Implementation and Validation of Engagement Monitoring in an Engagement Enhancing Rehabilitation System

Chong Li, *Student Member, IEEE*, Zoltán Rusák, Imre Horváth, Adrie Kooijman, and Linhong Ji

Abstract—Enhancing engagement of patients during stroke rehabilitation exercises are in the focus of current research. Various methods and computer supported tools have been developed for this purpose, which try to avoid mundane exercising that is prone to become a routine or even boring for the patients and leads to ineffective training. This paper introduces an engagement enhancing cyber-physical stroke rehabilitation system (CP-SRS) aiming at enhancing the patient's engagement during rehabilitation training exercises. This paper focuses on introducing the implementation and validation of the engagement monitoring subsystem (EMS) in the CP-SRS. The EMS is expected to evaluate the patient's actual engagement levels in motor, perceptive, cognitive and emotional aspects. Experiments in these four aspects were conducted separately, in order to characterize the range and accuracy of the engagement indicators by influencing the subjects into different engaged states. During the experiments, different setups were created to mimic the situations in which the subject was engaged or not engaged. The subjects involved in the experiments were healthy subjects. Results showed that the measurement in motor, perceptive, cognitive, and emotional aspects can represent the corresponding engagement level. More experiments will be conducted in the future to validate the efficiency of the CP-SRS in enhancing the engagement with stroke patients.

Index Terms—Cyber-physical system, engagement monitoring, stroke rehabilitation, system implementation, system validation.

I. Introduction

MAINTAINING and enhancing patient's engagement in stroke rehabilitation exercises are in the focus of current research. It has been demonstrated through evidence that it can remarkably improve the functional outcome of technology-assisted stroke rehabilitation [1]–[3]. There have been various

methods and computer supported tools developed for trying to avoid mundane exercising that is prone to become a routine or even boring for the patients and leads to ineffective training. The most common method is integrating rehabilitation training with gamification. Many researchers have studied user engagement with technology, especially with video games. O'Brien and Toms identified four theories have been utilized to explain engagement, including flow theory [4], aesthetic theory [5], play theory [6], and information interaction [7]; they also concluded the attributes of engagement in technology applications, namely, aesthetics, affective appeal, attention, challenge, feedback, goal-directed. meaningfulness, motivation, perceived control, and sensory appeal [8]. Engagement is a complex construct which also depends on the context. In the context of rehabilitation exercises, engagement has been defined as a construct that is driven by motivation and executed through active and effortful participation [9]. Lequerica and Kortte proposed the components influencing therapeutic engagement includes 1) perceived need of treatment, 2) perceived likelihood of a successful outcome, 3) perceived self-efficacy to complete the tasks, and 4) reassessment of beliefs, attitudes, and expectations [9]. Kortte et al. developed an engagement rating scale, which took rehabilitation engagement, therapy absences, functional status, emotional functioning, affective state, level of functioning and denial into consideration [10].

A major limitation is that the current methods developed to evaluate and measure engagement are subjective and qualitative. The same patient's engagement could be different according to different therapists, which leads to inaccurate assessment. In addition, without precise measurement, the methods to engage the user cannot be validated. The trend in rehabilitation robotics is to develop robotic systems which can give direct quantitative feedback to the patient about her/his performance during the therapy, thus enhancing motivation and self-appraisal of the value of the proposed exercises [11]. In education, several studies evaluated student's engagement using postures [12]–[13], body motion [13] or log files in e-learning system [14]. Szafir and Mutlu developed an adaptive agent that monitors the cognitive engagement of students and improves their engagement during learning [15]. They use EEG to quantitatively represent the engagement level of students. The adaptive agent is implemented as robotic instructor capable to give immediacy cues triggered by drops in EEG-monitored engagement levels to regain participant attention. Although this method is very promising in education, in the
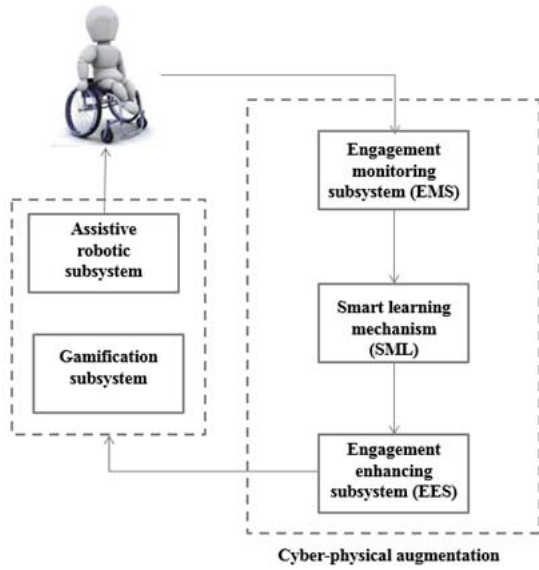
Fig. 1. System concept of the cyber-physical stroke rehabilitation system (CP-SRS).

Fig. 2. System architecture.

context of rehabilitation training integrated with video games, a more comprehensive characterization of the engagement that monitors not only the cognitive engagement of patients, but also the motor, emotional and perceptive engagement according to the influencing factors for therapeutic engagement.

Several studies applied serious gaming [16], collaborative tele-rehabilitation [17] and virtual reality [18] to increase patient's engagement during the training exercises. Another limitation is that none of these studies applied intervention during training exercises in order to maintain the user's engagement. Even the best engagement methods such as serious games need to cope with decrease of engagement as the users get familiar with the game [19]. The ultimate goal of the proposed CP-SRS is to enhance and maintain the engagement by continuous monitoring of the engagement level and introducing interventions during rehabilitation exercises [20].

This paper aims to validate if the EMS embedded in the CP-SRS can capture the actual engagement level of the users. The next section first introduces the concept, functions, architecture, and operation flow of the whole system, CP-SRS. The third section introduces the implementation of the engagement monitoring subsystem (EMS). Finally, experiments were designed and conducted. In the experiments, different setups were created to mimic the situations in which the subject was in engaged, unengaged, or neutral engagement state. The aim of the experiments was to test if the proposed indicators can reflect the expected differences in the engagement level in different setups.

II. SYSTEM CONCEPT

A. Objectives

The overall objective of our background research is development and validation of a cyber-physical stroke rehabilitation system (CP-SRS) which is composed of five subsystems, namely, assistive robotic subsystem, gamification subsystem, engagement monitoring subsystem and smart learning mechanism (SLM), and engagement enhancement
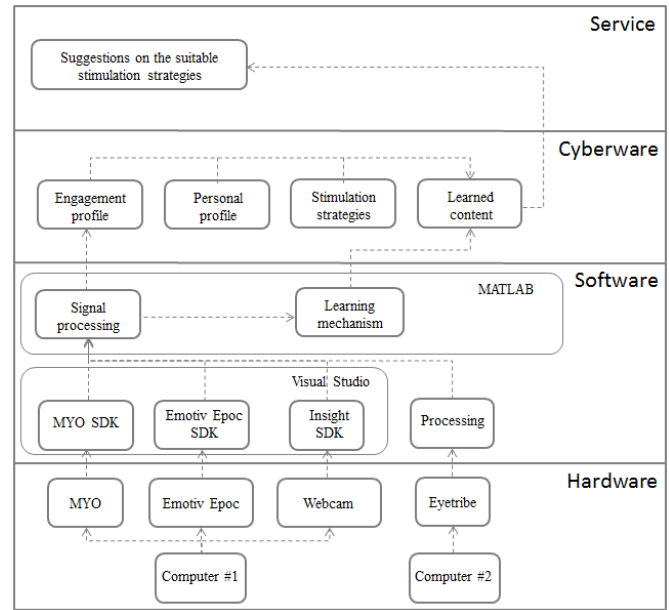
subsystem (EES) (Fig. 1). Specifically, the assistive robotic subsystem assists the stroke survivors to make exercises during rehabilitation program in order to reduce deficits in their motor ability. Gamification subsystem integrates video games with the training exercises and enables human computer interaction [21].

The main function of the cyber-physical augmentation in the system is to enhance the patient's engagement by introducing interventions during rehabilitation exercises. To determine on when to introduce the interventions, the EMS should monitor patient's engagement level. Basically, when the patient's engagement level decreases, the system introduces interventions. The interventions should re-engage the patients and maintain the high engagement of the patient. Based on our previous research [22], stimulations in motor, perceptive, cognitive, and emotional aspects are able to re-engage the patient. The EES apply the stimulations by adjusting the parameters of the training exercises. There are several stimulations in each aspect. In order to determine which stimulation or which combination of stimulations are the most suitable ones to apply, the SLM can make suggestions on the stimulations to apply based on the knowledge it has learned in previous cases [20].

In Sections II-B and II-C the architecture and operation flow of the system are introduced. This paper focuses on the implementation and validation of the EMS, which is presented in the Section III and Section IV.

B. Architecture

The architecture of this cyber physical augmentation in Fig. 2 is represented in four levels of hardware, software, cyberware, and service components. The first level includes commercialized hardware components capable to monitor the activities and physiological properties of the patients. Due to the computation capacity of a regular computer, two computers were used for implementing the prototype. MYO, Eyetribe,
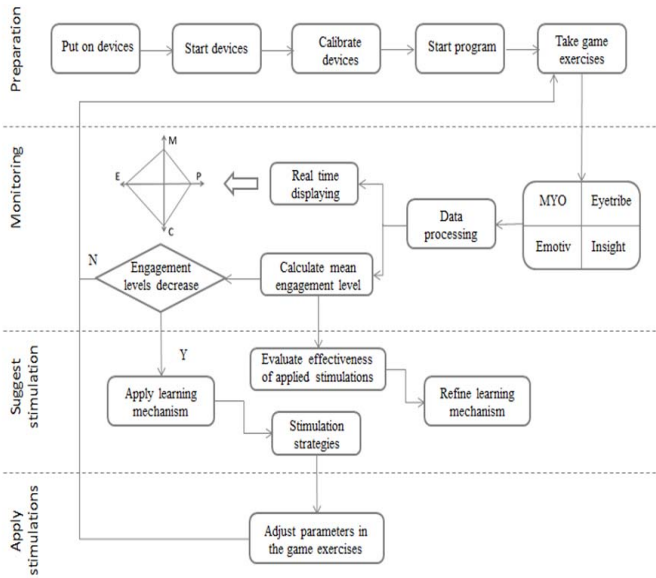
Fig. 3. Operation flow.

Emotiv Epoc, and a webcam were used to monitor the patient's muscle activities, eye movement, brain activities, and facial expressions of emotion respectively. The second level, which includes third party software SDK provided with hardware components convert the measured signals into a time stamped data stream. Then the data is streamed from Visual Studio and Processing to MATLAB via TCP/IP. Then MATLAB processes the data and interprets them as the actual engagement levels. Based on analyzing the current engagement level, the learning mechanism makes suggestions on the most suitable stimulation strategies to apply when the engagement has decreased. In the cyberware level, it consists of the engagement profile, personal profile, stimulation strategies, and the learned content in the leaning mechanism. In the fourth level, the cyber physical augmentation of the CP-SRS is able to make suggestions on the suitable stimulation strategies to apply based on the learned content. This augmentation can be easily integrated with the rehabilitation robot by applying the stimulation strategies by the means of adjusting the parameters in the robot assisted training, such as the assisting force from the robotic arm, threshold of the force sensor on the user interface, size of the moving space and so on. Then the robot system assists the patient in the adjusted exercise.

C. Operation Flow

There are four phases in the operation flow of the proposed system. In preparation, the physical therapist assists the patient to put on the devices, start the devices, calibrate the devices, and start the program to monitor the patient's status. Then in the next phase the patient's engagement levels are monitored during the game exercises. If the engagement level decreases, the trained learning mechanism suggests the stimulations in the third phase. In the last phase, the stimulations are applied and the patient continues with the adjusted exercise. Additionally, after the stimulations are applied, the effectiveness of the stimulation is evaluated and used to refine the learning

mechanism to make it more accurate. Moreover, when a new patient begins to use this system, the learning mechanism can also give recommendations of the exercise to take based on the knowledge it has learned.

III. IMPLEMENTATION OF EMS

A. Overview of the EMS

As shown in the sequence diagram of the EMS (Fig. 4), the system reads data from five sources, namely, EEG signals from Emotiv Epoc, emotion by analyzing the facial expressions, EMG signals from MYO, eye movement from Eyetribe, and content changes on the screen. This data is streamed to MATLAB in parallel, and is interpreted as the mean engagement level of every 10 s. Fig. 5 shows the variables and their sources for calculating the engagement level in the four aspects. Details about processing each variable are discussed in the next part.

B. Realization of the EMS in Each Aspect

1) Motor Engagement Monitoring: Motor engagement (E_m) is defined as a state in which the patient moves with active and effortful motion. It is calculated as the normalized EMG signals. RMS is considered to be the most meaningful calculation of the amplitude of the EMG signal, since it gives a measure of the power of the signal. In our previous study, the EMG signal was normalized by having RMS of EMG divided by the average velocity; we found that when the subjects were smiling or they were focusing on the exercises, the normalized EMG were more intense than that when the subjects were less engaged [19]. Therefore, we used this measurement as the indicator of motor engagement.

$$E_m = \frac{EMG_{RMS}}{veloctiy} \quad (1)$$

MYO Gesture Control Armband[1] (Thalmic Labs) consists of eight EMG sensing modules that can be strapped onto user's forearm. These EMG signals can be interpreted as different hand gestures that can be translated into various types of computerized input via Bluetooth. There is also a combination of a gyroscope, an accelerometer, and a magnetometer to detect the arm motion.

We used MYO as a low-cost device to monitor the muscle activities on the upper arm. Visual Studio 2013 was used to read the data from MYO and stream the data to MATLAB via TCP/IP in real time. The data consisted of eight channels of EMG signals. The sample rate was 200 Hz. The EMG signals were filtered using a bandpass filter (20–500 Hz). Maximum voluntary contraction (MVC) was used to normalize the data [23]. Before the experiment, each participant was required to do isometric contractions of the bicep for three times. The maximum values of each repetition are averaged to compute the MVC. The data acquired from the experiment was compared to the MVC and were rescaled to percentage of the MVC. In MATLAB, these data were used to calculate RMS.
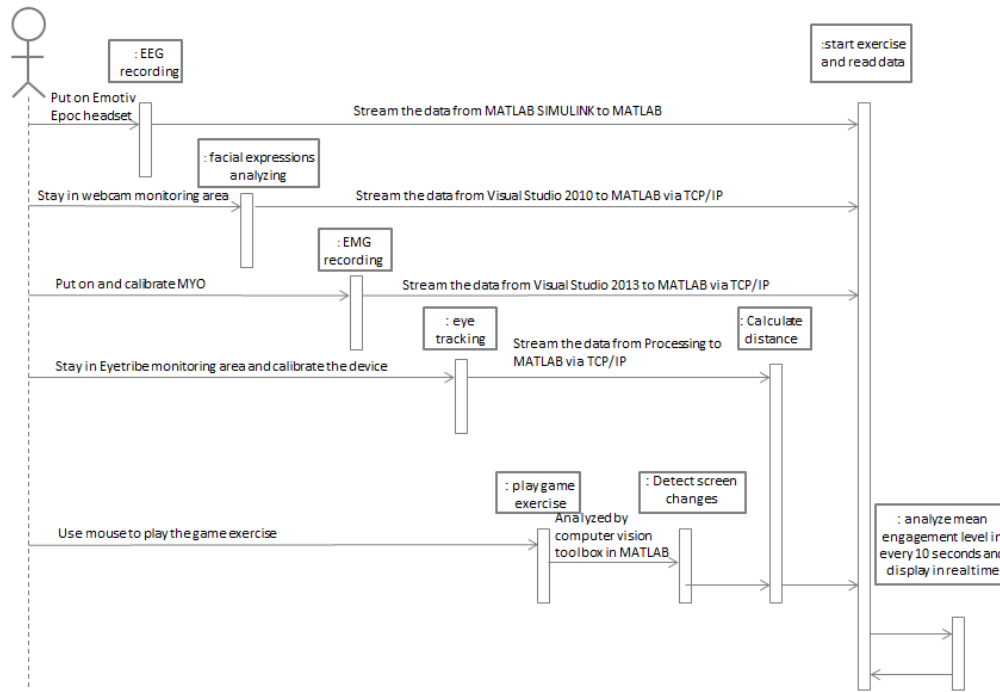
---

[1] https://www.myo.com/
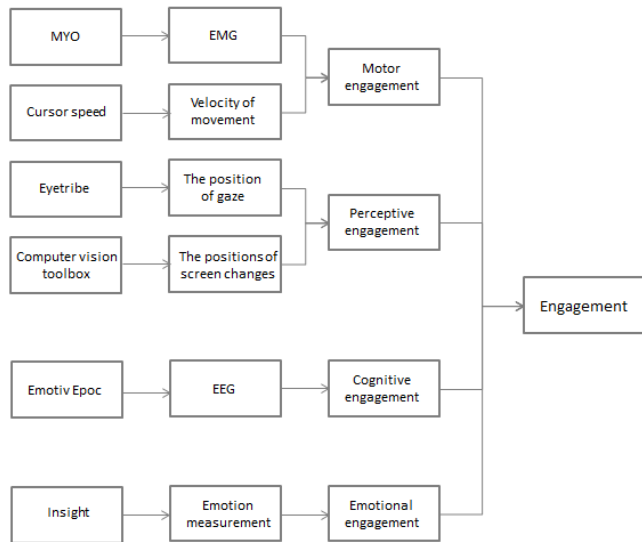
Fig. 4. Sequence diagram of EMS.

Fig. 5. Data sources and the monitored variables for evaluating engagement.

As for the velocity of movement, the moving speed of the cursor was applied as the velocity of movement, since the movement of the cursor reflected the patient's movement. The positions of the cursor were read in Processing together with the unix timestamp[2] of the system time. Then these data were streamed to MATLAB via TCP/IP. In MATLAB, the velocity of the movement was estimated by dividing the distance between two cursors by the time interval of 1 s.

2) Perceptive Engagement Monitoring: Perceptive engagement (E_p) is defined as sensory concentration, such as visual,

auditory and tactile. As for the visual aspect, eye tracking has been widely used as the method to evaluate user's concentration, and eye movement speed, eye movement total displacement, and validity of eye data have been used as the indicator to represent concentration [24]. However, we argue that when the user is interacting with the video game, the user's gaze will follow the content changes on the screen. Although the user may sometimes look at the display unit for an overall overview, for the most of the time, the focus of the patient's gaze should be on either the system cursor or the content change to interact with the video game. This happens because the content change usually demonstrates the tasks to the patient, and the system cursor indicates what the patient controls to complete the tasks. To complete the tasks, the gaze should be concentrated on either of these two locations on the screen. When the patient's attention is decaying, we argue that the patient will lose track of the content changes on the screen.

Since the patient should concentrate on the display unit and interact with the video game, the visual engagement is evaluated by analyzing patient's gaze (PG), the position of the system cursor (PSC), and the position of the content change (PCC) in the video game.

PG was monitored by the Eyetribe.[3] PCC was identified by comparing the screen contents at different times. PSC was read in Processing program. Let d_{v1} be the distance between PG and PCC, and d_{v2} the distance between the PG and PSC. The system evaluated the visual engagement by comparing the sum of d_{v1} and d_{v2}.

$$E_p = d_{v1} + d_{v2} \qquad (2)$$

In this case, the cursor was always one of the content changes in the screen. Therefore, we can only use d_{v1} to

---

[2]Unix time is a system for describing instants in time, defined as the number of seconds that have elapsed since 00:00:00 Coordinated Universal Time (UTC), Thursday, 1 January 1970, not counting lea seconds.

[3]http://theeyetribe.com/

represent the perceptive engagement. The Eyetribe was used to monitor the position of gaze on the display unit. It relies on infrared illumination and uses advanced mathematical models to determine the point of gaze. The sample rate was 30 Hz. We used Processing to read the eye tracking data and PSC. The data also contained system time in the unit of unix timestamp. These data were streamed to MATLAB via TCP/IP.

In MATLAB, example of Motion-Based Multiple Object Tracking[4] was used to detect the moving objects in a video. The input video was made by capturing the area of the game on the screen using Screen Capture tool.[5] The unix timestamp of the system time was also read in MATLAB with screen changes. Then the two timestamps read from MATLAB and Processing were used to synchronize these two signals so that d_{v1} can be calculated.

3) Cognitive Engagement Monitoring: Cognitive engagement (E_c) is considered to be proportional to the level of concentration performing cognitive tasks. Previous research has offered the following formula for calculating a signal, E_c, based on the α, β and θ waves that are highly correlated with participant's cognitive engagement [25]. This formula has been widely used by researchers in analyzing task engagement [15], [26]–[28].

$$E_c = \frac{\beta}{\alpha + \theta} \qquad (3)$$

Emotiv Epoc,[6] a headset with 14 nodes, was used to monitor the EEG signal. The EEG measurements were gathered from AF_3, AF_4, F_3, F_4, FC_5, and FC_6 on the frontal lobe which is known to manage attention, mental states and motor planning. The sampling rate was 128 Hz. Similar to MYO, Visual Studio 2010 was used to read the data from the headset and streamed and stored in MATLAB. We adopted the processing method from Freeman's study [28]. The power spectrum was calculated using a fast Fourier transformation. Bandwidth powers of α (8–12 Hz), β (13–30 Hz), and θ (4–7 Hz) were calculated by combining the bin powers in these three bandwidths. Then bandwidth powers were divided by total power (0–80 Hz) to produce percent power. E_c was first computed over a 20-s period and then updated every 2 s using a moving 20-s window.

4) Emotional Engagement Monitoring: Research has shown that positive emotion is associated with gains in functional status post-stroke [29]. Therefore, the goal of the game exercise is to arouse the positive emotion of the patient. Emotional engagement (E_e) is defined as emotion involvement during the exercise. If the exercise can influence the patient's emotion, it means that the patient is emotionally engaged in the exercise. In our system, the emotion is monitored by Insight, which is one software that can measure the user's emotion based on facial expression and represent it into seven categories, namely, neutral, happiness, surprise, anger, disgust, fear, and sadness. If the patient is emotionally engaged, the dominant

---
[4]http://nl.mathworks.com/help/vision/examples/motion-based-multiple-object-tracking.html
[5]http://nl.mathworks.com/matlabcentral/fileexchange/24323-screencapture-get-a-screen-capture-of-a-figure-frame-or-component
[6]https://emotiv.com/

TABLE I
PROPERTIES OF TWO COMPUTERS

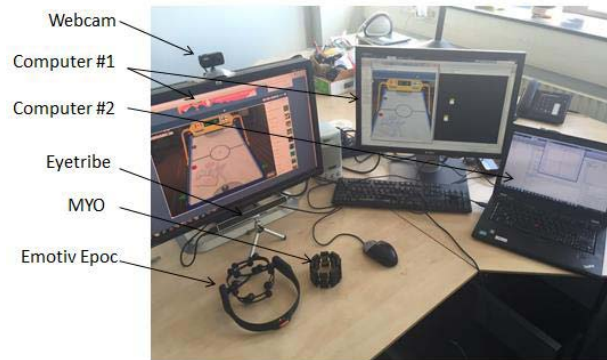| Computer	Operation system	Processor	RAM	System type	Monitor type
Computer #1	Windows 7 Enterprise	Inter(R) Core(TM) i7-2770 CPU @3.40GHz	16GB	64-bit	HP LP2475w
Computer #2	Windows 7 Ultimate	Inter(R) Core(TM) i5-2520M CPU @2.50GHz	3GB	64-bit	Thinkpad Display

Fig. 6. Overview of the prototype.

emotion will change due to different event in the game exercise. Therefore, we can define the indicator for emotional engagement is the ratio between time duration when positive emotion is dominant and the time duration when negative emotion is dominant.

$$E_e = T_{positive} / T_{negative} \qquad (4)$$

Insight was run in Visual Studio 2010. So we used the same method as with MYO to send the results of Insight to MATLAB via TCP/IP.

C. Test Setup of the EMS

Due to the computation capacity of a regular computer, two computers were used for implementing the prototype. Basic information of these two computers was shown in Table I. As you can see in Fig.6, computer #1 ran the game, and Eyetribe, and MATLAB in its background. We selected several games from the Internet, with focuses on motor training and cognitive training. All the games were played with mouse in computer #1. Chrome was used for running the game. MATLAB interpreted the perceptive engagement level based on the measurement of the gaze from Processing and the measurement of screen changes.

Computer #2 ran MATLAB, MYO, Insight, and EmotivEpoc. The data from these three sources were streamed to MATLAB via TCP/IP from Visual Studio. The program in MATLAB interpreted these data as motor engagement, cognitive engagement, and emotional engagement. Together with the data sent from MATLAB in computer #1 via TCP/IP,
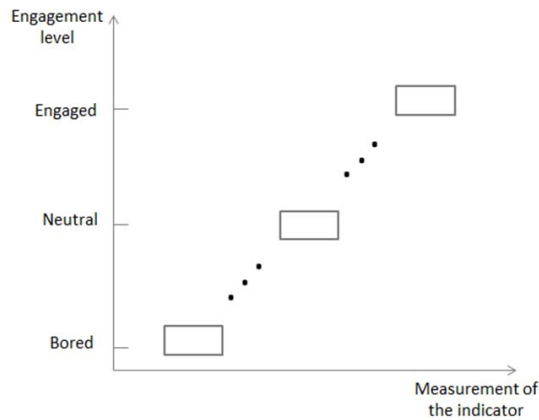
Fig. 7. Overview of the prototype.

Fig. 8. The picture for the motor exercises displayed on the screen.

the MATLAB program in computer #2 were able to represent all the four levels of engagement in real time.

IV. VALIDATION OF THE EMS

A. Objective

In this pre-medical experiments, the goal was to characterize the range and accuracy of the engagement indicators by influencing the subjects into different engaged states. The reliability of the indicators was an important issue. We expected that healthy subjects may provide a more consistent sample set, which allows us to see if the selected indicators of engagement can be used. Five healthy subjects were recruited in four experiments, which were designed and conducted to individually validate system modules of motor, perceptive, cognitive and emotional engagement monitoring. Different setups were created to mimic the situations in which the subject was in engaged, unengaged, or neutral engagement state. The prediction of the measurement in each setup is demonstrated in Fig. 7. Our assumption was the measurement of the indicator for the motor and cognitive engagement would be higher and the measurement of the indicator for perceptive engagement would be lower in the engaged state than that in the neutral and the bored state. Due to the small sampling, Friedman Test, non-parametric method of analysis, was applied in each experiment to see if the results were consistent for all the subjects.

B. Experiment to Validate Indicator of Motor Engagement

1) Setup of the Experiment: 5 subjects were recruited in this experiment, age from 26 to 50, two men and three women. According to the definition, when the subjects were making the movement with attention and effort, the level of their motor engagement was supposed to be higher. Therefore, in the setup mimicking the engaged state, the subjects were required to follow a strict accuracy requirement that needed the subject's attention. While in the neutral state, the subjects were required to track the same circle without the accuracy requirement.

Whereas in the created unengaged state, instead of making active movement, the subjects were passively moved by the experimenter. We also investigated the influence of velocity of
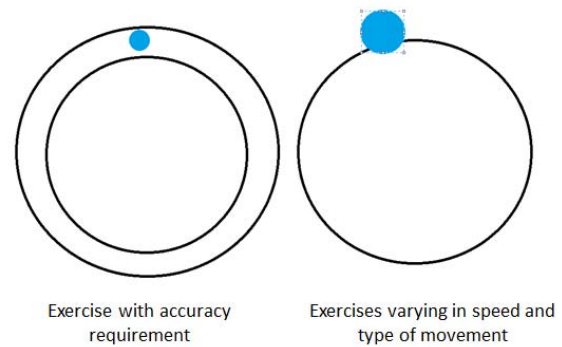

Fig. 9. The equipment used in the passive exercises.

the movement by requiring the subjects to move at different speeds.

In this experiment, the exercise was displayed on the screen for the subjects (Fig. 8). All the five subjects were required to use their right arm to do different exercises, in which they controlled the computer mouse to drag the object (the small solid circle) to track the given circle. During each exercise, MYO armband was put on the upper arm of the subject to measure the EMG signals from the bicep. There were three variables in different exercises. The first one was accuracy requirement. In the exercise requiring accuracy, the subject had to move the solid circle in the area between two given circles. The second one was speed requirement. The subject was required to move at three speeds respectively. At high speed, the subject had to complete each circle in about 3 s, while at medium and low speed, the time of completing a circle was 5 s and 10 s, respectively. The speed was controlled by the subjects. The third variable was the type of movement, active or passive. In the active movement, the subject moved voluntarily, while in passive movement, the subject was moved by the experimenter to move using the equipment shown in Fig. 9. The subject put their right hand on one of the handles, and the experimenter grasped the other handle to drive the subjects to complete the exercise. The mouse in the middle of the equipment was also controlled by the experimenter by another hand.

In the exercise with accuracy constraints, the subject was required to move the solid circle in the area between two given circles without speed requirement. In the second to
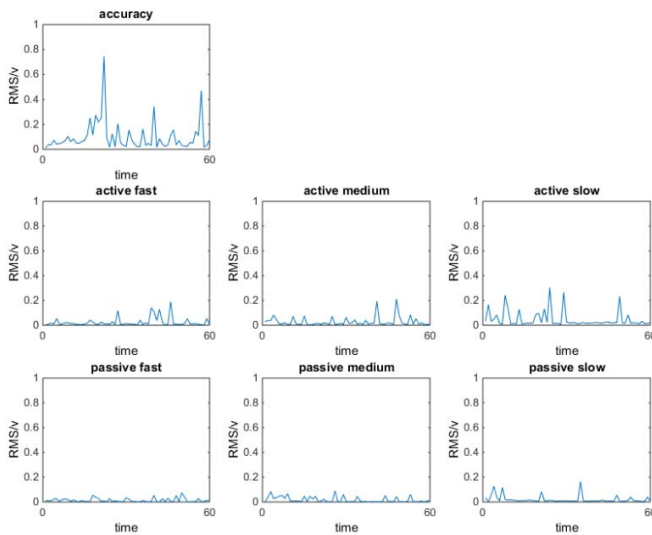
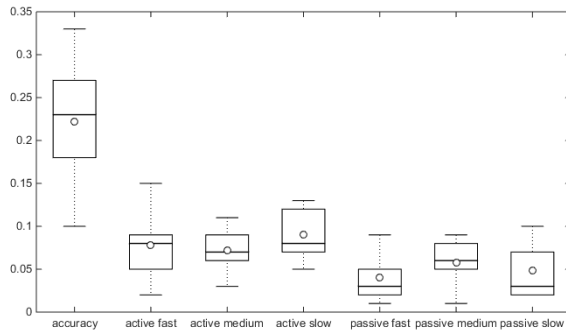Fig. 10. Typical pattern of the indicator in different exercises.

Fig. 11. Boxplot of the measurement of all the subjects in different setups.

fourth exercise, the subject was required to move the solid circle to track the big circle at high speed, medium speed and low speed respectively but without accuracy constraints. In the fifth to seventh exercise, the experimenter drove the subject to track a big circle at three speeds without accuracy requirement. Each exercise lasted 1 min. The order of the exercise was random for different subjects.

2) Results: The mean of RMS of EMG and the velocity of the movement were calculated in every second. The typical patterns of the indicator of motor engagement, that is the RMS of the EMG signal divided by velocity, in all the seven exercises are shown in Fig. 10. Boxplot of the measurement of all the subjects in different setups was shown in Fig. 11.

In the box, the central line is the median, the circle is the mean, and the edges of the box are the 25th and 75th percentiles.

A Friedman Test was conducted to test for differences between different setups. There was a statistically significant difference in the measured indicator depending on different setups, $\chi 2\ (2) = 19.736$, $p = 0.003$. Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied. Due to the reason that Bonferroni corrections are overly conservative, only one comparison between
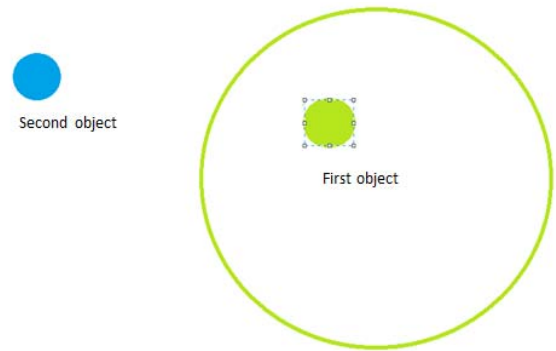

Fig. 12. The picture for the perceptive exercise.

the active exercise with accuracy requirement and active exercise with a high speed. Post hoc analysis showed there was a significant difference in the measurement between these two exercises ($Z = -2.023$, $p = 0.043$). Median (IQR) motor engagement for the exercise with accuracy requirement, the exercise with active movement at a high speed, the exercise with active movement at a medium speed, the exercise with active movement at a low speed, the exercise with passive movement at a high speed, the exercise with passive movement at a medium speed, the exercise with passive movement at a high speed, were 0.23 (0.14 to 0.3), 0.08 (0.035 to 0.12), 0.07 (0.045 to 0.1), 0.08 (0.06 to 0. 125), 0.03 (0.015 to 0.07), 0.06 (0.03 to 0.085), and 0.03 (0.02 to 0.085), respectively.

C. Experiment to Validate Indicator of Perceptive Engagement

1) Setup of the Experiment: Five subjects were recruited in this experiment, age from 24 to 50, three men and two women. According to the definition of the perceptive engagement, when the subject was following the content changes on the screen, he/she was regarded to be engaged perceptively. Therefore, in the simulated engaged state, the subject was required to follow the content changes all the time. While in the created neutral and unengaged states, the subject was asked to get fixated on non-changed part of the screen for different duration.

In this experiment, the picture for the exercise was shown on the screen (Fig. 12). During each exercise, the first object was always moving in the given circle on the screen. The subjects were required to look at the first object and follow its location until the second object appeared on the screen. Then the subjects were asked to look at the second object until it disappeared from the screen. The second object stayed at the same position during its appearance. A controlled time of appearance was used as independent variable of this experiment. When the second object appeared for a short time (i.e., 1 s) it was regarded as a screen change by comparing the two frames of the screen in MATLAB. When it appeared on the screen for longer time (e.g., more than 3 s), it was not considered as a screen change, but as a loss of attention. The gaze of the subjects was monitored by Eyetribe, which was put in front of the subject. The Eyetribe was calibrated for the subject before the exercise.
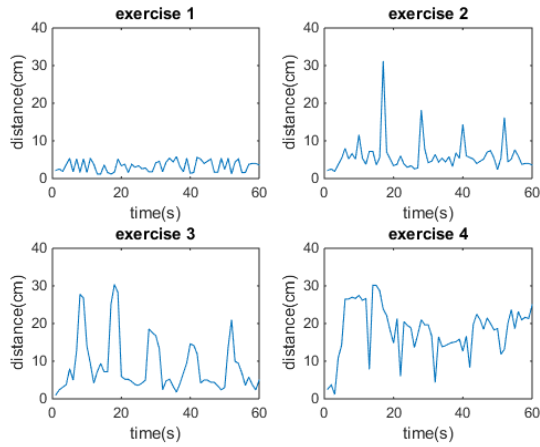
Fig. 13. Typical patterns of the distance in different exercises.

Fig. 14. Boxplot of the measurement of all subjects in different setups.

Fig. 15. Corsi block tasks.

There were four exercises in this experiment. Each exercise lasted 1 min. The order of the exercise was random for different subjects. In the first exercise, the second object stayed on the display for 1 s, and the interval between its appearances was 9 s. In the second exercise, the second object stayed on the display for 3 s, and the interval was 7 s. In the third and fourth exercise, the second object stayed on the screen for 5 s and 9 s, respectively, and the interval was 5 s and 1 s, respectively. Therefore, in the first exercise, the subjects had full perceptive engagement, and the indicator for perceptive engagement was expected to increase in the other exercises.

2) Results: The mean of the distance between the subject's gaze and the screen changes was analyzed in every second. Typical patterns of the indicator in different exercises were shown in Fig. 13.

In the typical patterns, we can see that in the first exercise, since the subject was perceptively engaged all the time, the distance was between 1 cm and 6 cm in the whole exercise. While in the second setup, when the second object stayed longer at the same position on the screen, the subject got fixated on the non-changed content. As the first object moved while the subjects were fixated, the distance between the focus of the subjects' gaze and the location of the first object has increased in the corresponding time period. As the second object appeared longer in the third and fourth exercises, the distance remained high for a longer duration.

Boxplot of the measurement of all the subjects in different setups was shown in Fig. 14. In the box, the central line is the median, the circle is the mean, and the edges of the box are the 25th and 75th percentiles.

A Friedman Test showed that there was a statistically significant difference in the measured indicator depending on different setups, $\chi 2$ (3) = 15.00, p = 0.002. Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied. Similar as in the analysis above, only one comparison between exercise 1 and exercise 2 was conducted. Result showed a significant difference in the perceptive engagement between these two exercises (Z = 2.023, p = 0.043). Median (IQR) perceptive engagement for exercise 1, exercise 2, exercise 3, exercise 4, were
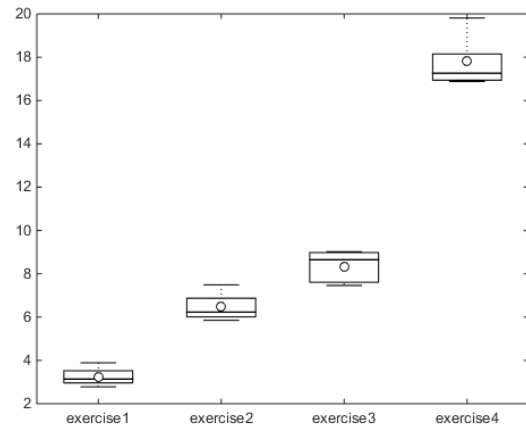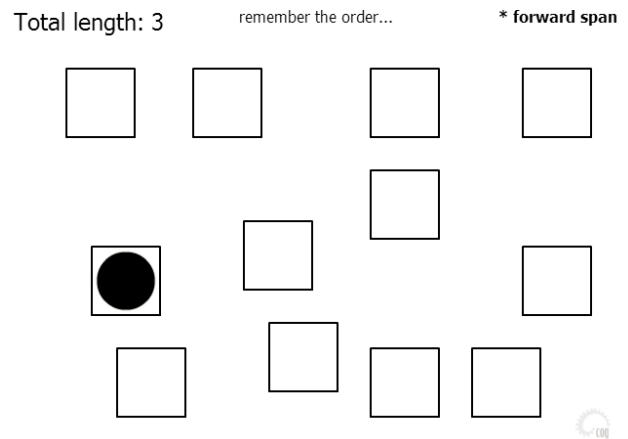
3.14 (2.87–3.71), 6.23 (5.93–7.18), 8.65 (7.54–9.00), and 16.91 (17.26–18.98), respectively. Separate Wilcoxon signed-rank tests were conducted as the Post Hoc Tests. However, no significant differences were found between different setups.

D. Experiment to Validate Indicator of Cognitive Engagement

1) Setup of the Experiment: Five subjects were recruited in this experiment, age from 14 to 42, three men and two women. Our assumption was that when the subject is doing the cognitive tasks intensively with effort, the indicator for the cognitive engagement is higher. Therefore, in the created cognitively engaged state, the subject was involved with challenging cognitive tasks. While in the created neutral and unengaged states, the subject was involved with less intensive and less challenging cognitive tasks.

In the experiment of cognitive engagement, all the five subjects were asked to play a cognitive game, Corsi block task (Fig. 15), in which the subjects were required to remember order of the blocks that got marked, then used the computer mouse to click them in the same order. During the exercises, EEG signals from six locations on the frontal lobe were measured by Emotiv Epoc headset. There were two independent
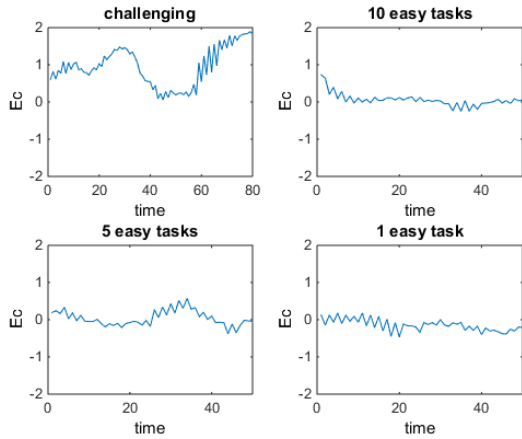
Fig. 16. Typical patterns of E_c.

Fig. 17. Boxplot of the measurement for all the subjects in different setups.

variables in different exercises. The first one was difficulty of the cognitive tasks. The second one was the intensity of the cognitive tasks.

In the first exercise, the subjects were required to complete 10 challenging tasks. During the first exercise, the subject can only play the next level with more blocks to remember, if he/she completed two tasks at the current difficulty level. The cognitive task was always challenging for the subject in this way. When the subject completed 10 tasks, the exercise stopped. While in the second exercise, the subject was required to play 10 tasks at the easiest level (three blocks to remember) during 2 min. In the third and fourth exercise, the subject was required to complete 5 and 1 task at the easiest level respectively during 2 min. After the subject completed the task in the second to the fourth exercise, in the rest of the time, the subject was asked to move the mouse between different blocks, but without doing any cognitive tasks. The aim of this requirement was to make the subject to make similar movement during each exercise to eliminate the influence of the movement on the measurement. The order of the exercises was random for different subjects.

2) Results: The results were analyzed from the 6 locations, AF_3, AF_4, F_3, F_4, FC_5, and FC_6. We found that there was a clear tendency in the data from AF3 to represent cognitive engagement of the subjects, but no clear tendency in the data from the other five locations. So the results analyzed from the data of AF_3 were presented below. Typical patterns of E_c were shown in Fig. 16. Boxplot of the measurement of all the subjects in different setups was shown in Fig. 17. In the box, the central line is the median, the circle is the mean, and the edges of the box are the 25th and 75th percentiles.

A Friedman Test showed that there was a statistically significant difference in the measured indicator depending on different setups, $\chi 2 (3) = 10.92$, $p = 0.012$. Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied. According to our hypotheses, two comparisons were conducted, between exercise 1 and exercise 2, and between exercise 2 and exercise 3, resulting in a significance level set at $p < 0.025$. However, no significant difference was found in these two pairs, between exercise 1 and exercise 2
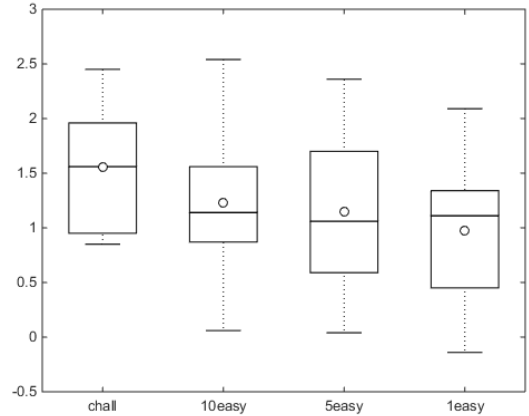
($Z = -1.214$, $p = 0.225$), between exercise 2 and exercise 3 ($Z = -1.214$, $p = 0.225$). Median (IQR) perceptive engagement for exercise 1, exercise 2, exercise 3, exercise 4, were 1.56 (0.90–2.21), 1.14 (0.47–2.05), 1.06 (0.32–2.03), and 1.11 (0.16–1.72), respectively.

E. Experiment in Emotional Aspect

1) Setup of the Experiment: 5 subjects were recruited in this experiment, age from 24 to 35, three men and two women. According to our definition of the emotional engagement, the subject is emotionally engaged when the rehabilitation exercise can influence the subject's emotion. The indicator for emotional engagement was the ratio of the time duration when positive emotion (happiness and surprise) was dominant with the time duration when negative emotion (anger, disgust, fearful, and sad) was dominant. The objective of this experiment was to validate if the emotion analyzed by Insight can reliably measure the subject's emotion. If the emotions analyzed by Insight can match the emotions indicated by the subject, it can be inferred that the indicator for emotional engagement will be measured correctly using Insight. To achieve the objective, the subjects were required to complete a questionnaire in which they had to reflect their emotions in the given moments during each exercise. Then the emotions reflected by the subjects were compared to the emotions measured by Insight.

During this experiment, each subject was required to play three online games (Fig. 18). In Air hockey, the subjects had to play against the computer. They were expected to use the hockey stick (the one below), which was controlled by the mouse, to shoot the hockey in the opposite goal. The subjects were required to play the first two components in this game. In the game Connect, the subjects had to connect the same two objects with the mouse and finish in 4 min. In the Cooking game, the subjects were required to move the mouse to cook kebab according to the game instructions. The order of these three games were random for different subjects. These games were controlled by the mouse. During each game, the subject's emotion was analyzed by Insight

After each game, the subject was required to fill in a questionnaire in Table II. The questionnaire of each exercise

TABLE II
QUESTIONNAIRE

Please fill in your emotion at each moment during the game you just played. You can choose one of the following emotions: neutral (😐), happy (😃), surprised (😮), anger (😠), disgust (🤢), fearful (😨), sad (😞), or other (name it).

| Air hockey | | Time | Connect | | Time | Cooking | | Time |
Moments	Emotion	(s)	Moments	Emotion	(s)	Moments	Emotion	(s)
Beginning	Neutral	0	Beginning	Neutral	0	Beginning	Neutral	0
Score first goal	Surprised	19	Finding pairs fast	Neutral	25	Slicing the meat	Neutral	150
Lose first goal	Anger	12	Struggling finding pairs	Anger	52	Finish slicing	Happy	230
Pass first level	Happy	54	Computer gives you a hint	Happy	86	Marinating the chicken	Bored	320
Score behind the component	Fearful	65	One minute left	Fearful	300	Finishing marinating the vegetables	Bored	340
Lose at a level (if any)	Anger	103	Finished the task in time (if you did)	No		Heating the kebab	Bored	470
End	Motivated to return	120	Cannot finish in time	Relaxed	360	End	Relaxed	570

Fig. 18. Online games.

was given to the subject separately after each exercise. In the questionnaire, the subject had to choose one of the emotions from a list of Emoji faces, including neutral (😐), happy (😃), surprised (😮), anger (😠), disgust (🤢), fearful (😨), sad (😞), or other (named by the subject),at each given moment during the exercise, while the time of each moment was filled in by the experimenter. During the exercise, the experimenter recorded the time at each moment in the questionnaire. After the exercise, the emotion analyzed by Insight at each moment was identified, which was then used to be compared with the emotion reflected by the subject at the same moment. The result of one subject was shown in Table II.

2) Results: The mean of emotion measured by Insight was calculated in every second. The dominant emotion in each second, which means the highest value of the seven emotions at that moment, was identified. The analyzed dominant emotions of the same subject who filled in the questionnaire above were shown in Fig. 19. Then these analyzed emotions were compared with the results of the questionnaires. We found that the overall accuracy of the emotions analyzed by Insight was 56%. However, the accuracy was 68% for four subjects, who seemed to express their emotions explicitly. For the other subject, although he indicated in the questionnaire that his internal emotion changed during the games, neutral emotion was the dominant emotion for 86% of the time. Moreover, Insight was found to be more accurate in identifying happiness,

with an accuracy of 82%, much higher than the negative emotions, such as anger, disgust, and fearful. It may because that the subjects tend to express positive emotion more easily and naturally than negative emotions.

Then the proportion of the emotion at each moment was analyzed. The most dominant three emotions at each moment in each were shown in Table III. We found that although the dominant emotion identified by Insight was not the same as the emotion indicated by the subject at some moments, the emotion indicated by the subject exist in the most dominant three emotions at some moments, such as "lose at a level" in Air hockey, "one-minute left" in Connect, and "finish slicing" in Cooking. Using the most dominant three emotions identified by Insight was found to improve the accuracy to 89% in matching the emotion indicated by the subjects. Therefore, the most dominant three emotions can be used in analyzing the indicator for emotional engagement. The indicator for emotional engagement can be refined as

$$E_e = \frac{T_{p1} * w_1 + T_{p2} * w_2 + T_{p3} * w_3}{T_{n1} * w_1 + T_{n2} * w_2 + T_{n3} * w_3}$$

(T_{p1}: time duration when positive emotion is the dominant emotion, T_{p2}: time duration when positive emotion is the second emotion, T_{p3}: time duration when positive emotion is the third emotion; T_{n1}: time duration when negative emotion is the dominant emotion, T_{n2}: time duration when negative emotion is the second emotion, T_{n3}: time duration when negative emotion is the third emotion; w_1, w_2, and w_3: weights are different for the top three emotions, here they are 0.5, 0.3, and 0.2, respectively.)

These results indicated that Insight was able to identify the subject's emotion reliably using the most dominant three emotions identified by Insight. Therefore, the refined indicator for emotional engagement can distinguish the situation when the subject is in a positive emotion and the situation when the subject is in a negative emotion.
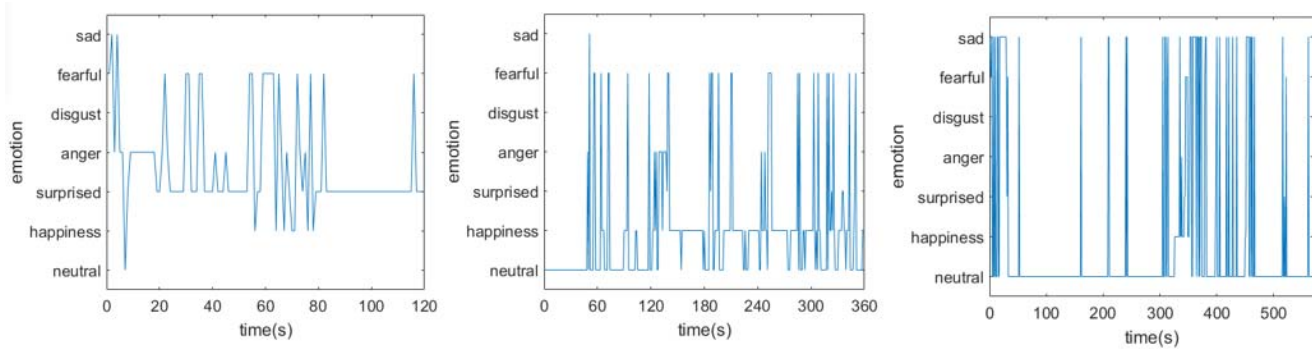
Fig. 19. Typical patterns of emotion analyzed by Insight in each game (From left to right: Air hockey, Connect, and Cooking).

TABLE III
PROPORTION OF THE COMPONENTS OF EMOTION AT EACH MOMENT

Air hockey						
Beginning	neutral	33%	fearful	62%	surprised	1%
Score first goal	neutral	2%	surprised	98%	others	0%
Lose first goal	neutral	25%	happy	49%	surprised	9%
Pass first level	neutral	3%	happy	94%	others	3%
Score behind the component	neutral	21%	surprised	36%	fearful	41%
Lose at a level	surprised	41%	happy	10%	anger	3%
End	surprised	39%	happy	27%	anger	6%
Connect						
Beginning	neutral	97%	surprised	2%	others	1%
Finding pairs fast	neutral	73%	surprised	13%	fearful	6%
Struggling finding pairs	neutral	25%	anger	33%	sad	19%
Computer gives you a hint	happy	89%	fearful	10%	others	1%
One minute left	happy	46%	surprised	14%	fearful	25%
Cannot finish in time	neutral	85%	happy	6%	fearful	6%
Cooking						
Beginning	neutral	87%	surprised	6%	sad	1%
Slicing the meat	neutral	93%	anger	4%	sad	3%
Finish slicing	neutral	52%	happy	26%	surprised	7%
Finish marinating the chicken	neutral	36%	happy	62%	others	2%
Finish marinating the vegetables	neutral	8%	fearful	1%	sad	88%
Heating the kebab	neutral	2%	anger	5%	sad	91%
End	neutral	13%	happy	41%	sad	37%

V. DISCUSSIONS

The results have shown that there were significant differences in the measured indicator between different setups in the experiments of the motor, perceptive, and cognitive aspects, which has achieved the objective of the experiments to test if the proposed indicators can reflect the expected results in different setups.

Regarding the motor engagement, it was found that there was a significant difference between the active exercise with accuracy requirement with the active exercise with a high speed. However, the results showed that there is only small difference between the active exercise without accuracy requirement and passive exercise. This indicated that movement with

accuracy requirement requires the subjects to pay much more attention and effort to complete the task. Active movement without accuracy requirement does not engage users on high level. Therefore, voluntary movement during rehabilitation training does not necessarily mean the patients are engaged. Other measures of the engagement in the perceptive, cognitive, and emotional aspects are also needed to have a comprehensive and accurate view on patients' engagement.

As for the experiment in the perceptive aspect, the results showed a significant difference in the perceptive engagement between the created engaged state and created neutral state, and there were also recognizable differences between other exercises. Experiments are still needed to be conducted to

measure the indicator of perceptive engagement with real patients in the future.

In the cognitive aspect, no significant difference was found between exercise 1 and exercise 2, and between exercise 2 and exercise 3, which may be caused by the small sample size. Also, no clear difference was found between the exercises with different cognitive loads. It may because after the subject completed the required cognitive tasks during the third and fourth exercise, the subject's attention was diverted when doing the mundane movement from block to another block, according to some subjects. This diverted attention can lead to subject's cognitive action which was represented by the measurement. Therefore, this may be the explanation that there was no difference in the measurement between the exercise with more intensive cognitive tasks and less cognitive tasks.

In the emotional aspect, the subjects may be subjective in answering the questionnaires. But their answer should reflect their emotion at the moment, at least the emotion they believe they were in. Since it is difficult, if not impossible, to detect the subject's internal emotion, we believe this is the best way to monitor the emotion in this context. However, the current results from the Insight was not accurate enough to measure the subject's emotion. In the future, maybe it is a good idea to profile the engaged and unengaged facial expressions based on the results from the Insight, and then use a machine learning technique to classify the facial expressions into different engaged states.

The primary reason of monitoring the engagement from motor, perceptive, cognitive, and emotional aspects is to provide a comprehensive understanding of the engaged states of the patients in the context of rehabilitation. Moreover, the ultimate goal of the CP-SRS is to enhance the patient's engagement during rehabilitation training. Based on monitoring the engagement, the SLM can make suggestions on a suitable stimulation strategy in order to enhance the engagement in the aspects which have decreased.

Although the experiment has shown promising results with these healthy subjects, it is too early to draw conclusion if the identified indicators are able to represent the real engagement level of real stroke patients. Problems may occur in a real environment with stroke patients. For instance, there might be abnormality in the EMG or EEG measurement due to spasm or lesions in the region of the cortex where the EEG is measured from. Experiments with real stroke patients are still needed to validate the accuracy of the refined indicator because post-stroke patients may have different facial expressions due to muscle paralysis. Moreover, a trade-off between cost and accuracy could be reached if the indicator is not reliable. For instance, if the overall accuracy of facial expression analysis by Insight with real patients is low, then the CP-SRS could function accurately without including the Insight.

VI. CONCLUSION

This paper proposes a cyber-physical stroke rehabilitation system (CP-SRS) aiming at enhancing the patient's engagement during rehabilitation training exercises. This paper focuses on introducing the implementation and validation of the EMS. The proposed EMS can read the data from four sources, namely, MYO, Eyetribe, Emotiv Epoc, and Insight, to analyze the subject's motor, perceptive, cognitive and emotional engagement during the exercise. Experiments were conducted to test if the identified indicators can reflect the expected results in different setups, which were created to mimic the situations in which the subject was in different engaged states. The results have shown that there were significant differences in the measured indicator between different setups in the experiments of the motor, perceptive, and cognitive aspects. Therefore, it is promising to use the identified indicators to evaluate the engaged states of the patients during rehabilitation. Then, in the next step, based on monitoring and evaluating the engagement level during the exercise, the effectiveness of the stimulation strategies in enhancing the engagement and the accuracy of the suggestions from the SLM on the most suitable stimulation strategies will be investigated.

REFERENCES

[1] G. B. Prange, M. J. Jannink, C. G. M. Groothuis-Oudshoorn, H. J. Hermens, and M. J. Ijzerman, "Systematic review of the effect of robot-aided therapy on recovery of the hemiparetic arm after stroke," *J. Rehabil. Res. Develop.*, vol. 43, no. 2, pp. 171–184, 2006.

[2] A. Henderson, N. Korner-Bitensky, and M. Levin, "Virtual reality in stroke rehabilitation: A systematic review of its effectiveness for upper limb motor recovery," *Topics Stroke Rehabil.*, vol. 14, no. 2, pp. 52–61, 2007.

[3] G. Kwakkel, B. J. Kollen, and H. I. Krebs, "Effects of robot-assisted therapy on upper limb recovery after stroke: A systematic review," *Neurorehabil. Neural Repair*, vol. 22, no. 2, pp. 111–121, 2008.

[4] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. New York, NY, USA: Harper Row, 1990.

[5] M. Beardsley, *The Aesthetic Point of View*. Ithaca, NY, USA: Cornell Univ. Press, 1982.

[6] W. Stephenson, "Play theory," in *The Play Theory of Mass Communication*. Chicago, IL, USA: Univ. Chicago Press, 1967, pp. 45–65.

[7] E. G. Toms, "Information interaction: Providing a framework for information architecture," *J. Am. Soc. Inf. Sci. Technol.*, vol. 53, no. 10, pp. 855–862, Aug. 2002.

[8] H. L. O'Brien and E. G. Toms, "What is user engagement? A conceptual framework for defining user engagement with technology," *J. Am. Soc. Inf. Sci. Technol.*, vol. 59, no. 6, pp. 938–955, Apr. 2008.

[9] A. H. Lequerica and K. Kortte, "Therapeutic engagement: A proposed model of engagement in medical rehabilitation," *Am. J. Phys. Med. Rehabil.*, vol. 89, no. 5, pp. 415–422, 2010.

[10] K. B. Kortte, L. D. Falk, R. C. Castillo, D. Johnson-Greene, and S. T. Wegener, "The Hopkins rehabilitation engagement rating scale: Development and psychometric properties," *Arch. Phys. Med. Rehabil.*, vol. 88, no. 7, pp. 877–884, 2007.

[11] N. Garcia, J. M. Sabater-Navarro, E. Gugliemeli, and A. Casals, "Trends in rehabilitation robotics," *Med. Biol. Eng. Comput.*, vol. 49, no. 10, pp. 1089–1091, 2011.

[12] S. D'Mello, P. Chipman, and A. C. Graesser, "Posture as a predictor of learner's affective engagement," in *Proc. 29th Annu. Conf. Cognit. Sci. Soc.*, Austin, TX, USA, 2007, pp. 905–910.

[13] J. Sanghvi *et al.*, "Automatic analysis of affective postures and body motion to detect engagement with a game companion," in *Proc. 6th ACM/IEEE Int. Conf. Human-Robot Interact.*, Mar. 2011, pp. 305–311.

[14] M. Cocea and S. Weibelzahl, "Cross-system validation of engagement prediction from log files," in *Creating New Learning Experiences on a Global Scale*. Berlin, Germany: Springer, 2007, pp. 14–25.

[15] D. Szafir and B. Mutlu, "Pay attention!: Designing adaptive agents that monitor and improve user engagement," in *Proc. SIGCHI Conf. Human Factors Comput. Syst*, 2012, pp. 11–20.

[16] J. W. Burke *et al.*, "Optimising engagement for stroke rehabilitation using serious games," *Vis. Comput.*, vol. 25, no. 12, pp. 1085–1099, 2009.

[17] R. C. V. Loureiro, M. J. Johnson, and W. S. Harwin, "Collaborative tele-rehabilitation: A strategy for increasing engagement," in *Proc. 1st IEEE/RAS-EMBS Int. Conf. Biomed. Robot. Biomechatron.*, Feb. 2006, pp. 859–864.

[18] L. Zimmerli, M. Jacky, L. Lünenburger, R. Riener, and M. Bolliger, "Increasing patient engagement during virtual reality-based motor rehabilitation," *Arch. Phys. Med. Rehabil.*, vol. 94, no. 9, pp. 1737–1746, Sep. 2013.

[19] C. Li, Z. Rusák, I. Horváth, and L. Ji, "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.*, vol. 37, no. 4, pp. 334–342, 2014.

[20] C. Li, Z. Rusák, I. Horváth, and L. Ji, "Development of engagement evaluation method and learning mechanism in an engagement enhancing rehabilitation system," *Eng. Appl. Artif. Intell.*, vol. 51, pp. 182–190, May 2016.

[21] C. Li, Z. Rusák, Y. Hou, C. Young, and L. Ji, "Upper limb motor rehabilitation integrated with video games focusing on training fingers' fine movements," *Int. J. Robot. Autom.*, vol. 29, no. 4, pp. 359–368, 2014.

[22] C. Li, Z. Rusák, I. Horváth, and L. Ji, "Current status of robotic stroke rehabilitation and opportunities for a cyber-physically assisted upper limb stroke rehabilitation," in *Proc. TMCE*, Budapest, Hungary, May 2014, pp. 899–914.

[23] F. J. Vera-Garcia, J. M. Moreside, and S. M. McGill, "MVC techniques to normalize trunk muscle EMG in healthy women," *J. Electromyogr. Kinesiol.*, vol. 20, no. 1, pp. 10–16, Feb. 2010.

[24] S. Li, and X. Zhang, "Eye-movement-based objective real-time quantification of patient's mental engagement in rehabilitation: A preliminary study," in *Proc. IEEE Int. Conf. Mechatron. Autom. (ICMA)*, Aug. 2014, pp. 180–185.

[25] A. T. Pope, E. H. Bogart, and D. S. Bartolome, "Biocybernetic system evaluates indices of operator engagement in automated task," *Biol. Psychol.*, vol. 40, nos. 1–2, pp. 187–195, May 1995.

[26] K. A. Brookhuis and D. De Waard, "The use of psychophysiology to assess driver status," *Ergonomics*, vol. 36, no. 9, pp. 1099–1110, 1993.

[27] P. J. Mikulka, M. W. Scerbo, and F. G. Freeman, "Effects of a biocybernetic system on vigilance performance," *Human Factors*, vol. 44, no. 4, pp. 654–664, 2002.

[28] F. G. Freeman, P. J. Mikulka, M. W. Scerbo, and L. Scott, "An evaluation of an adaptive automation system using a cognitive vigilance task," *Biol. Psychol.*, vol. 67, no. 3, pp. 283–297, Nov. 2004.

[29] G. V. Ostir, I.-M. Berges, M. E. Ottenbacher, A. Clow, and K. J. Ottenbacher, "Associations between positive emotion and recovery of functional status following stroke," *Psychosom. Med.*, vol. 70, no. 4, pp. 404–409, Oct. 2008.

Authors' biographies and photographs not available at the time of publication.