Neural Netw Learn Syst* 2014; 25: 609–620.
- Hughes AM, Freeman CT, Burrage JH, et al. Shoulder and elbow muscle activity during fully supported trajectory tracking in people who have had a stroke. *J Elect Kines* 2010; 20: 465–476.
- Xu G, Song A, Pan L, et al. Clinical experimental research on adaptive robot-aided therapy control methods for upper-limb rehabilitation. *Robotica* 2014; 32: 1081–1100.
- Xu G, Gao X, Chen S, et al. A novel approach for robot-assisted upper-limb rehabilitation: progressive resistance training as a paradigm. *Int J Adv Robot Syst* 2017; 14: 1729881417736670.
- Kuriakose S and Lahiri U. Understanding the psychophysiological implications of interaction with a virtual reality-based system in adolescents with autism: a feasibility study. *IEEE Trans Neural Syst Rehabil Eng* 2015; 23(4): 665–675.
- Kim J and André E. Emotion recognition based on physiological changes in music listening. *IEEE Trans Pattern Anal Mach Int* 2008; 30: 2067–2083.
- Liu C, Agrawal P, Sarkar N, et al. Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. *Int J Hum Comput Int* 2009; 25: 506–529.
- Liu C, Conn K, Sarkar N, et al. Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE Trans Robot* 2008; 24: 883–896.
- Chanel G, Rebetz C, Bétrancourt M, et al. Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Trans Syst Man Cybern A Syst Hum* 2011; 41: 1052–1063.
- Cai H and Lin Y. Modeling of operators' emotion and task performance in a virtual driving environment. *Int J Hum Comput St* 2011; 69: 571–586.
- Chang JJ, Tung WL, Wu WL, et al. Effects of robot-aided bilateral force-induced isokinetic arm training combined with conventional rehabilitation on arm motor function in patients with chronic stroke. *Arch Phys Med Rehabil* 2007; 88: 1332–1338.

25. Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Disc* 1998; 2: 121–167.
26. Vapnik VN and Vapnik V. *Statistical learning theory*. New York: Wiley, 1998.
27. Shirzad N and Van der Loos HFM. Evaluating the user experience of exercising reaching motions with a robot that predicts desired movement difficulty. *J Motor Behav* 2016; 48: 31–46.

Appendix I

Given the linearly non-separable training data pairs (x_i, y_i) ($i = 1, 2, \dots, n$), where x_i represents input data sample in p -dimensional space, $y_i \in \{+1, -1\}$ is the class label, and n is the number of training data samples, a nonlinear mapping \varnothing was conceptually needed to map the input data into a higher m -dimensional ($m > n$) feature space in which the mapped data samples could be approximately linearly classified. The classification could then be described as finding the following separating hyperplane f in higher feature space

$$f(x_i) = w^T \varnothing(x_i) + b \quad (1A)$$

where $w = [w_1, w_2, \dots, w_m]$ is the weight vector, $\varnothing = [\psi_1, \psi_2, \dots, \psi_m]$ is the nonlinear mapping vector, and b is the hyperplane displacement.

Unfortunately, the abovementioned mapping solution might give rise to curse of dimensionality. To solve this issue, kernel functions κ were introduced satisfying

$$\kappa(x_i, x) = \langle \varnothing(x_i), \varnothing(x) \rangle \quad (1B)$$

where x_i and x are the data samples in p -dimensional input space, respectively. RBF kernel function is adopted here and described as

$$\kappa(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \quad (1C)$$

where γ is the Gaussian kernel parameter.

Further, the problem of finding the optimal hyperplane could be transformed as the following convex quadratic optimization

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & \left\{ \begin{array}{l} y_i \left(\sum_{j=1}^m w_j \psi_j(x_i) + b \right) \geq 1 - \xi_i \\ \xi_i \geq 0, \quad i = 1, 2, \dots, n \end{array} \right\} \end{aligned} \quad (1D)$$

where C is the soft margin parameter and ξ_i is the slack variable.

By introducing the Lagrange multipliers α_i , the convex quadratic optimization described by Equation (1D) could be further reduced to the following dual quadratic programming problem

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \kappa(x_i, x_j) \\ \text{s.t.} \quad & \left\{ \begin{array}{l} 0 \leq \alpha_i \leq C \\ \sum_{i=1}^n \alpha_i y_i = 0 \quad i = 1, 2, \dots, n \end{array} \right\} \end{aligned} \quad (1E)$$